A Tour of Representation Theory

October 11, 2017

Martin Lorenz
2010 Mathematics Subject Classification. Primary 16Gxx, 16Txx, 17Bxx,
20Cxx, 20Gxx
For Maria
## Contents

Notations and Conventions xiii

**Part I. Algebras**

Chapter 1. Representations of Algebras 3
  1.1. Algebras 3
  1.2. Representations 24
  1.3. Primitive Ideals 42
  1.4. Semisimplicity 51
  1.5. Characters 65

Chapter 2. Further Topics on Algebras 79
  2.1. Projectives 79
  2.2. Frobenius and Symmetric Algebras 96

**Part II. Groups**

Chapter 3. Groups and Group Algebras 113
  3.1. Generalities 113
  3.2. First Examples 124
  3.3. More Structure 131
  3.4. Semisimple Group Algebras 143
  3.5. Further Examples 150
  3.6. Some Classical Theorems 159
  3.7. Characters, Symmetric Polynomials, and Invariant Theory 169
3.8. Decomposing Tensor Powers

Chapter 4. Symmetric Groups
4.1. Gelfand-Zetlin Algebras
4.2. The Branching Graph
4.3. The Young Graph
4.4. Proof of the Graph Isomorphism Theorem
4.5. The Irreducible Representations
4.6. The Murnaghan-Nakayama Rule
4.7. Schur-Weyl Duality

Part III. Lie Algebras

Chapter 5. Lie Algebras and Enveloping Algebras
5.1. Lie Algebra Basics
5.2. Types of Lie Algebras
5.3. Three Theorems about Linear Lie Algebras
5.4. Enveloping Algebras
5.5. Generalities on Representations of Lie Algebras
5.6. The Nullstellensatz for Enveloping Algebras
5.7. Representations of $\mathfrak{sl}_2$

Chapter 6. Semisimple Lie Algebras
6.1. Characterizations of Semisimplicity
6.2. Complete Reducibility
6.3. Cartan Subalgebras and the Root Space Decomposition
6.4. The Classical Lie Algebras

Chapter 7. Root Systems
7.1. Abstract Root Systems
7.2. Bases of a Root System
7.3. Classification
7.4. Lattices Associated to a Root System

Chapter 8. Representations of Semisimple Lie Algebras
8.1. Reminders
8.2. Finite-Dimensional Representations
8.3. Highest Weight Representations
8.4. Finite-Dimensional Irreducible Representations
## Contents

8.5. The Representation Ring 388  
8.6. The Center of the Enveloping Algebra 391  
8.7. Weyl’s Character Formula 406  
8.8. Schur Functors and Representations of $sl(V)$ 416

### Part IV. Hopf Algebras

Chapter 9. Coalgebras, Bialgebras and Hopf Algebras 425  
  9.1. Coalgebras 425  
  9.2. Comodules 438  
  9.3. Bialgebras and Hopf Algebras 444

Chapter 10. Representations and Actions 461  
  10.1. Representations of Hopf Algebras 462  
  10.2. First Applications 472  
  10.3. The Representation Ring of a Hopf Algebra 481  
  10.4. Actions and Coactions of Hopf Algebras on Algebras 488

Chapter 11. Affine Algebraic Groups 499  
  11.1. Affine Group Schemes 499  
  11.2. Affine Algebraic Groups 504  
  11.3. Representations and Actions 508  
  11.4. Linearity 511  
  11.5. Irreducibility and Connectedness 516  
  11.6. The Lie Algebra of an Affine Algebraic Group 522  
  11.7. Algebraic Group Actions on Prime Spectra 526

Chapter 12. Finite-Dimensional Hopf Algebras 537  
  12.1. Frobenius Structure 537  
  12.2. The Antipode 545  
  12.3. Semisimplicity 547  
  12.4. Divisibility Theorems 554  
  12.5. Frobenius-Schur Indicators 561

Appendix A. The Language of Categories and Functors 567  
  A.1. Categories 567  
  A.2. Functors 570  
  A.3. Naturality 571  
  A.4. Adjointness 575
Appendix B. Background from Linear Algebra 579
   B.1. Tensor Products 579
   B.2. Hom-⊗ Relations 585
   B.3. Vector Spaces 586

Appendix C. Some Commutative Algebra 591
   C.1. The Nullstellensatz 591
   C.2. The Generic Flatness Lemma 593
   C.3. The Zariski Topology on a Vector Space 593

Appendix D. The Diamond Lemma 597
   D.1. The Goal 597
   D.2. The Method 598
   D.3. First Applications 600
   D.4. A Simplification 602
   D.5. The Poincaré-Birkhoff-Witt Theorem 603

Appendix E. The Symmetric Ring of Quotients 607
   E.1. Definition and Basic Properties 607
   E.2. The Extended Center 609
   E.3. Comparison with Other Rings of Quotients 611

Bibliography 613

Index 625
Notations and Conventions

Functions and actions will generally be written on the left. In particular, all modules are left modules unless stated otherwise. Rings need not be commutative, but every ring \( R \) is understood to have an identity element, denoted by \( 1_R \) or simply \( 1 \), and ring homomorphisms \( f : R \to S \) are assumed to be unital, that is, \( f(1_R) = 1_S \). Throughout, we work over a commutative base field \( \mathbb{k} \). Any specific assumptions on \( \mathbb{k} \) will be explicitly stated, usually at the beginning of a chapter, and occasionally \( \mathbb{k} \) will be explicitly allowed to be a commutative ring.

Here is a list of the main abbreviations and symbols used in the text.

**General**

- \( [\cdot] \) disjoint union of sets
- \( X^I \) the set of functions \( I \to X \)
- \( \#X \) the cardinality of \( X \) if \( X \) is finite; otherwise \( \infty \)
- \( A^{(I)} \) the subset of \( A^I \) consisting of all finitely supported functions: \( f(i) = 0 \) for all but finitely many \( i \in I \) (\( A \) an abelian group)
- \( \mathbb{Z}_+, \mathbb{R}_+, \ldots \) the non-negative integers, reals, \ldots
- \( \mathbb{N} \) the natural numbers, \( \{1, 2, \ldots\} \)
- \( \mathbb{F}_q \) the field with \( q \) elements
- \( \mathbb{k} \) the base field
- \( \mu_N \) the group of \( N^{\text{th}} \) roots of unity
- \( [n] \) the set \( \{1, 2, \ldots, n\} \) for \( n \in \mathbb{N} \)
- \( \mathcal{P}_n \) the set of partitions of \( n \)
- \( \lambda \vdash n \) \( \lambda \) is a partition of \( n \)
Categories

Names of categories are generally chosen so as to be self-explanatory: Sets for the category of sets, Groups for the category of groups, AbGroups for abelian groups, etc. Here are some frequently occurring categories.

\[
\begin{align*}
\text{Vect}_k, \text{Vect}_k^\Delta & \quad \text{k-vector spaces, } \Delta\text{-graded } k\text{-vector spaces for a monoid } \Delta \\
\text{Alg}_k, \text{Alg}_k^\Delta & \quad \text{k-algebras, } \Delta\text{-graded } k\text{-algebras} \\
\text{CommAlg}_k & \quad \text{commutative } k\text{-algebras} \\
A\text{Mod}, \text{Mod}_A & \quad \text{left, right } A\text{-modules for } A \in \text{Alg}_k \\
A\text{Mod}_B & \quad (A, B)\text{-bimodules for } A, B \in \text{Alg}_k \\
\text{Rep } A, \text{Rep}_{\text{fin}} A & \quad \text{representations, finite-dimensional representations of } A \\
A\text{Proj} & \quad - \\
A\text{proj} & \quad - \\
\text{Proj}_{\text{fin}} A & \quad - \\
\end{align*}
\]

Vector spaces

\[
\begin{align*}
 V^* & = \text{Hom}_k(V, k) \quad \text{the dual space of the } k\text{-vector space } V \\
 V^\oplus I & \quad \text{the } I\text{-fold direct sum of } V \\
 V^\otimes n & \quad \text{the } n\text{th tensor power of } V \\
\langle \cdot, \cdot, \cdot \rangle & \quad \text{the evaluation form } V^* \times V \to k \\
T V, T^n V & \quad \text{the tensor algebra of } V \text{ and its } n\text{th component, } T^n V = V^\otimes n \\
\text{Sym } V, \text{Sym}^n V & \quad \text{the symmetric algebra of } V \text{ and its } n\text{th component} \\
\Lambda V, \Lambda^n V & \quad \text{the exterior algebra of } V \text{ and its } n\text{th component} \\
O(V) = \text{Sym}(V^*) & \quad \text{the algebra of polynomial functions on } V \\
V_\lambda, V_\phi & \quad \text{eigenspace, space of semi-invariants, weight space} \\
\mathbb{k}X \text{ or } \mathbb{k}[X] & \quad \text{the vector space of all formal } \mathbb{k}\text{-linear combinations of a set } X \\
\text{GL}(V) & \quad \text{the group of invertible linear endomorphisms of } V
\end{align*}
\]

Algebras

\[
\begin{align*}
\mathbb{k}\langle X \rangle & \quad \text{the free } \mathbb{k}\text{-algebra generated by the set } X \\
A^{\text{op}} & \quad \text{the opposite algebra of an algebra } A \\
A^\times & \quad \text{the group of units (invertible elements) of } A \\
C_A(X) & \quad \{a \in A \mid ax = xa \text{ for all } x \in X\}, \text{ the centralizer of } X \subseteq A \\
\mathcal{Z}A \text{ or } \mathcal{Z}(A) & \quad C_A(A), \text{ the center of the algebra } A \\
\text{Mat}_n(A) & \quad \text{the algebra of } n \times n\text{-matrices over } A \\
\text{Irr } A & \quad \text{the set of equivalence classes of irreducible representations of } A; \\
& \quad \text{alternatively, a full representative set of irreducible representations}
\end{align*}
\]
MaxSpec $A$ the set of maximal ideals of $A$
Spec $A$ the prime ideals of $A$
Prim $A$ the primitive ideals of $A$
rad $A$

Groups

$kG$ or $k[G]$ the group algebra of the group $G$ over $k$
$C_n$ the cyclic group of order $n$
$D_n$ the dihedral group of order $2n$
$S_n$ the symmetric group on $\{1, 2, \ldots, n\}$
$A_n$ the alternating subgroup of $S_n$
$GL_n(A)$ Mat$_{n}(A)^{\times}$, the group of invertible $n \times n$-matrices over $A \in \text{Alg}_k$
$G \subseteq X$ short for $G \times X \rightarrow X$, a left action of $G$ on the set $X$
$G \backslash X$ the set of orbits for an action $G \subseteq X$ or, alternatively, a transversal for these orbits
$G/H$ the collection of all left cosets $gH$ ($g \in G$) of a subgroup $H \leq G$; alternatively, a transversal for the left cosets
$G_x$ the conjugacy class of an element $x \in G$

Symmetric Groups

$s_i$ the transposition $(i, i + 1) \in S_n$
Irr $S_n$ Irr $kS_n$
$GZ_n$ the Gelfand-Zetlin subalgebra of $kS_n$
$X_1, \ldots, X_n$ the Jucys-Murphy elements of $kS_n$
$d_n = \dim_k GZ_n$ also the sum of all $\dim_k V$ with $V \in \text{Irr } S_n$
$\mathcal{Z}_n = \mathcal{Z}(kS_n)$ the center of the group algebra $kS_n$
$V^A$ and $V^{\lambda/\mu}$ the irreducible representation of $S_n$ corresponding to the partition $\lambda \vdash n$ and the representation of $S_n$ given by the skew shape $\lambda/\mu$ with $|\lambda/\mu| = n$
$f^A$ and $f^{A/\mu}$ the numbers of $\lambda$-tableaux and $\lambda/\mu$-tableaux; also $\dim_k V^A$ and $\dim_k V^{A/\mu}$
$\chi^A$ and $\chi^{A/\mu}$ the characters of $\dim_k V^A$ and $\dim_k V^{A/\mu}$
$S^W$, $S^A$ Schur functors

Lie Algebras
U_\mathfrak{g} or U(\mathfrak{g}) the enveloping algebra of \mathfrak{g}

\mathfrak{h}

\mathfrak{g}_{\alpha}

\rho the half-sum of the positive roots

\Phi

\Delta

\mathcal{W} = \mathcal{W}_{\Phi}

\mathcal{L} = \mathcal{L}_{\Phi} the root lattice

\Lambda = \Lambda_{\Phi} the weight lattice

M(\lambda) Verma module

V(\lambda) the unique irreducible image of M(\lambda)

w \cdot \lambda \quad w(\lambda + \rho) - \rho for w \in \mathcal{W}, \lambda \in \mathfrak{h}^*.

**Hopf Algebras**

\Delta the comultiplication

\varepsilon the counit

S the antipode

H^* the finite dual of the Hopf algebra H
Part I

Algebras
Chapter 1

Representations of Algebras

This chapter develops the basic themes of representation theory in the setting of algebras. We establish notation to be used throughout the remainder of the book and prove some fundamental results of representation theory such as Wedderburn's Structure Theorem. The focus will be on irreducible and completely reducible representations. The reader is referred to Appendices A and B for a brief introduction to the language of categories and for the requisite background material from linear algebra.

Throughout this chapter, $\mathbb{k}$ denotes an arbitrary field. The category of $\mathbb{k}$-vector spaces and $\mathbb{k}$-linear maps is denoted by $\text{Vect}_\mathbb{k}$ and $\otimes$ will stand for $\otimes_\mathbb{k}$.

1.1. Algebras

In the language of rings, a $\mathbb{k}$-algebra can be defined as a ring $A$ (with 1) together with a given ring homomorphism $\mathbb{k} \to A$ that has image in the center $\mathcal{Z}A$ of $A$. Below, we recast this definition in an equivalent form, starting over from scratch in the setting of $\mathbb{k}$-vector spaces. The basics laid out in the following apply to any commutative ring $\mathbb{k}$, working in $\mathbb{k}\text{Mod}$ (with the conventions of §B.1.3) rather than in $\text{Vect}_\mathbb{k}$. However, we will consider this more general setting only very occasionally, and hence $\mathbb{k}$ is understood to be a field.

1.1.1. The Category of $\mathbb{k}$-Algebras

A $\mathbb{k}$-algebra can equivalently be defined as a vector space $A \in \text{Vect}_\mathbb{k}$ that is equipped with two $\mathbb{k}$-linear maps, the multiplication $m = m_A: A \otimes A \to A$ and the unit
\[ u = u_A : \mathbb{k} \to A, \text{ such that the following diagrams commute:} \]

\[
\begin{array}{ccc}
A \otimes A \otimes A & \xrightarrow{m \otimes \text{Id}} & A \otimes A \\
\text{Id} \otimes m & \downarrow & m \\
A \otimes A & \xrightarrow{m} & A
\end{array}
\quad \text{and} \quad
\begin{array}{ccc}
A \otimes A & \xrightarrow{u \otimes \text{Id}} & A \otimes \mathbb{k} \\
\text{Id} \otimes u & \downarrow & m \\
A & \xrightarrow{u} & A \otimes \mathbb{k}
\end{array}
\]

Here, \( \text{Id} = \text{Id}_A \) denotes the identity map of \( A \). The isomorphism \( \mathbb{k} \otimes A \xrightarrow{\sim} A \) in (1.1) is the standard one, given by the scalar multiplication, \( \lambda \otimes a \mapsto \lambda a \) for \( \lambda \in \mathbb{k} \) and \( a \in A \); similarly for \( A \otimes \mathbb{k} \xrightarrow{\sim} A \). Multiplication will generally be written as juxtaposition: \( m(a \otimes b) = ab \) for \( a, b \in A \). Thus, \( ab \) depends \( \mathbb{k} \)-linearly on both \( a \) and \( b \). The algebra \( A \) is said to be commutative if \( ab = ba \) for all \( a, b \in A \); or, equivalently, \( m = m \circ \tau \) where \( \tau \in \text{End}_\mathbb{k}(A \otimes A) \) is given by \( \tau(a \otimes b) = b \otimes a \).

The first diagram in (1.1) amounts to the associative law: \( (ab)c = a(bc) \) for all \( a, b, c \in A \). The second diagram expresses the unit laws: \( u(1_A) a = a = a u(1_A) \) for all \( a \in A \); so \( A \) has the identity element \( u(1_A) = 1_A \). If \( u = 0 \), then it follows that \( A = \{0\} \); otherwise, the unit map \( u \) is injective and it is often notationally suppressed, viewing it as an inclusion \( \mathbb{k} \subseteq A \). Then \( 1_A = 1_A \), the scalar operation of \( \mathbb{k} \) on \( A \) becomes multiplication in \( A \), and \( \mathbb{k} \subseteq \mathfrak{Z}(A) \).

Given \( \mathbb{k} \)-algebras \( A \) and \( B \), a homomorphism from \( A \) to \( B \) is a \( \mathbb{k} \)-linear map \( f : A \to B \) that respects multiplications and units in the sense that the following diagrams commute:

\[
\begin{array}{ccc}
A \otimes A & \xrightarrow{f \otimes f} & B \otimes B \\
m_A & \downarrow & m_B \\
A & \xrightarrow{f} & B
\end{array}
\quad \text{and} \quad
\begin{array}{ccc}
A & \xrightarrow{u_B} & B \\
u_A & \downarrow & u_B \\
A & \xrightarrow{u} & A \otimes \mathbb{k}
\end{array}
\]

These diagrams are equivalent to the equations \( f(aa') = f(a)f(a') \) for all \( a, a' \in A \) and \( f(1_A) = 1_B \).

The category whose objects are the \( \mathbb{k} \)-algebras and whose morphisms are the homomorphisms between \( \mathbb{k} \)-algebras will be denoted by \( \text{Alg}_\mathbb{k} \).

Thus, \( \text{Hom}_{\text{Alg}_\mathbb{k}}(A, B) \) is the set of all \( \mathbb{k} \)-algebra homomorphisms \( f : A \to B \). Algebra homomorphisms are often simply called algebra maps. The variants isomorphism and monomorphism have the same meaning as in \( \text{Vect}_\mathbb{k} \): algebra homomorphisms that are bijective and injective, respectively; similarly for automorphism and endomorphism. \(^1\) A subalgebra of a given \( \mathbb{k} \)-algebra \( A \) is a \( \mathbb{k} \)-subspace \( B \) of \( A \) that is

\(^1\)Every surjective algebra map is an epimorphism in \( \text{Alg}_\mathbb{k} \) in the categorical sense, but the converse does not hold \([140, \text{Section I.5}]\).
a \( \mathbb{k} \)-algebra in its own right in such a way that the inclusion \( B \hookrightarrow A \) is an algebra map.

**Tensor Products of Algebras.** The tensor product of two algebras \( A, B \in \text{Alg}_k \) is obtained by endowing \( A \otimes B \in \text{Vect}_k \) with the multiplication

\[
(a \otimes b)(a' \otimes b') := aa' \otimes bb'
\]

for \( a, a' \in A \) and \( b, b' \in B \). It is easy to check that this multiplication is well-defined. Taking \( u_A \otimes u_B : k \equiv k \otimes k \to A \otimes B \) as unit map, the vector space \( A \otimes B \) turns into a \( k \)-algebra. Observe that the switch map \( a \otimes b \mapsto b \otimes a \) is an isomorphism \( A \otimes B \cong B \otimes A \) in \( \text{Alg}_k \). Exercise 1.1.10 spells out some functorial properties of this construction and explores some examples.

**Extending the Base Field.** A \( k \)-algebra that is a field is also called a \( k \)-field. For any \( A \in \text{Alg}_k \) and any \( k \)-field \( K \), we may regard \( K \otimes A \) as a \( k \)-algebra as in the preceding paragraph and also as a \( K \)-vector space as in §B.3.4. The multiplication \((K \otimes A) \otimes (K \otimes A) \to K \otimes A\) in (1.3), given by \((\lambda \otimes a)(\lambda' \otimes a') = \lambda \lambda' \otimes aa'\) for \( \lambda, \lambda' \in K \) and \( a, a' \in A \), passes down to a \( K \)-linear map

\[
(K \otimes A) \otimes_K (K \otimes A) \cong A \otimes (K \otimes_K K) \otimes A \cong K \otimes (A \otimes A) \longrightarrow K \otimes A,
\]

where the last map is \( K \otimes m_A \). With this multiplication and with

\[
K \otimes u_A : K \cong K \otimes k \to K \otimes A
\]

as unit map, \( K \otimes A \) becomes a \( K \)-algebra.

The above construction is functorial: any map \( f : A \to B \) in \( \text{Alg}_k \) gives rise to the map \( K \otimes f : K \otimes A \to K \otimes B \) in \( \text{Alg}_K \). Thus, the field extension functor \( K \otimes \cdot : \text{Vect}_k \to \text{Vect}_K \) of §B.3.4 restricts to a functor \( K \otimes \cdot : \text{Alg}_k \to \text{Alg}_K \).

1.1.2. Some Important Algebras

We now describe a selection of algebras that will play prominent roles later on in this book, taking the opportunity to mention some standard concepts from the theory of algebras and from category theory along the way.

**Endomorphism Algebras**

The archetypal algebra from the viewpoint of representation theory is the algebra \( \text{End}_k(V) \) of all \( k \)-linear endomorphisms of a vector space \( V \in \text{Vect}_k \). Multiplication in \( \text{End}_k(V) \) is given by composition of endomorphisms and the unit map sends each \( \lambda \in k \) to the scalar transformation \( \lambda \text{Id}_V \). If \( \text{dim}_k V = n < \infty \), then any choice of \( k \)-basis for \( V \) gives rise to a \( k \)-linear isomorphism \( V \cong k^n \) and to an isomorphism of \( k \)-algebras \( \text{End}_k(V) \cong \text{Mat}_n(k) \), the \( n \times n \) matrix algebra over \( k \).
The matrix algebra $\text{Mat}_n(\mathbb{k})$ and the endomorphism algebra $\text{End}_\mathbb{k}(V)$ of a finite-dimensional vector space $V$ are examples of \textit{finite-dimensional} algebras, that is, algebras that are finite dimensional over the base field $\mathbb{k}$.

\textbf{Free and Tensor Algebras}

We will also on occasion work with the \textit{free $\mathbb{k}$-algebra} that is generated by a given set $X$; this algebra will be denoted by $\mathbb{k}\langle X \rangle$. One can think of $\mathbb{k}\langle X \rangle$ as a noncommutative polynomial algebra over $\mathbb{k}$ with the elements of $X$ as noncommuting variables. Assuming $X$ to be indexed, say $X = (x_i)_{i \in I}$, a $\mathbb{k}$-basis of $\mathbb{k}\langle X \rangle$ is given by the collection of all finite products

$$x_{i_1} x_{i_2} \cdots x_{i_k},$$

where $(i_1, i_2, \ldots, i_k)$ is a finite (possibly empty) sequence of indices from $I$. These products are also called \textit{monomials} or \textit{words} in the alphabet $X$; the order of the symbols $x_{i_j}$ in words does matter. Multiplication in $\mathbb{k}\langle X \rangle$ is defined by concatenation. The empty word is the identity element $1 \in \mathbb{k}\langle X \rangle$.

Formally, $\mathbb{k}\langle X \rangle$ can be constructed as the tensor algebra $T(\mathbb{k}X)$, where $\mathbb{k}X$ is the $\mathbb{k}$-vector space of all formal $\mathbb{k}$-linear combinations of the elements of $X$ (Example A.5). Here, the \textit{tensor algebra} of an arbitrary vector space $V \in \text{Vect}_\mathbb{k}$ is defined as the direct sum

$$TV \overset{\text{def}}{=} \bigoplus_{k \in \mathbb{Z}_+} V^{\otimes k},$$

where $V^{\otimes k}$ is the $k$th tensor power of $V$ as in (B.10); so (B.9) gives

$$\dim_\mathbb{k} V^{\otimes k} = (\dim_\mathbb{k} V)^k.$$

The unit map of $TV$ is given by the canonical embedding $\mathbb{k} = V^{\otimes 0} \hookrightarrow TV$, and multiplication in $TV$ comes from the associativity isomorphisms (B.11) for tensor powers:

$$(v_1 \otimes \cdots \otimes v_k)(v'_1 \otimes \cdots \otimes v'_l) = v_1 \otimes \cdots \otimes v_k \otimes v'_1 \otimes \cdots \otimes v'_l$$

for $v_i, v'_j \in V$. This multiplication is distributively extended to define products of arbitrary elements of $TV$. In this way, $TV$ becomes a $\mathbb{k}$-algebra. Note that the subspace $V = V^{\otimes 1} \subseteq TV$ \textit{generates} the algebra $TV$ in the sense that the only $\mathbb{k}$-subalgebra of $TV$ containing $V$ is $TV$ itself. Equivalently, every element of $TV$ is a $\mathbb{k}$-linear combination of finite products with factors from $V$. In fact, any generating set of the vector space $V$ will serve as a set of generators for the algebra $TV$.

The importance of tensor algebras stems from their functorial properties, which we shall now explain in some detail. Associating to a given $\mathbb{k}$-vector space $V$ the $\mathbb{k}$-algebra $TV$, we obtain a functor

$$T : \text{Vect}_\mathbb{k} \longrightarrow \text{Alg}_\mathbb{k}.$$
As for morphisms, let \( f \in \text{Hom}_k(V, W) \) be a homomorphism of \( k \)-vector spaces. Then we have morphisms \( f^\otimes k \in \text{Hom}_k(V^\otimes k, W^\otimes k) \) for each \( k \in \mathbb{Z}_+ \) as in §B.1.3:

\[
f^\otimes k(v_1 \otimes \cdots \otimes v_k) = f(v_1) \otimes \cdots \otimes f(v_k).
\]

The \( k \)-linear map

\[
Tf \overset{\text{def}}{=} \bigoplus_{k \in \mathbb{Z}_+} f^\otimes k : TV = \bigoplus_{k \in \mathbb{Z}_+} V^\otimes k \rightarrow \bigoplus_{k \in \mathbb{Z}_+} W^\otimes k = TW
\]

is easily seen to be a \( k \)-algebra map and it is equally straightforward to check that \( T \) satisfies all requirements of a functor.

The property of the tensor algebra that is expressed in the following proposition is sometimes referred to as the universal property of the tensor algebra; it determines the tensor algebra up to isomorphism (Exercise 1.1.1).

**Proposition 1.1.** For \( V \in \text{Vect}_k \) and \( A \in \text{Alg}_k \), there is a natural bijection of sets

\[
\text{Hom}_{\text{Alg}_k}(TV, A) \cong \text{Hom}_k(V, A|_{\text{Vect}_k})
\]

Here, \( f|_V \) denotes the restriction of \( f \) to \( V = V^\otimes 1 \subseteq TV \). The notation \( A|_{\text{Vect}_k} \) indicates that the algebra \( A \) is viewed merely as a \( k \)-vector space, with all other algebra structure being ignored. We use the symbol \( \cong \) for an isomorphism in any category (Section A.1); in \( \text{Sets} \), this is a bijection. The bijection in Proposition 1.1 behaves well with respect to varying the input data, \( V \) and \( A \)—this is what “naturality” of the bijection is meant to convey. Technically, the functor \( T: \text{Vect}_k \rightarrow \text{Alg}_k \) and the forgetful functor \( \cdot|_{\text{Vect}_k}: \text{Alg}_k \rightarrow \text{Vect}_k \) are a pair of adjoint functors. The reader wishing to see the specifics spelled out is referred to Section A.4. We also mention that any two left adjoint functors of a given functor are naturally isomorphic [140, p. 85]; see also Exercise 1.1.1.

**Proof of Proposition 1.1.** The map in the proposition is injective, because \( V \) generates the algebra \( TV \). For surjectivity, let \( \phi \in \text{Hom}_k(V, A|_{\text{Vect}_k}) \) be given. Then the map \( V^k \rightarrow A, (v_1, v_2, \ldots, v_k) \mapsto \phi(v_1)\phi(v_2)\cdots\phi(v_k) \) is \( k \)-multilinear, and hence it gives rise to a unique \( k \)-linear map \( \phi^k: V^\otimes k \rightarrow A, v_1 \otimes \cdots \otimes v_k \mapsto \phi(v_1)\cdots\phi(v_k) \) by (B.13). The maps \( \phi^k \) yield a unique \( k \)-linear map \( f: TV \rightarrow A \) such that \( f|_V^{\otimes k} = \phi^k \) for all \( k \). In particular, \( f|_V = \phi \) as needed, and it is also immediate that \( f \) is in fact a \( k \)-algebra map. This establishes surjectivity. \( \square \)

**Grading.** Tensor algebras are examples of graded algebras, that is, algebras that are equipped with a meaningful notion of “degree” for their nonzero elements. In
A -algebra $A$ is said to be graded if

$$A = \bigoplus_{k \in \mathbb{Z}_+} A^k$$

for $\mathbb{k}$-subspaces $A^k \subseteq A$ such that $A^k A^{k'} \subseteq A^{k+k'}$ for all $k, k' \in \mathbb{Z}_+$. More precisely, such algebras are called $\mathbb{Z}_+$-graded; grading by monoids other than $(\mathbb{Z}_+, +)$ are also often considered. The nonzero elements of $A^k$ are called homogeneous of degree $k$.

An algebra map $f : A \rightarrow B$ between graded algebras $A$ and $B$ is called a homomorphism of graded algebras if $f$ respects the gradings in the sense that $f(A^k) \subseteq B^k$ for all degrees $k$. All this applies to the tensor algebra $T(V)$, with $V^\otimes k$ being the component of degree $k$. The algebra maps $Tf : TV \rightarrow TW$ constructed above are in fact homomorphisms of graded algebras. Numerous algebras that we shall encounter below carry a natural grading. See Exercise 1.1.11 for more background on gradings.

Returning to the case where $V = \mathbb{k}X$ is the vector space with basis $X = (x_i)_{i \in I}$, the $k^\text{th}$ tensor power $(\mathbb{k}X)^\otimes k$ has a basis given by the tensors

$$x_{i_1} \otimes x_{i_2} \otimes \cdots \otimes x_{i_k}$$

for all sequences of indices $(i_1, i_2, \ldots, i_k)$ of length $k$. Sending the above $k$-tensor to the corresponding word $x_{i_1} x_{i_2} \cdots x_{i_k}$, we obtain an isomorphism of $T(\mathbb{k}X)$ with the free algebra $\mathbb{k}\langle X \rangle$. The grading of $T(\mathbb{k}X)$ by the tensor powers $(\mathbb{k}X)^\otimes k$ makes $\mathbb{k}\langle X \rangle$ a graded algebra as well: the homogeneous component of degree $k$ is the $\mathbb{k}$-subspace of $\mathbb{k}(X)$ that is spanned by the words of length $k$. This grading is often referred to as the grading by “total degree.” Proposition 1.1 in conjunction with the (natural) bijection $\text{Hom}_{\mathbb{k}}(\mathbb{k}X, A |_{\text{Vect}}) \cong \text{Hom}_{\text{Sets}}(X, A |_{\text{Sets}})$ from (A.4) gives a natural bijection, for any $\mathbb{k}$-algebra $A$,

$$\text{Hom}_{\mathbb{k}\text{Alg}}(\mathbb{k}\langle X \rangle, A) \overset{\sim}{\longrightarrow} \text{Hom}_{\text{Sets}}(X, A |_{\text{Sets}})$$

Thus, an algebra map $f : \mathbb{k}\langle X \rangle \rightarrow A$ is determined by the values $f(x) \in A$ for the generators $x \in X$ and these values can be freely assigned in order to define $f$. If $X = \{x_1, x_2, \ldots, x_n\}$ is finite, then we will also write $\mathbb{k}\langle x_1, x_2, \ldots, x_n \rangle$ for $\mathbb{k}\langle X \rangle$. In

\[\text{Hom}_{\mathbb{k}\text{Alg}}(\mathbb{k}\langle X \rangle, A) \overset{\sim}{\longrightarrow} \text{Hom}_{\text{Sets}}(X, A |_{\text{Sets}})\]

\[f \quad \quad f \mid_X\]

It will be clear from the context whether $A^k$ denotes the $k^\text{th}$ homogeneous component or the $k$-fold cartesian product $A \times \cdots \times A$ of $A$.\[\]
this case, (1.4) becomes

\[ \text{Hom}_{\text{Alg}}(k\langle x_1, x_2, \ldots, x_n \rangle, A) \sim A^n \]

(1.5)

Algebras having a finite set of generators are called affine. They are exactly the homomorphic images of free algebras \( k\langle X \rangle \) generated by a finite set \( X \) or, equivalently, the homomorphic images of tensor algebras \( T^V \) with \( V \) finite dimensional.

### Polynomial and Symmetric Algebras

Our next example is the familiar commutative polynomial algebra \( k[x_1, x_2, \ldots, x_n] \), with unit map sending \( k \) to the constant polynomials. Formally, the polynomial algebra can be defined by

\[ k[x_1, x_2, \ldots, x_n] \overset{\text{def}}{=} k\langle x_1, x_2, \ldots, x_n \rangle/(x_i x_j - x_j x_i \mid 1 \leq i < j \leq n), \]

where \((\ldots)\) denotes the ideal that is generated by the indicated elements. Since these elements are all homogeneous (of degree 2), the total degree grading of the free algebra \( k\langle x_1, x_2, \ldots, x_n \rangle \) passes down to a grading of \( k[x_1, x_2, \ldots, x_n] \) (Exercise 1.1.11); the grading thus obtained is the usual total degree grading of the polynomial algebra. The universal property (1.5) of the free algebra yields a corresponding universal property for \( k[x_1, x_2, \ldots, x_n] \). Indeed, for any \( k \)-algebra \( A \), the set \( \text{Hom}_{\text{Alg}}(k\langle x_1, x_2, \ldots, x_n \rangle, A) \) can be identified with the set of all algebra maps \( f : k\langle x_1, x_2, \ldots, x_n \rangle \rightarrow A \) such that \( f(x_i x_j - x_j x_i) = 0 \) or, equivalently, \( f(x_i) f(x_j) = f(x_j) f(x_i) \) for all \( i, j \). Thus, for any \( k \)-algebra \( A \), sending an algebra map \( f \) to the \( n \)-tuple \((f(x_i))\) yields a natural bijection of sets

\[ \text{Hom}_{\text{Alg}}(k\langle x_1, x_2, \ldots, x_n \rangle, A) \sim \{(a_i) \in A^n \mid a_i a_j = a_j a_i \forall i, j\}. \]

(1.6)

Letting \( \text{CommAlg}_k \) denote the full subcategory of \( \text{Alg}_k \) consisting of all commutative \( k \)-algebras, then this becomes a natural bijection, for any \( A \in \text{CommAlg}_k \),

\[ \text{Hom}_{\text{CommAlg}}(k\langle x_1, x_2, \ldots, x_n \rangle, A) \sim A^n \]

(1.7)

Since this bijection is analogous to (1.5), but in the world of commutative algebras, \( k[x_1, x_2, \ldots, x_n] \) is also called the free commutative \( k \)-algebra generated by the \( x_i \)s.

Exactly as the tensor algebra \( TV \) of a \( k \)-vector space \( V \) can be thought of as a more general basis-free version of the free algebra \( k\langle x_1, x_2, \ldots, x_n \rangle \), the symmetric
algebra of $V$ generalizes the polynomial algebra $\mathbb{k}[x_1, x_2, \ldots, x_n]$; it is defined by

\[
\text{Sym} V \overset{\text{def}}{=} (TV)/I \quad \text{with} \quad I = I(V) = (v \otimes v' - v' \otimes v \mid v, v' \in V)
\]

Since the ideal $I$ is generated by homogeneous elements of $TV$, it follows that $I = \bigoplus_k I \cap V^\otimes k$, thereby making $\text{Sym} V$ a graded algebra (Exercise 1.1.11):

\[
\text{Sym} V = \bigoplus_{k \in \mathbb{Z}_+} \text{Sym}^k V \quad \text{with} \quad \text{Sym}^k V = V^\otimes k / I \cap V^\otimes k.
\]

Since the nonzero generators of $I$ have degree $> 1$, it follows that $I \cap V = 0$. Thus, we may again view $V \subseteq \text{Sym} V$ and we can write the image of $v_1 \otimes \cdots \otimes v_k \in V^\otimes k$ in $\text{Sym} V$ as $v_1 v_2 \cdots v_k \in \text{Sym}^k V$.

The foregoing yields a functor

\[
\text{Sym} : \text{Vect}_k \rightarrow \text{CommAlg}_k.
\]

Indeed, $\text{Sym} V$ is clearly a commutative $k$-algebra for every $V \in \text{Vect}_k$. Moreover, if $f \in \text{Hom}_k(V, W)$ is a homomorphism of vector spaces, then the image of a typical generator $v \otimes v' - v' \otimes v \in I(V)$ under the map $Tf \in \text{Hom}_{\text{Alg}}(TV, TW)$ is the element $f(v) \otimes f(v') - f(v') \otimes f(v) \in I(W)$. Thus $Tf$ maps $I(V)$ to $I(W)$, and hence $Tf$ passes down to an algebra map $\text{Sym} f : \text{Sym} V \rightarrow \text{Sym} W$. This is in fact a homomorphism of graded algebras:

\[
\text{Sym}^k f = (\text{Sym} f)|_{\text{Sym}^k V} : \text{Sym}^k V \longrightarrow \text{Sym}^k W
\]

For any commutative $k$-algebra $A$, there is a natural bijection

\[
\text{Hom}_{\text{CommAlg}_k}(\text{Sym} V, A) \rightarrow \text{Hom}_{\text{Vect}_k}(V, A)
\]

(1.8)

This follows from Proposition 1.1 exactly as (1.7) was derived from (1.5) earlier. As in Proposition 1.1, the bijection (1.8) states, more formally, that the functor $\text{Sym}$ is left adjoint to the forgetful functor $\text{Vect}_k : \text{CommAlg}_k \rightarrow \text{Vect}_k$. If $V = kX$ for a set $X$, then $\text{Hom}_k(kX, A)|_{\text{Vect}_k} \equiv \text{Hom}_{\text{Sets}}(X, A)|_{\text{Sets}}$ by (A.4) and so (1.8) gives a natural bijection, for any $A \in \text{CommAlg}_k$,

\[
\text{Hom}_{\text{CommAlg}_k}(\text{Sym} kX, A) \rightarrow \text{Hom}_{\text{Sets}}(X, A)
\]

(1.9)
If $X = \{x_1, x_2, \ldots, x_n\}$, then $\text{Hom}_{\text{Sets}}(X, A)_{\text{Sets}} \cong A^n$. Comparing the above bijection with (1.7), it follows that (Exercise 1.1.1)

$$\text{Sym} \ k X \cong k[x_1, x_2, \ldots, x_n].$$

As is well known (see also Exercise 1.1.12), a $k$-basis of the homogeneous component of degree $k$ of the polynomial algebra $k[x_1, x_2, \ldots, x_n]$ is given by the so-called standard monomials of degree $k$,

$$x_1^{k_1} x_2^{k_2} \ldots x_n^{k_n} \quad \text{with } k_i \in \mathbb{Z}_+ \text{ and } \sum_i k_i = k. \tag{1.10}$$

In particular,

$$\dim_k \text{Sym}^k V = \binom{k + n - 1}{n - 1} \quad (n = \dim_k V),$$

as can be seen by identifying each standard monomial with a pattern consisting of $k$ stars and $n-1$ bars:

$$\underbrace{\star \cdots \star}_{k_1} \underbrace{\star \cdots \star}_{k_2} \cdots \underbrace{\star \cdots \star}_{k_n}.$$

**Exterior Algebras**

The exterior algebra $\Lambda V$ of a $k$-vector space $V$ is defined by

$$\Lambda V \overset{\text{def}}{=} (TV)/J \quad \text{with } J = J(V) = \langle v \otimes v \mid v \in V \rangle.$$

Exactly as for the ideal $I$ of $\text{Sym} V$, one sees that $J = \bigoplus_k J \cap V^\otimes k$ and $J \cap V = 0$. Thus, we may again view $V \subseteq \Lambda V$ and $\Lambda V$ is a graded algebra:

$$\Lambda V = \bigoplus_{k \in \mathbb{Z}_+} \Lambda^k V \quad \text{with } \Lambda^k V = V^\otimes k / J \cap V^\otimes k.$$

Writing the canonical map $V^\otimes k \to \Lambda^k V$ as $v_1 \otimes \cdots \otimes v_k \mapsto v_1 \wedge \cdots \wedge v_k$, multiplication in $\Lambda V$ becomes

$$(v_1 \wedge \cdots \wedge v_k)(v'_1 \wedge \cdots \wedge v'_l) = v_1 \wedge \cdots \wedge v_k \wedge v'_1 \wedge \cdots \wedge v'_l.$$

Following the reasoning for $\text{Sym}$, one obtains a functor,

$$\wedge : \text{Vect}_k \to \text{Alg}_k,$$

and the map $\wedge f : \Lambda V \to \Lambda W$ for $f \in \text{Hom}_k(V, W)$ is in fact a homomorphism of graded algebras:

$$\Lambda^k f = \wedge f|_{\Lambda^k V} : \quad \Lambda^k V \xrightarrow{\wedge} \Lambda^k W \quad v_1 \wedge \cdots \wedge v_k \mapsto f(v_1) \wedge \cdots \wedge f(v_k)$$

The defining relations of the exterior algebra state that $v \wedge v = 0$ for all $v \in V$; in words, the exterior product is **alternating** on elements of $V$. Expanding the
product \((v + v') \wedge (v + v') = 0\) and using \(v \wedge v = v' \wedge v' = 0\), one obtains the rule \(v \wedge v' = -v' \wedge v\) for all \(v, v' \in V\). Conversely, with \(v = v'\), this rule gives \(v \wedge v = -v \wedge v\) for all \(v \in V\), which in turn forces \(v \wedge v = 0\) in case \(\text{char } k \neq 2\). Thus, in this case, \(\wedge V = (TV)/(v \otimes v' + v' \otimes v \mid v, v' \in V)\). In general, the rule \(v \wedge v' = -v' \wedge v\) for \(v, v' \in V\) implies by induction that \(ab = (-1)^{|a||b|} ba\) for all \(a \in \wedge^k V\) and \(b \in \wedge^l V\). Using \(|.|\) to denote degrees of homogeneous elements, the latter relation gives the following property, which is referred to as anticommutativity or graded-commutativity of the exterior algebra:

\[
\text{(1.11) } ab = (-1)^{|a||b|} ba.
\]

It follows that, for any given collection of elements \(v_1, v_2, \ldots, v_n \in V\) and any permutation \(s\) of the indices \(\{1, 2, \ldots, n\}\),

\[
\text{(1.12) } v_{s(1)} \wedge v_{s(2)} \cdot \cdot \cdot v_{s(k)} = \text{sgn}(s) v_1 \wedge v_2 \cdot \cdot \cdot v_k,
\]

where \(\text{sgn}(s)\) denotes the sign of the permutation \(s\). Indeed, (1.12) is clear from anticommutativity in case \(s\) is a transposition interchanging two adjacent indices; the general case is a consequence of the standard fact that these transpositions generate the symmetric group \(S_n\) (Example 7.10).

Anticommutativity implies that if \(V\) has basis \((x_i)_{i \in I}\), then the elements

\[
\text{(1.13) } x_{i_1} \wedge x_{i_2} \cdot \cdot \cdot x_{i_k} \quad \text{with } i_1 < i_2 < \cdot \cdot \cdot < i_k
\]
generate the \(k\)-vector space \(\wedge^k V\). These elements do in fact form a basis of \(\wedge^k V\); see Exercise 1.1.12. Therefore, if \(\dim_k V = n < \infty\), then

\[
\text{dim}_k \wedge^k V = \binom{n}{k} \quad \text{and} \quad \dim_k \wedge V = 2^n.
\]

In particular, \(\wedge^n V\) is 1-dimensional and, for any \(f \in \text{End}_k(V)\), the endomorphism \(\wedge^n f \in \text{End}_k(\wedge^n V) = k\) is given by the determinant (see also Lemma 3.33):

\[
\text{(1.14) } \wedge^n f = \det f.
\]

The Weyl Algebra

The following algebra is called the (first) Weyl algebra over \(k\):

\[
\text{(1.15) } A_1(k) \overset{\text{def}}{=} k\langle x, y \rangle/(yx - xy - 1)
\]

Committing a slight abuse of notation, let us keep \(x\) and \(y\) to denote their images in \(A_1(k)\); so \(yx = xy + 1\) in \(A_1(k)\). This relation allows us to write each finite product in \(A_1(k)\) with factors \(x\) or \(y\) as a \(k\)-linear combination of ordered products of the form \(x^i y^j\) \((i, j \in \mathbb{Z}_+\)). These standard monomials therefore generate \(A_1(k)\) as \(k\)-vector space. One can show that they are in fact linearly independent (Exercise 1.1.14; see also Examples 1.8 and D.3); hence the standard monomials form a \(k\)-basis of the Weyl algebra.
If \( f : A_1(\mathbb{k}) \to A \) is any \( \mathbb{k} \)-algebra map and \( a = f(x), b = f(y) \), then we must have \( ba - ab - 1 = 0 \) in \( A \). However, this relation is the only restriction, because it guarantees that the homomorphism \( \mathbb{k}(x, y) \to A \) that corresponds to the pair \((a, b) \in A^2\) in (1.5) does in fact factor through \( A_1(\mathbb{k}) \). Thus, we have a bijection, natural in \( A \in \text{Alg}_\mathbb{k} \),

\[
\text{Hom}_{\text{Alg}_\mathbb{k}}(A_1(\mathbb{k}), A) \cong \{(a, b) \in A^2 \mid ba - ab - 1 = 0\}.
\]

### 1.1.3. Modules

In order to pave the way for the dual concept of a “comodule,” to be introduced later in §9.2.1, we now review the basic definitions concerning modules over \( \mathbb{k} \)-algebras in the diagrammatic style of §1.1.1, working in the category \( \text{Vect}_\mathbb{k} \). We will also briefly discuss some issues related to switching sides.

#### Left Modules

Let \( A = (A, m, u) \) be a \( \mathbb{k} \)-algebra. A **left module** over \( A \), by definition, is an an abelian group \((V, +)\) that is equipped with a left action of \( A \), that is, a biadditive map \( A \times V \to V, (a, v) \mapsto a.v \), satisfying the conditions

\[
a.(b.v) = (ab).v \quad \text{and} \quad 1_A.v = v
\]

for all \( a, b \in A \) and \( v \in V \). Putting \( \lambda v := u(\lambda).v \) for \( \lambda \in \mathbb{k} \), the group \( V \) becomes a \( \mathbb{k} \)-vector space. The action map is easily seen to be \( \mathbb{k} \)-bilinear, and hence it corresponds to a \( \mathbb{k} \)-linear map \( A \otimes V \to V \) by (B.12). Thus, a left \( A \)-module may equivalently be defined as a vector space \( V \in \text{Vect}_\mathbb{k} \) together with a linear map \( \mu = \mu_V : A \otimes V \to V \) such that the following two diagrams commute:

\[\begin{align*}
A \otimes A \otimes V & \xrightarrow{m \otimes \text{Id}_V} A \otimes V \\
\mu & \downarrow \\
A \otimes V & \xrightarrow{\text{Id}_A \otimes \mu} V
\end{align*}\]

\[
\begin{align*}
\mathbb{k} \otimes V & \xrightarrow{u \otimes \text{Id}_V} A \otimes V \\
\mu & \downarrow \\
V & \xrightarrow{\mu}
\end{align*}
\]

We will generally suppress \( \mu \), writing instead \( \mu(a \otimes v) = a.v \) as above or else use simple juxtaposition, \( \mu(a \otimes v) = av \).

Given left \( A \)-modules \( V \) and \( W \), a **homomorphism** from \( V \) to \( W \) is the same as a \( \mathbb{k} \)-linear map \( f : V \to W \) such that the following diagram commutes:

\[
\begin{align*}
A \otimes V & \xrightarrow{\text{Id}_A \otimes f} A \otimes W \\
\mu_V & \downarrow \quad f \quad \mu_W \\
V & \xrightarrow{f} W
\end{align*}
\]
In terms of elements, this states that \( f(a.v) = a.f(v) \) for all \( a \in A \) and \( v \in V \). As in Appendices A and B, the set of all \( A \)-module homomorphisms \( f : V \to W \) will be denoted by \( \text{Hom}_A(V, W) \) and the resulting category of left \( A \)-modules by

\[
\mathcal{A}\text{Mod}.
\]

Thus, \( \mathcal{A}\text{Mod} \) is a subcategory of \( \text{Vect}_k \). Note also that \( \text{End}_A(V) := \text{Hom}_A(V, V) \) is always a \( k \)-subalgebra of \( \text{End}_k(V) \).

We refrain from reminding the reader in tedious detail of the fundamental module theoretic notions such as submodule, factor module, . . . and we shall also assume familiarity with the isomorphism theorems and other standard facts. We will however remark that, by virtue of the bijection \( \text{Hom}_k(A \otimes V, V) \cong \text{Hom}_k(A, \text{End}_k(V)) \) that is given by Hom-\( \otimes \) adjunction (B.15), a left module action \( \mu : A \otimes V \to V \) corresponds to an algebra map,

\[
\rho : A \to \text{End}_k(V).
\]

In detail, for a given \( \rho \), we may define an action \( \mu \) by \( \mu(a \otimes v) := \rho(a)(v) \) for \( a \in A \) and \( v \in V \). Conversely, from a given action \( \mu \), we obtain \( \rho \) by defining \( \rho(a) := (v \mapsto \mu(a \otimes v)) \).

---

### Changing Sides: Opposite Algebras

Naturally, the category \( \text{Mod}_A \) of all right modules over a given algebra \( A \) (as in Appendices A and B) can be also described by diagrams in \( \text{Vect}_k \) analogous to (1.17) and (1.18). However, it turns out that right \( A \)-modules are essentially the same as left modules over a related algebra, the so-called opposite algebra \( A^{\text{op}} \) of \( A \). As a \( k \)-vector space, \( A^{\text{op}} \) is identical to \( A \) but \( A^{\text{op}} \) is equipped with a new multiplication \(*\) that is given by \( a * b = ba \) for \( a, b \in A \). Alternatively, we may realize \( A^{\text{op}} \) as a vector space isomorphic to \( A \) via \( A \to A^{\text{op}} \), and with multiplication given by \( a^{\text{op}} * b^{\text{op}} = (ba)^{\text{op}} \). Clearly, \( A^{\text{op}}^{\text{op}} \equiv A \).

Now suppose that \( V \) is a right \( A \)-module with right action \( \mu : V \otimes A \to V \). Then we obtain a left \( A^{\text{op}} \)-module structure on \( V \) by defining \( \mu^{\text{op}} : A^{\text{op}} \otimes V \to V 
\mu^{\text{op}}(a^{\text{op}} \otimes v) = \mu(v \otimes a) \). Likewise, any left \( A \)-module action \( \mu : A \otimes V \to V \) gives rise to a right \( A^{\text{op}} \)-action via \( \mu^{\text{op}} : V \otimes A^{\text{op}} \to V, \mu^{\text{op}}(v \otimes a^{\text{op}}) = \mu(a \otimes v) \). Left \( A^{\text{op}} \)-modules become right modules over \( A^{\text{op}}^{\text{op}} \equiv A \) in this way. Therefore, we obtain an equivalence of categories (§A.3.3)

\[
\text{Mod}_A \equiv A^{\text{op}}\text{Mod}.
\]

Alternatively, in terms of algebra maps, it is straightforward to check as above that a right \( A \)-module action \( V \otimes A \to V \) corresponds to an algebra map \( A \to \text{End}_k(V)^{\text{op}} \). Such a map in turn clearly corresponds to an algebra map \( A^{\text{op}} \to \text{End}_k(V)^{\text{op}} \equiv \text{End}_k(V) \), and hence to a left \( A^{\text{op}} \)-module action on \( V \).
Bimodules: Tensor Products of Algebras

We will almost exclusively work in the context of left modules, but occasionally we shall also encounter modules that arise naturally as right modules or even as bimodules (§B.1.2). If $A$ and $B$ are $k$-algebras, then an $(A, B)$-bimodule is the same as a $k$-vector space $V$ that is both a left $A$-module and a right $B$-module, with module actions $\mu: A \otimes V \to V$ and $\mu': V \otimes B \to V$, such that the following diagram commutes:

$$
\begin{array}{ccc}
A \otimes V \otimes B & \xrightarrow{\mu \otimes \text{Id}_B} & V \otimes B \\
\text{Id}_A \otimes \mu' & & \mu' \\
A \otimes V & \xrightarrow{\mu} & V
\end{array}
$$

Defining morphisms between $(A, B)$-bimodules to be the same as $k$-linear maps that are left $A$-module as well as right $B$-module maps, we once again obtain a category, $A_{\mathbf{Mod}}_B$. As with right modules, $(A, B)$-bimodules are in fact left modules over some algebra, the algebra in question being the tensor product $A \otimes B^\text{op}$ (§1.1.1). Indeed, suppose that $V$ is an $(A, B)$-bimodule. As we have remarked above, the module actions correspond to algebra maps $\alpha : A \to \text{End}_k(V)$ and $\beta : B^\text{op} \to \text{End}_k(V)$. Condition (1.19) can be expressed by stating that the images of these maps commute elementwise. The “universal property” of the tensor product of algebras (Exercise 1.1.10), therefore provides us with a unique algebra map $A \otimes B^\text{op} \to \text{End}_k(V)$, $a \otimes b^\text{op} \mapsto \alpha(a)\beta(b^\text{op})$, and this algebra map in turn corresponds to a left $A \otimes B^\text{op}$-module structure on $V$. In short, we have an equivalence of categories,

$$A_{\mathbf{Mod}}_B \cong A\otimes B^\text{op}_{\mathbf{Mod}}$$

**Example 1.2** (The regular bimodule). Every algebra $A$ carries a natural $(A, A)$-bimodule structure, with left and right $A$-module actions given by left and right multiplication, respectively. Commutativity of (1.19) for these actions is equivalent to the associative law of $A$. The resulting left, right and bimodule structures will be referred to as the **regular** structures. We will be primarily concerned with the left regular module structure; it will be denoted by $A_{\text{reg}} \in A_{\mathbf{Mod}}$ so as to avoid any confusion with the algebra $A \in \text{Alg}_k$. By the foregoing, we may view the regular $(A, A)$-bimodule $A$ as a left module over the algebra $A \otimes A^\text{op}$.

**Example 1.3** (Bimodule structures on Hom-spaces). For $A, B \in \text{Alg}_k$ and given modules $V \in A_{\mathbf{Mod}}$ and $W \in B_{\mathbf{Mod}}$, the $k$-vector space $\text{Hom}_k(W, V)$ becomes an $(A, B)$-bimodule by defining

$$(a.f.b)(w) := a.f(b.w)$$
for \( a \in A, b \in B, f \in \text{Hom}_k(W, V) \) and \( w \in W \). Thus, \( \text{Hom}_k(W, V) \) becomes a left \( A \otimes B^{\text{op}} \)-module. We may also regard \( V \) as a left module over the endomorphism algebra \( \text{End}_A(V) \) and likewise for \( W \). If \( A = B \), then the above bimodule action equips \( \text{Hom}_A(W, V) \) with a \((\text{End}_A(V), \text{End}_A(W))\)-bimodule structure, with actions given by composition in \( \text{AMod} \).

### 1.1.4. Endomorphism Algebras and Matrices

This subsection provides some technicalities for later use; it may be skipped at a first reading and referred to as the need arises. Throughout, \( A \) denotes an arbitrary \( k \)-algebra.

#### Direct Sums

Our first goal is to describe the endomorphism algebra of a finite direct sum \( \bigoplus_{i=1}^n V_i \) with \( V_i \in \text{AMod} \). If all \( V_i = V \), then we will write \( \bigoplus_{i=1}^n V_i = V^{\otimes n} \).

In general, the various embeddings and projections are module maps

\[
\mu_i : V_i \hookrightarrow \bigoplus_{i=1}^n V_i \quad \text{and} \quad \pi_i : \bigoplus_{i=1}^n V_i \twoheadrightarrow V_i.
\]

Explicitly, \( \pi_i(v_1, v_2, \ldots, v_n) = v_i \) and \( \mu_i(v) = (0, \ldots, 0, v, 0, \ldots, 0) \) with \( v \) occupying the \( i \)th component on the right. Consider the generalized \( n \times n \) matrix algebra,

\[
\left( \text{Hom}_A(V_j, V_i) \right)_{i,j} = \begin{pmatrix}
\text{Hom}_A(V_1, V_1) & \cdots & \text{Hom}_A(V_n, V_1) \\
\vdots & \ddots & \vdots \\
\text{Hom}_A(V_1, V_n) & \cdots & \text{Hom}_A(V_n, V_n)
\end{pmatrix}.
\]

The \( k \)-vector space structure of this set is “entrywise,” using the standard \( k \)-linear structure on each \( \text{Hom}_A(V_j, V_i) \subseteq \text{Hom}_k(V_j, V_i) \) and identifying the generalized matrix algebra with the direct sum of vector spaces \( \bigoplus_{i,j} \text{Hom}_A(V_j, V_i) \). Multiplication comes from composition:

\[
(f_{ik})(g_{kj}) = (\sum_k f_{ik} \circ g_{kj}).
\]

Note the reversal of indices: \( f_{ij} \in \text{Hom}_A(V_j, V_i) \). The identity element of the generalized matrix algebra is the diagonal matrix with entries \( \text{Id}_{V_i} \).

**Lemma 1.4.**

(a) For \( V_1, \ldots, V_n \in \text{AMod} \), there is an isomorphism in \( \text{Alg}_k \),

\[
\text{End}_A(\bigoplus_{i=1}^n V_i) \xrightarrow{\sim} \left( \text{Hom}_A(V_j, V_i) \right)_{i,j} \wrt \mu_j
\]

\[
f \longmapsto \left( \pi_i \circ f \circ \mu_j \right)_{i,j}
\]
(b) Let $V \in A\text{Mod}$. Then $V^\otimes n$ becomes a left module over $\text{Mat}_n(A)$ via matrix multiplication and there is an isomorphism in $\text{Alg}_k$,

$$\text{End}_A(V) \xrightarrow{\sim} \text{End}_{\text{Mat}_n(A)}(V^\otimes n)$$

Similarly, $f^\otimes n = \sum_i \mu_i \circ f \circ \pi_i$.

**Proof.** (a) Let us put $V = \bigoplus_{i=1}^n V_i$ and denote the map in (a) by $\alpha$; it is clearly $k$-linear. In fact, $\alpha$ is an isomorphism by (B.14). In order to show that $\alpha$ is an algebra map, we note the relations $\sum_k \mu_k \circ \pi_k = \text{Id}_{V}$ and $\pi_i \circ \mu_j = \delta_{i,j} \text{Id}_{V_j}$ (Kronecker $\delta$). Using this, we compute

$$\alpha(f \circ g) = \left( \pi_i \circ f \circ g \circ \mu_j \right)$$

$$= \left( \pi_i \circ f \circ (\sum_k \mu_k \circ \pi_k) \circ g \circ \mu_j \right)$$

$$= \left( \sum_k (\pi_i \circ f \circ \mu_k) \circ (\pi_k \circ g \circ \mu_j) \right)$$

$$= \alpha(f) \alpha(g).$$

Similarly $\alpha(1) = 1$. This shows that $\alpha$ is a $k$-algebra homomorphism, proving (a).

(b) In componentwise notation, the map $f^\otimes n$ is given by $(v_i) \mapsto (f(v_i))$ and the “matrix multiplication” action of $\text{Mat}_n(A)$ on $V^\otimes n$ by $(a_{ij})(v_j) = (\sum_j a_{ij}v_j)$. It is straightforward to check that $f^\otimes n \in \text{End}_{\text{Mat}_n(A)}(V^\otimes n)$ and that $f \mapsto f^\otimes n$ is a $k$-algebra map. The inverse map is given by the $n\times n$-matrix $g$.

$$\text{End}_{\text{Mat}_n(A)}(V^\otimes n) \rightarrow \text{End}_A(V), \quad g \mapsto \pi_1 \circ g \circ \mu_1.$$}

For example, in order to check that $\sum_i \mu_i \circ \pi_1 \circ g \circ \mu_1 \circ \pi_i = g$, observe that $g$ commutes with the operators $\mu_i \circ \pi_j : V^\otimes n \rightarrow V^\otimes n$, because $\mu_i \circ \pi_j$ is given by the action of the matrix $e_{i,j} \in \text{Mat}_n(A)$, with 1 in the $(i,j)$-position and 0s elsewhere. Therefore,

$$\sum_i \mu_i \circ \pi_1 \circ g \circ \mu_1 \circ \pi_i = \sum_i \mu_i \circ \pi_1 \circ \mu_1 \circ \pi_i \circ g = \text{Id}_{V^\otimes n} \circ g = g.$$}

This completes the proof of the lemma. $\square$

**Free Modules**

We now turn to a generalization of the familiar fact from linear algebra that the $n \times n$-matrix algebra $\text{Mat}_n(k)$ is the endomorphism algebra of the vector space $k^n$. In place of $k^n$, we consider the $n$-fold direct sum $\bigoplus_{\text{reg}}^n$ of the regular module (Example 1.2). Left $A$-modules isomorphic to $\bigoplus_{\text{reg}}^n$ for some $n \in \mathbb{Z}_+$ are called **finitely generated free** (Example A.5).

**Lemma 1.5.**  
(a) $\text{Mat}_n(A)^{\text{op}} \cong \text{Mat}_n(A^{\text{op}})$ in $\text{Alg}_k$, via the matrix transpose.
(b) There is an isomorphism in $\text{Alg}_\kappa$, given by right matrix multiplication,

$$\begin{array}{cc}
\text{Mat}_n(A)^\text{op} & \sim \text{End}_A(A_{\text{reg}}) \\
\psi & \psi \\
(x_{ij}) & \left( (a_i) \mapsto (\sum_i a_i x_{ij}) \right)
\end{array}$$

**Proof.** (a) We will identify opposite algebras with the originals, but with multiplication $\ast$. Consider the map $\cdot^T: \text{Mat}_n(A)^\text{op} \to \text{Mat}_n(A^\text{op})$ sending each matrix to its transpose; this is clearly a $\kappa$-linear bijection fixing the identity matrix $I_{n\times n}$. We need to check that, for $X = (x_{ij}), Y = (y_{ij}) \in \text{Mat}_n(A)^\text{op}$, the equation $(X \ast Y)^T = X^TY^T$ holds in $\text{Mat}_n(A^\text{op})$. But the matrix $(X \ast Y)^T = (YX)^T$ has $(i,j)$-entry $\sum_k y_{jk}x_{ki}$, whereas the $(i,j)$-entry of $X^TY^T$ equals $\sum_k x_{ki}y_{kj}$. By definition of the multiplication in $A^\text{op}$, these two entries are identical.

(b) Right multiplication by $x \in A$ gives map $r_x = \cdot x \in \text{End}_A(A_{\text{reg}})$. Since $r_x \circ r_y = r_{xy} = r_{yx}$ for $x, y \in A$, the assignment $x \mapsto r_x$ is an algebra map $A^\text{op} \to \text{End}_A(A_{\text{reg}})$. This map has inverse $\text{End}_A(A_{\text{reg}}) \to A^\text{op}, f \mapsto f(1)$. Hence, $\text{End}_A(A_{\text{reg}})^\text{op} \cong A^\text{op}$ as $\kappa$-algebras and so

$$\text{End}_A(A_{\text{reg}})^\text{op} \cong \text{Mat}_n(\text{End}_A(A_{\text{reg}})) \cong \text{Mat}_n(A^\text{op}) \cong \text{Mat}_n(A)^\text{op}. $$

It is readily checked that this isomorphism is explicitly given as in the lemma. □

**Exercises for Section 1.1**

1.1.1 (Universal properties). (a) Let $TV \in \text{Alg}_\kappa$ be equipped with a $\kappa$-linear map $t: V \to TV|_{\text{Vect}_k}$ such that the map $t^* = \cdot t: \text{Hom}\kappa(TV, A) \to \text{Hom}\kappa(V, A)|_{\text{Vect}_k}$ is a bijection for any $A \in \text{Alg}_\kappa$. Show that $TV \cong TV$.

(b) Deduce from (1.7) and (1.9) that $\text{Sym}\kappa[n] \cong \kappa[x_1, \ldots, x_n]$, where $[n] = \{1, 2, \ldots, n\}$.

1.1.2 (Splitting maps). (a) Let $U \xrightarrow{f} V \xrightarrow{g} W$ be maps in $A\text{Mod}$. Show that $g \circ f$ is an isomorphism if and only if $f$ is mono, $g$ is epi and $V = \text{Im} f \oplus \text{Ker} g$. If $U = W$ and $g \circ f = \text{Id}_W$, then one says that the maps $f$ and $g$ **split** each other.

(b) Let $0 \to U \xrightarrow{f} V \xrightarrow{g} W \to 0$ be a short exact sequence in $A\text{Mod}$ (§B.1.1) and put $S := \text{Im} f = \text{Ker} g$. Show that the following conditions are equivalent; if they hold, the given short exact sequence is said to be **split**:

(i) $f' \circ f = \text{Id}_U$, for some $f' \in \text{Hom}_A(V, U)$;

(ii) $g \circ g' = \text{Id}_W$, for some $g' \in \text{Hom}_A(W, V)$;

(iii) $S$ has a **complement**, that is, $V = S \oplus C$ for some $A$-submodule $C \subseteq V$. 
1.1.3 (Generators of a module). Let $A \in \text{Alg}_k$ and $V \in \mathcal{A}\text{Mod}$. A subset $\Gamma \subseteq V$ is said to generate $V$ if the only submodule of $V$ containing $\Gamma$ is $V$ itself. Modules that have a finite generating set are called finitely generated; modules that are generated by one element are called cyclic.

(a) Let $V$ be finitely generated. Use Zorn’s Lemma to show that every proper submodule $U \subseteq V$ is contained in a maximal proper submodule $M$, that is, $M \nsubseteq V$ and $M \subseteq M' \nsubseteq V$ implies $M = M'$.

(b) Let $0 \to U \to V \to W \to 0$ be a short exact sequence in $\mathcal{A}\text{Mod}$. Show that if both $U$ and $W$ are both finitely generated, then $V$ is finitely generated as well. Conversely, assuming $V$ to be finitely generated, show that $W$ is finitely generated but this need not hold for $U$. (Give an example to that effect.)

(c) Show that the following are equivalent:

(i) $V$ has a generating set consisting of $n$ elements;

(ii) $V$ is a homomorphic image of the free left $A$-module $A^{\oplus n}$;

(iii) $V^{\oplus n}$ is a cyclic left $\text{Mat}_n(A)$-module (Lemma 1.4).

1.1.4 (Noetherian modules). A left module $V$ over an algebra $A$ is said to be noetherian if $V$ satisfies the Ascending Chain Condition (ACC) on its submodules: if $U_1 \subseteq U_2 \subseteq U_3 \subseteq \ldots$ are submodules of $V$, then $U_n = U_{n+1} = \ldots$ for some $n$.

(a) Show that ACC has the following equivalent reformulations:

(i) All submodules of $V$ are finitely generated.

(ii) Every nonempty collection of submodules of $V$ has at least one maximal member (Maximum Condition on submodules).

In fact, ACC and the Maximum Condition may be formulated for any partially ordered set and they are equivalent in this more general setting (assuming the Axiom of Choice).

(b) Let $0 \to U \to V \to W \to 0$ be a short exact sequence in $\mathcal{A}\text{Mod}$. Show that $V$ is noetherian if and only if both $U$ and $W$ are.

1.1.5 (Noetherian algebras). The algebra $A$ is called left noetherian if $A_{\text{reg}} \in \mathcal{A}\text{Mod}$ is noetherian, that is, $A$ satisfies ACC on left ideals. Right noetherian algebras are defined likewise using right ideals. Algebras that are both right and left noetherian are simply called noetherian.

(a) Assuming $A$ to be left noetherian, show that all finitely generated left $A$-modules are noetherian.

(b) Let $B$ be a subalgebra of $A$ such that the $k$-algebra $A$ is generated by $B$ and an element $x$ such that $Bx + B = xB + B$. Adapt the proof of the Hilbert Basis Theorem to show that if $B$ is left (or right) noetherian, then so is $A$.

1.1.6 (Skew polynomial algebras). Let $A \in \text{Alg}_k$. Like the ordinary polynomial algebra $A[x]$, a skew polynomial algebra over $A$ is a $k$-algebra, $B$, containing $A$ as
a subalgebra and an additional element \( x \in B \) whose powers form a basis of \( B \) as left \( A \)-module. Thus, as in \( A[x] \), every element of \( B \) can be uniquely written as a finite sum \( \sum a_i x^i \) with \( a_i \in A \). However, we now only insist on the inclusion \( xA \subseteq Ax + A \) to hold; so all products \( xa \) with \( a \in A \) can be written in the form
\[
xa = \sigma(a)x + \delta(a)
\]
with unique \( \sigma(a), \delta(a) \in A \).

(a) Show that the above rule leads to a \( k \)-algebra multiplication on \( B \) if and only if \( \sigma \in \text{End}_{\text{Alg}_k}(A) \) and \( \delta \) is a \( k \)-linear endomorphism of \( A \) satisfying
\[
\delta(aa') = \sigma(a)\delta(a') + \delta(a)a' \quad (a, a' \in A).
\]

Maps \( \delta \) as above are called \textit{left} \( \sigma \)-\textit{derivations} of \( A \) and the resulting algebra \( B \) is denoted by \( A[x; \sigma, \delta] \). If \( \sigma = \text{Id}_A \), then one simply speaks of a \textit{derivation} \( \delta \) and writes \( A[x; \delta] \) for \( A[x; \text{Id}_A, \delta] \). Similarly, \( A[x; \sigma] = A[x; \sigma, 0] \). If \( \sigma \in \text{Aut}_{\text{Alg}_k}(A) \), then we may define the \textit{skew Laurent polynomial algebra} \( A[x^\pm 1; \sigma] \) as above except that negative powers of the variable \( x \) are permitted: \( A[x^\pm 1; \sigma] = \bigoplus_{i \in \mathbb{Z}} Ax^i \) and \( x^i a = \sigma^i(a)x^i \) for \( a \in A \). Assuming \( \sigma \in \text{Aut}_{\text{Alg}_k}(A) \), show:

(b) If \( A \) is a domain, that is, \( A \neq 0 \) and products of nonzero elements of \( A \) are nonzero, then \( A[x; \sigma, \delta] \) is likewise.

(c) If \( A \) is left (or right) noetherian, then so is \( A[x; \sigma, \delta] \). (Use Exercise 1.1.5.)

1.1.7 (Artin-Tate Lemma). Let \( A \in \text{Alg}_k \) be affine and let \( B \subseteq A \) be a \( k \)-subalgebra such that \( A \) is finitely generated as a left \( B \)-module, say \( A = \sum_{i=1}^m Ba_i \). Show:

(a) There exists an affine \( k \)-subalgebra \( B' \subseteq B \) such that \( A = \sum_{i=1}^m B'a_i \).

(b) If \( B \) is commutative, then \( B \) is affine. (Use (a) and Hilbert’s Basis Theorem.)

1.1.8 (Affine algebras and finitely generated modules). Let \( A \in \text{Alg}_k \) be affine and let \( M \in \mathcal{A}_\text{Mod} \) be finitely generated. Show that if \( N \) is an \( A \)-submodule of \( M \) such that \( \dim_k M/N < \infty \), then \( N \) is finitely generated.

1.1.9 (Subalgebras as direct summands). Let \( A \in \text{Alg}_k \) and let \( B \subseteq A \) be a \( k \)-subalgebra such that \( A \) is free as a left \( B \)-module. Show that \( A = B \oplus C \) for some left \( B \)-submodule \( C \subseteq A \).

1.1.10 (Tensor product of algebras). Let \( A, B \in \text{Alg}_k \). Prove:

(a) The tensor product \( A \otimes B \in \text{Alg}_k \) has the following \textit{universal property}: the maps \( a : A \rightarrow A \otimes B \), \( x \mapsto x \otimes 1 \), and \( b : B \rightarrow A \otimes B \), \( y \mapsto 1 \otimes y \), are \( k \)-algebra maps such that \( \text{Im}a \) commutes elementwise with \( \text{Im}b \). Moreover, if \( \alpha : A \rightarrow C \) and \( \beta : B \rightarrow C \) are any \( k \)-algebra maps such that \( \text{Im} \alpha \) commutes elementwise with \( \text{Im} \beta \), then there exists a unique \( k \)-algebra map \( t : A \otimes B \rightarrow C \) satisfying \( t \circ a = \alpha \) and \( t \circ b = \beta \). In particular, the tensor product gives a bifunctor
\[
\otimes : \text{Alg}_k \times \text{Alg}_k \rightarrow \text{Alg}_k.
\]
(b) \( \mathcal{F}(A \otimes B) \cong \mathcal{F}A \otimes \mathcal{F}B \).

(c) \( \mathbb{C} \otimes \mathbb{H} \cong \text{Mat}_2(\mathbb{C}) \) as \( \mathbb{C} \)-algebras.

(d) \( A \otimes \mathbb{k}[x_1, \ldots, x_n] \cong A[x_1, \ldots, x_n] \) as \( \mathbb{k} \)-algebras. In particular,
\[
\mathbb{k}[x_1, \ldots, x_n] \otimes \mathbb{k}[x_1, \ldots, x_m] \cong \mathbb{k}[x_1, \ldots, x_{n+m}].
\]

(e) \( A \otimes \text{Mat}_n(\mathbb{k}) \cong \text{Mat}_n(A) \) as \( \mathbb{k} \)-algebras. In particular,
\[
\text{Mat}_n(\mathbb{k}) \otimes \text{Mat}_m(\mathbb{k}) \cong \text{Mat}_{nm}(\mathbb{k}).
\]

1.1.11 (Graded vector spaces, algebras and modules). Let \( \Delta \) be a monoid, with binary operation denoted by juxtaposition and with identity element \( 1 \). A \( \Delta \)-grading of a \( \mathbb{k} \)-vector space \( V \) is given by a direct sum decomposition
\[
V = \bigoplus_{k \in \Delta} V^k
\]
with \( \mathbb{k} \)-subspaces \( V^k \). The nonzero elements of \( V^k \) are said to be homogeneous of degree \( k \). If \( V \) and \( W \) are \( \Delta \)-graded, then a morphisms \( f : V \rightarrow W \) of \( \Delta \)-graded vector spaces, by definition, is a \( \mathbb{k} \)-linear map that preserves degrees in the sense that
\[
f(V^k) \subseteq W^k \quad \text{for all} \quad k \in \Delta.
\]
In this way, \( \Delta \)-graded \( \mathbb{k} \)-vector spaces form a category, \( \text{Vect}_\mathbb{k}^\Delta \). For any \( V, W \in \text{Vect}_\mathbb{k}^\Delta \), the tensor product \( V \otimes W \) inherits a \( \Delta \)-grading with
\[
(V \otimes W)^k = \bigoplus_{ij=k} V^i \otimes W^j.
\]

A \( \mathbb{k} \)-algebra \( A \) is said to be \( \Delta \)-graded if the underlying \( \mathbb{k} \)-vector space of \( A \) is \( \Delta \)-graded and multiplication \( A \otimes A \rightarrow A \) as well as the unit map \( \mathbb{k} \rightarrow A \) are morphisms of graded vector spaces. Here, \( \mathbb{k} \) has the trivial grading: \( \mathbb{k} = \mathbb{k}^1 \). Explicitly, this means that \( A = \bigoplus_{k \in \Delta} A^k \) for \( \mathbb{k} \)-subspaces \( A^k \) satisfying \( A^k A^{k'} \subseteq A^{kk'} \) for \( k, k' \in \Delta \) and \( 1_A \in A^1 \). In particular \( A^1 \) is a \( \mathbb{k} \)-subalgebra of \( A \). Taking as morphisms the \( \mathbb{k} \)-algebra maps that preserve the \( \Delta \)-grading, we obtain a category, \( \text{Alg}_\mathbb{k}^\Delta \). Let \( A \in \text{Alg}_\mathbb{k}^\Delta \). A module \( V \in \text{\text{Mod}} \) is called \( \Delta \)-graded if the underlying \( \mathbb{k} \)-vector space of \( V \) is \( \Delta \)-graded and the action map \( A \otimes V \rightarrow V \) is a morphism of graded vector spaces: \( V = \bigoplus_{k \in \Delta} V^k \) for \( \mathbb{k} \)-subspaces \( V^k \) such that \( A^k V^{k'} \subseteq V^{kk'} \) for all \( k, k' \in \Delta \).

(a) Let \( A \in \text{Alg}_\mathbb{k}^\Delta \) be such that the underlying \( \mathbb{k} \)-vector space of \( A \) is \( \Delta \)-graded. Assuming that \( k \neq 1 \) implies \( kk' \neq k' \) for all \( k' \in \Delta \), show that \( 1 \in A^1 \) is in fact automatic if multiplication of \( A \) is a map of graded vector spaces.
(b) Let \( A \in \text{Alg}_k \) and \( U \in \text{AMod} \) be \( \Delta \)-graded, and let \( U \) be an \( A \)-submodule of \( V \). Show that \( U = \bigoplus_k (U \cap V^k) \) if and only if \( U \) is generated, as \( A \)-module, by homogeneous elements. In this case, the \( A \)-module \( V/U \) is graded with homogeneous components \( (V/U)^k = V^k / U \cap V^k \).

(c) Let \( A \in \text{Alg}_k \) and let \( I \) be an ideal of \( A \). Show that \( I = \bigoplus_k (I \cap A^k) \) if and only if \( I \) is generated, as an ideal of \( A \), by homogeneous elements. In this case, the algebra \( A/I \) is graded with homogeneous components \( (A/I)^k = A^k / I \cap A^k \).

1.1.12 (Some properties of symmetric and exterior algebras). Let \( \text{Alg}_k \) denote the category of \( k \)-algebras as in Exercise 1.1.11, with \( \mathbb{Z} = (\mathbb{Z}, +) \).

(a) For any \( V, W \in \text{Vect}_k \), show that Sym\( (V \oplus W) \cong \text{Sym} V \oplus \text{Sym} W \) in \( \text{Alg}_k \). (Use Exercise 1.1.10(a) and the universal property (1.8) of the symmetric algebra.)

(b) An algebra \( A \in \text{Alg}_k \) is called anticommutative or graded commutative if \( ab = (-1)^{|a||b|} ba \) for all homogeneous \( a, b \in A \) as in (1.11). If, in addition, \( a^2 = 0 \) for all homogeneous \( a \in A \) of odd degree, then \( A \) is called alternating. (Anticommutative algebras are automatically alternating if \( \text{char} k \neq 2 \).) Show that the exterior algebra \( \Lambda V \) is alternating and that, for any alternating \( k \)-algebra \( A \), there is a natural bijection of sets

\[
\text{Hom}_{\text{Alg}_k}(\Lambda V, A) \cong \text{Hom}_k(V, A^1)
\]

\[
f \mapsto f|_V
\]

(c) Let \( A, B \in \text{Alg}_k \) be alternating. Define \( A \otimes B \) to be the usual tensor product \( A \otimes B \) as a \( k \)-vector space, with the \( \mathbb{Z} \)-grading of Exercise 1.1.11. However, multiplication is not given by (1.3) but rather by the \textbf{Koszul sign rule}:

\[
(a \otimes b)(a' \otimes b') := (-1)^{|b'||a'} aa' \otimes bb'.
\]

Show that this makes \( A \otimes B \) an alternating \( k \)-algebra.

(d) Conclude from (b) and (c) that \( \Lambda (V \oplus W) \cong \Lambda V \otimes \Lambda W \) as graded \( k \)-algebras.

(e) Deduce the bases of \( \text{Sym} V \) and \( \Lambda V \) as stated in (1.10) and (1.13) from the isomorphisms in (a) and (d).

1.1.13 (Central simple algebras). A \( k \)-algebra \( A \neq 0 \) is called \textbf{simple} if 0 and \( A \) are the only ideals of \( A \). Show that this implies that the center \( Z \) of \( A \) is a \( k \)-field. A simple algebra \( A \) is called \textbf{central simple} if \( Z = k \), viewing the unit map \( k \rightarrow A \) as an embedding. (In the literature, central simple \( k \)-algebras are often also understood to be finite dimensional, but we will not assume this here.)

(a) Show that if \( A \in \text{Alg}_k \) is central simple and \( B \in \text{Alg}_k \) is arbitrary, then the ideals of the algebra \( A \otimes B \) are exactly the subspaces of the form \( A \otimes I \), where \( I \) is an ideal of \( B \).
can be realized as the skew polynomial algebra \( O \) \((x, y, \sigma) \) of \( A \). Show that \( \text{End}_k(V) \) is central simple if and only if \( \dim_k V < \infty \). Conclude from (a) and Exercise 1.1.10(e) that the ideals of the matrix algebra \( \text{Mat}_n(B) \) are exactly the subspaces \( \text{Mat}_n(I) \), where \( I \) is an ideal of \( B \).

(c) Conclude from (a) and Exercise 1.1.10(b) that the tensor product of any two central simple algebras is again central simple.

1.1.14 (Weyl algebras). Let \( A_1(k) \) denote the Weyl algebra, with standard algebra generators \( x \) and \( y \) and defining relation \( xy = yx + 1 \) as in (1.15).

(a) Consider the skew polynomial algebra \( B = A[\eta; \delta] \) (Exercise 1.1.6) with \( A = k[\xi] \) the ordinary polynomial algebra and with derivation \( \delta = \frac{d}{d\xi} \). Show that \( \eta \xi = \xi \eta + 1 \) holds in \( B \) and conclude from (1.16) that there is a unique algebra map \( f : A_1(k) \to B \) with \( f(x) = \xi \) and \( f(y) = \eta \). Conclude further that \( f \) is an isomorphism and that the standard monomials \( x^i y^j \) form a \( k \)-basis of \( A_1(k) \). Finally, conclude from Exercise 1.1.6 that \( A_1(k) \) is a noetherian domain.

(b) Assuming \( \text{char} k = 0 \), show that \( A_1(k) \) is central simple in the sense of Exercise 1.1.13. Conclude from Exercise 1.1.13 that the algebra \( A_n(k) := A_1(k)^{\otimes n} \) is central simple for every positive integer \( n \); this algebra is called the \( n \)th Weyl algebra over \( k \).

(c) Now let \( \text{char} k = p > 0 \) and put \( \mathcal{Z} := \mathcal{Z}(A_1(k)) \). Show that \( \mathcal{Z} = k[x^p, y^p] \) is a polynomial algebra over \( k \) and that \( A_1(k) \equiv \mathcal{Z}^{\otimes p^2} \) as \( \mathcal{Z} \)-module: the standard monomials \( x^i y^j \) with \( 0 \leq i, j < p \) form a \( \mathcal{Z} \)-basis of \( A_1(k) \).

1.1.15 (Quantum plane and quantum torus). Fix a scalar \( q \in k^\times \) and consider the following algebra, called the quantum plane.

\[
O_q(k^2) \overset{\text{def}}{=} k(x, y)/(xy - qyx).
\]

As in the case of the Weyl algebra \( A_1(k) \), denote the images of \( x \) and \( y \) in \( O_q(k^2) \) by \( x \) and \( y \) as well; so \( xy = qyx \) holds in \( O_q(k^2) \).

(a) Adapt the method of Exercise 1.1.14(a) to show that the quantum plane can be realized as the skew polynomial algebra \( O_q(k^2) \equiv k[x][y; \sigma] \), where \( k[x] \) is the ordinary polynomial algebra and \( \sigma \in \text{Aut}_{k[x]}(k[x]) \) is given by \( \sigma(x) = q^{-1}x \). Conclude from Exercise 1.1.6 that \( O_q(k^2) \) is a noetherian domain.

(b) Observe that \( \sigma \) extends to an automorphism of the Laurent polynomial algebra \( k[x^{\pm 1}] \). Using this fact, show that there is a tower of \( k \)-algebras,

\[
O_q(k^2) \equiv k[x][y; \sigma] \subseteq k[x^{\pm 1}][y; \sigma] \subseteq O_q((k^\times)^2) \overset{\text{def}}{=} k[x^{\pm 1}][y^{\pm 1}; \sigma],
\]

where the last algebra is a skew Laurent polynomial algebra (Exercise 1.1.6). The algebra \( O_q((k^\times)^2) \) is called a quantum torus.
(c) Show: if the parameter \( q \) is not a root of unity, then \( O_q((k^x)^2) \) is a central simple \( k \)-algebra. Conclude that every nonzero ideal of the quantum plane \( O_q(k^2) \) contains some standard monomial \( x^i y^j \).

(d) If \( q \) is a root of unity of order \( n \), then show that \( \mathcal{Z} := \mathcal{Z}(O_q((k^x)^2)) \) is a Laurent polynomial algebra in the two variables \( x^{ zn}, y^{ zn} \) and the standard monomials \( x^i y^j \) with \( 0 \leq i, j < n \) form a basis of \( O_q((k^x)^2) \) as module over \( \mathcal{Z} \).

1.2. Representations

By definition, a **representation** of a \( k \)-algebra \( A \) is an algebra homomorphism \( \rho: A \to \text{End}_k(V) \) with \( V \in \text{Vect}_k \). If \( \dim_k V = n \) is finite, then the representation is called a finite-dimensional representation and \( n \) is referred to as its **dimension** or the **degree**. We will usually denote the operator \( \rho(a) \) by \( a_V \); so

\[
\rho: A \longrightarrow \text{End}_k(V)
\]

\[
\begin{array}{ccc}
\omega & \omega \\
a & \longmapsto & a_V
\end{array}
\]

(1.20)

The map \( \rho \) is often de-emphasized and the vector space \( V \) is referred to as a representation of \( A \). For example, instead of \( \text{Ker} \rho \), we will usually write

\[
\text{Ker} V \overset{\text{def}}{=} \{a \in A \mid a_V = 0\}.
\]

Representations with kernel 0 are called **faithful**. The image \( \rho(A) \) of a representation (1.20) will be written as \( A_V \); so

\[
A/\text{Ker} V \cong A_V \subseteq \text{End}_k(V).
\]

_Throughout the remainder of this section, \( A \) will denote an arbitrary \( k \)-algebra unless explicitly specified otherwise._

1.2.1. The Category \( \text{Rep} A \) and First Examples

As was explained in §1.1.3, representations of \( A \) are essentially the same as left \( A \)-modules: every representation \( A \to \text{End}_k(V) \) gives rise to a left \( A \)-module action \( A \otimes V \to V \) and conversely. This connection enables us to transfer familiar notions from the theory of modules into the context of representations. Thus, we may speak of **subrepresentations**, **quotients** and **direct sums** of representations and also of **homomorphisms**, **isomorphisms**, etc., of representations by simply using the corresponding definitions for modules. For example, a homomorphism from a representation \( \rho: A \to \text{End}_k(V) \) to a representation \( \rho': A \to \text{End}_k(V') \) is given by an \( A \)-module homomorphism \( f: V \to V' \), that is, is a \( k \)-linear map satisfying

\[
(1.21) \quad \rho'(a) \circ f = f \circ \rho(a) \quad (a \in A).
\]
This condition is sometimes stated as “$f$ intertwines $\rho$ and $\rho’.” Thus, the representations of an algebra $A$ form a category, $\text{Rep} \ A$, that is equivalent to the category of left $A$-modules:

$$\text{Rep} \ A \equiv \text{AMod}$$

An isomorphism $\rho \sim \rho’$ in $\text{Rep} \ A$ is given by an isomorphism $f : V \sim V’$ in $\text{Vect}_k$ satisfying the intertwining condition (1.21), which amounts to commutativity of the following diagram in $\text{Alg}_k$:

$$\begin{array}{ccc}
\text{End}_k(V) & \overset{\sim}{\longrightarrow} & \text{End}_k(V') \\
\downarrow_{\rho} & & \downarrow_{\rho'} \\
A & & A
\end{array}$$

(1.22)

Here, $f \circ f^{-1} = f \circ (f^{-1})^*$ in the notation of §B.2.1. Isomorphic representations are also called equivalent, and the symbol $\equiv$ is used for equivalence or isomorphism of representations. In the following, we shall freely use module theoretic terminology and notation for representations. For example, $\rho(a)(v) = a \cdot v$ will usually be written as $a \cdot v$ or $av$.

**Example 1.6** (Regular representations). The representation of $A$ that corresponds to the module $A_{\text{reg}} \in A\text{Mod}$ (Example 1.2) is called the regular representation of $A$; it is given by the algebra map $\rho_{\text{reg}} : A \to \text{End}_k(A)$ with

$$\rho_{\text{reg}}(a) = a_A := (b \mapsto ab) \quad (a, b \in A).$$

As in Example 1.2, we may also consider the right regular $A$-module as well as the regular $(A, A)$-bimodule; these correspond to representations $A^{\text{op}} \to \text{End}_k(A)$ and $A \otimes A^{\text{op}} \to \text{End}_k(A)$, respectively.

**Example 1.7** (The polynomial algebra). Let $A = \mathbb{k}[t]$ be the ordinary polynomial algebra. By (1.6) representations $\rho : \mathbb{k}[t] \to \text{End}_k(V)$, for a given $V \in \text{Vect}_k$, are in bijection with linear operators $\tau \in \text{End}_k(V)$ via $\rho(t) = \tau$. For a fixed positive integer $n$, we may describe the equivalence classes of $n$-dimensional representations of $\mathbb{k}[t]$ as follows. For any such representation, given by $V \in \text{Vect}_k$ and an operator $\tau \in \text{End}_k(V)$, we may choose an isomorphism $V \sim \mathbb{k}^n$ in $\text{Vect}_k$. This isomorphism and the resulting isomorphism $\text{End}_k(V) \cong \text{Mat}_n(\mathbb{k})$ allow us to replace $V$ by $\mathbb{k}^n$ and $\tau$ by a matrix $T \in \text{Mat}_n(\mathbb{k})$ without altering the isomorphism type. By (1.22) two representations of $\mathbb{k}[t]$ that are given by $T, T' \in \text{Mat}_n(\mathbb{k})$ are equivalent if and only if the matrices $T$ and $T'$ are conjugate to each other, that is, $T' = g T g^{-1}$ for some $g \in \text{GL}_n(\mathbb{k}) = \text{Mat}_n(\mathbb{k})^\times$. Thus, let $\text{GL}_n(\mathbb{k}) \setminus \text{Mat}_n(\mathbb{k})$ denote the set of orbits for the conjugation action $\text{GL}_n(\mathbb{k}) \curvearrowright \text{Mat}_n(\mathbb{k})$, we have a bijection of sets,

$$\begin{array}{ccc}
\left\{ \text{n-dimensional representations of } \mathbb{k}[t] \right\} & \overset{\sim}{\longrightarrow} & \text{GL}_n(\mathbb{k}) \setminus \text{Mat}_n(\mathbb{k}).
\end{array}$$

(1.23)
From linear algebra, we further know that a full representative set of the $\text{GL}_n(\mathbb{k})$-orbits in $\text{Mat}_n(\mathbb{k})$ is given by the matrices in **rational canonical form** or in **Jordan canonical form** over some algebraic closure of $\mathbb{k}$, up to a permutation of the Jordan blocks; see [64, Chapter 12, Theorems 16 and 23].

**Example 1.8** (Representations of the Weyl algebra). In view of (1.16), representations of the Weyl algebra $A_1(\mathbb{k}) = \mathbb{k}(x, y)/(yx - xy - 1)$ are given by a $V \in \text{Vect}_\mathbb{k}$ and a pair $(a, b) \in \text{End}_\mathbb{k}(V)^2$ satisfying the relation $ba = ab + \text{Id}_V$. As in Example 1.7, it follows from (1.22) that two such pairs $(a, b), (a', b') \in \text{End}_\mathbb{k}(V)^2$ yield equivalent representations if and only if

$$(a', b') = g(a, b) = (gag^{-1}, gbg^{-1})$$

for some $g \in \text{GL}(V) = \text{End}_\mathbb{k}(V)^X$, the group of invertible linear transformations of $V$. If $\text{char } \mathbb{k} = 0$, then $A_1(\mathbb{k})$ has no nonzero finite-dimensional representations $V$. Indeed, if $\dim_\mathbb{k} V < \infty$, then we may take the trace of both sides of the equation $ba = ab + \text{Id}_V$ to obtain that $\text{trace}(ba) = \text{trace}(ab) + \dim_\mathbb{k} V$. Since $\text{trace}(ab) = \text{trace}(ba)$ for any $a, b \in \text{End}_\mathbb{k}(V)$, this forces $\dim_\mathbb{k} V = 0$. The **standard representation** of $A_1(\mathbb{k})$, for any base field $\mathbb{k}$, is constructed by taking $V = \mathbb{k}[t]$, the polynomial algebra, and the two $\mathbb{k}$-linear endomorphisms of $\mathbb{k}[t]$ that are given by multiplication with the variable $t$ and by formal differentiation, $\frac{d}{dt}$. Denoting the former by just $t$, the product rule gives the relation $\frac{d}{dt} t = t \frac{d}{dt} + \text{Id}_{\mathbb{k}[t]}$. Thus, we obtain a representation $A_1(\mathbb{k}) \to \text{End}_\mathbb{k}(\mathbb{k}[t])$ with $x \mapsto t$ and $y \mapsto \frac{d}{dt}$. It is elementary to see that the operators $t^i \frac{d^j}{dt^j} \in \text{End}_\mathbb{k}(\mathbb{k}[t])$ $(i, j \in \mathbb{Z}_+)$ are $\mathbb{k}$-linearly independent if $\text{char } \mathbb{k} = 0$ (Exercise 1.2.9). It follows that the standard monomials $x^i y^j$ form a $\mathbb{k}$-basis of $A_1(\mathbb{k})$ when $\text{char } \mathbb{k} = 0$. This also holds for general $\mathbb{k}$ (Exercise 1.1.14 or Example D.3).

### 1.2.2. Changing the Algebra or the Base Field

In studying the representations of a given $\mathbb{k}$-algebra $A$, it is often useful to extend the base field $\mathbb{k}$—things tend to become simpler over an algebraically closed or at least sufficiently large field—and to take advantage of any available information concerning the representations of certain related algebras such as subalgebras or homomorphic images of $A$. Here we describe some standard ways to go about doing this. This material may appear dry and technical at first; it can be skipped or only briefly skimmed at a first reading and referred to later as needed.

**Pulling Back: Restriction.** Suppose we are given a $\mathbb{k}$-algebra map $\phi: A \to B$. Then any representation $\rho: B \to \text{End}_\mathbb{k}(V), \rho \mapsto b_V$, gives rise to a representation $\phi^*\rho := \rho \circ \phi: A \to \text{End}_\mathbb{k}(V); \; a_V = \phi(a)V$ for $a \in A$. We will refer to this process as **pulling back** the representation $\rho$ along $\phi$; the representation $\phi^*\rho$ of $A$ is also called the **restriction** of $\rho$ from $B$ to $A$. The “restriction” terminology is especially intuitive in case $A$ is a subalgebra of $B$ and $\phi$ is the embedding, or if $\phi$
is at least a monomorphism, but it is also used for general $\phi$. If $\phi$ is surjective, then $\phi'(\rho)$ is sometimes referred to as the inflation of $\rho$ along $\phi$. In keeping with the general tendency to emphasize $V$ over the map $\rho$, the pullback $\phi^*(\rho)$ is often denoted by $\phi^*V$. When $\phi$ is understood, we will also write $V\downarrow_A$ or $\text{Res}_A^B V$. The process of restricting representations along a given algebra map clearly is functorial: any morphism $\rho \rightarrow \rho'$ in $\text{Rep} B$ gives rise to a homomorphism $\phi^*(\rho) \rightarrow \phi^*(\rho')$ in $\text{Rep} A$, because the intertwining condition (1.21) for $\rho$ and $\rho'$ is clearly inherited by $\phi^*(\rho)$ and $\phi^*(\rho')$. In this way, we obtain the restriction functor,

$$\phi^* = \text{Res}_A^B : \text{Rep} B \rightarrow \text{Rep} A$$

**Pushing Forward: Induction and Coinduction.** In the other direction, we may also “push forward” representations along an algebra map $\phi : A \rightarrow B$. In fact, there are two principal ways to do this. First, for any $V \in \text{Rep} A$, the induced representation of $B$ is defined by

$$\text{Ind}_A^B V \overset{\text{def}}{=} B \otimes_A V$$

On the right, $B$ carries the $(B, A)$-bimodule structure that comes from the regular $(B, B)$-bimodule structure via $b.b'.a := bb'\phi(a)$. As in §B.1.2, this allows us to form the tensor product $B \otimes_A V$ and equip it with the left $B$-module action

$$b.(b' \otimes v) := bb' \otimes v.$$  

This makes $\text{Ind}_A^B V$ a representation of $B$. Alternative notations for $\text{Ind}_A^B V$ include $\phi_*V$ and $V\uparrow^B_A$. Again, this construction is functorial: if $f : V \rightarrow V'$ is a morphism in $\text{Rep} A$, then $\text{Ind}_A^B f := \text{Id}_B \otimes f : \text{Ind}_A^B V \rightarrow \text{Ind}_A^B V'$ is a morphism in $\text{Rep} B$. All this behaves well with respect to composition and identity morphisms; so induction gives a functor,

$$\phi_* = \text{Ind}_A^B : \text{Rep} A \rightarrow \text{Rep} B.$$  

Similarly, we may use the $(A, B)$-bimodule structure of $B$ that is given by $a.b'.b := \phi(a)b'b$ to form $\text{Hom}_A(B, V)$ and view it as left $B$-module as in §B.2.1:

$$(b.f)(b') = f(b'b).$$

The resulting representation of $B$ is called the coinduced representation:

$$\text{Coind}_A^B V \overset{\text{def}}{=} \text{Hom}_A(B, V)$$

If $f : V \rightarrow V'$ is a morphism in $\text{Rep} A$, then $\text{Coind}_A^B f := f_* : \text{Coind}_A^B V \rightarrow \text{Coind}_A^B V'$, $g \mapsto f \circ g$, is a morphism in $\text{Rep} B$. The reader will have no difficulty confirming that this gives a functor,

$$\text{Coind}_A^B : \text{Rep} A \rightarrow \text{Rep} B.$$
We shall primarily work with induction below. In some situations that we shall encounter, there is an isomorphism of functors $\text{Coind}_A^B \cong \text{Ind}_A^B$; see Exercise 2.2.6 and Proposition 3.4.

**Adjointness Relations.** It turns out that the functors $\text{Ind}_A^B$ and $\text{Coind}_A^B$ are left and right adjoint to $\text{Res}_A^B$, respectively, in the sense of Section A.4. These abstract relations have very useful consequences; see Exercises 1.2.7 and 1.2.8, for example. The isomorphism in part (a) of the proposition below, and various consequences thereof, are often referred to as **Frobenius reciprocity**.

**Proposition 1.9.** Let $\phi : A \to B$ be a map in $\text{Alg}_{k}$. Then, for any $V \in \text{Rep } A$ and $W \in \text{Rep } B$, there are natural isomorphisms in $\text{Vect}_{k}$,

(a) $\text{Hom}_B(\text{Ind}_A^B V, W) \cong \text{Hom}_A(V, \text{Res}_A^B W)$ and

(b) $\text{Hom}_B(W, \text{Coind}_A^B V) \cong \text{Hom}_A(\text{Res}_A^B W, V)$.

**Proof.** Both parts follow from $\text{Hom} \otimes$ adjunction (B.16).

For (a), we use the $(B, A)$-bimodule structure of $B$ that was explained above to form $\text{Hom}_B(B, W)$ and equip it with the left $A$-action $(a.f)(b) := f(b\phi(a))$. In particular, $(a.f)(1) = f(\phi(a)) = \phi(a).f(1)$; so the map $f \mapsto f(1)$ is an isomorphism $\text{Hom}_B(B, W) \cong \text{Res}_A^B W$ in $\text{Rep } A$. Therefore,

\[
\text{Hom}_B(\text{Ind}_A^B V, W) = \text{Hom}_B(B \otimes_A V, W)
\]

\[
\cong \text{Hom}_A(V, \text{Hom}_B(B, W)) \quad \text{(B.16)}
\]

\[
\cong \text{Hom}_A(V, \text{Res}_A^B W).
\]

Tracking a homomorphism $f \in \text{Hom}_A(V, \text{Res}_A^B W)$ through the above isomorphisms, one obtains the map in $\text{Hom}_B(\text{Ind}_A^B V, W)$ that is given by $b \otimes v \mapsto b.f(v)$ for $b \in B$ and $v \in V$.

Part (b) uses the above $(A, B)$-bimodule structure of $B$ and the standard $B$-module isomorphism $B \otimes_B W \cong W$. This isomorphism restricts to an isomorphism $B \otimes_B W \cong \text{Res}_A^B W$ in $\text{Rep } A$, giving

\[
\text{Hom}_B(W, \text{Coind}_A^B V) = \text{Hom}_B(W, \text{Hom}_A(B, V))
\]

\[
\cong \text{Hom}_A(B \otimes_B W, V) \quad \text{(B.16)}
\]

\[
\cong \text{Hom}_A(\text{Res}_A^B W, V).
\]

**Twisting.** For a given $V \in \text{Rep } A$, we may use restriction or induction along some $\alpha \in \text{Aut}_{\text{Alg}_{k}}(A)$ to obtain a new representation of $A$, called a **twist** of $V$. For
1.2. Representations

restriction, each \( a \in A \) acts via \( \alpha(a)_V \) on \( \alpha^*V = V \). Using induction instead, we have \( \alpha_*,V = A \otimes_A V \cong V \), with \( 1 \otimes v \leftrightarrow v \), and

\[
a.(1 \otimes v) = a.1 \otimes v = a \otimes v = 1.a^{-1}(a) \otimes v = 1 \otimes a^{-1}(a).v.
\]

Thus, identifying \( \alpha_*V \) with \( V \) as above, each \( a \in A \) acts via \( a^{-1}(a)_V \). Alternatively, putting \( a^*V := \alpha_*V \) and \( a^*v := 1 \otimes v \), we obtain an isomorphism \( a^* : V \xrightarrow{\sim} a^*V \) in \( \text{Vect}_k \) and the result of the above calculation can be restated as \( a^*v = a(ab^{-1}(a).v) \) or, equivalently,

\[
(1.24) \quad a^*(a.v) = a^*(a).
\]

Extending the Base Field. For any field extension \( K/k \) and any representation \( \rho : A \to \text{End}_k(V) \) of \( A \), we may consider the representation of the \( K \)-algebra \( K \otimes A \) that is obtained from \( \rho \) by “extension of scalars.” The resulting representation may be described as the representation \( \text{Ind}_A^{K \otimes A} V = (K \otimes A) \otimes_A V \cong K \otimes V \) that comes from the \( k \)-algebra map \( A \to K \otimes A, a \mapsto 1 \otimes a \). However, we view \( K \otimes A \) as a \( K \)-algebra as in §1.1.1, moving from \( \text{Alg}_k \) to \( \text{Alg}_K \) in the process. Explicitly, the action of \( K \otimes A \) on \( K \otimes V \) is given by

\[
(\lambda \otimes a).(\lambda' \otimes v) = \lambda \lambda' \otimes a.v
\]

for \( \lambda, \lambda' \in K, a \in A \) and \( v \in V \); equivalently, in terms of algebra homomorphisms,

\[
(1.25) \quad K \otimes \rho : K \otimes A \xrightarrow{\text{Id}_K \otimes \rho} K \otimes \text{End}_k(V) \xrightarrow{\text{can.}} \text{End}_K(K \otimes V)
\]

\[
\lambda \otimes f \xrightarrow{\rho_{\text{reg}}(\lambda) \otimes f}
\]

The “canonical” map in (1.25) is a special case of (B.27); this map is always injective, and it is an isomorphism if \( V \) is finite dimensional or the field extension \( K/k \) is finite.

Example 1.10 (The polynomial algebra). Recall from Example 1.7 that the equivalence classes of \( n \)-dimensional representations of \( k[t] \) are in bijection with the set of orbits for the conjugation action \( \text{GL}_n(k) \subset \text{Mat}_n(k) \). It is a standard fact from linear algebra (e.g., [64, Chapter 12, Corollary 18]) that if two matrices \( T, T' \in \text{Mat}_n(k) \) belong to the same \( \text{GL}_n(K) \)-orbit for some field extension \( K/k \), then \( T, T' \) also belong to the same \( \text{GL}_n(K) \)-orbit. In other words, if \( V \) and \( V' \) are finite-dimensional representations of \( k[t] \) such that \( K \otimes V \cong K \otimes V' \) in \( \text{Rep} K[t] \) for some field extension \( K/k \), then we must have \( V \cong V' \) to start with. This does in fact hold for any \( k \)-algebra in place of \( k[t] \), by the Noether-Deuring Theorem (Exercise 1.2.5).
1.2.3. Irreducible Representations

A representation \( \rho : A \to \text{End}_k(V) \), \( a \mapsto a_V \), is said to be \textit{irreducible}\(^3\) if \( V \) is an irreducible \( A \)-module. Explicitly, this means that \( V \neq 0 \) and no \( k \)-subspace of \( V \) other than \( 0 \) and \( V \) is stable under all operators \( a_V \) with \( a \in A \); equivalently, it is impossible to find a \( k \)-basis of \( V \) such that the matrices of all operators \( a_V \) have block upper triangular form

\[
\begin{pmatrix}
* & * \\
0 & *
\end{pmatrix}
\]

**Example 1.11** (Division algebras). Recall that a \textit{division} \( k \)-\textit{algebra} is a \( k \)-algebra \( D \neq 0 \) whose nonzero elements are all invertible: \( D^\times = D \setminus \{0\} \). Representations of \( D \) are the same as left \( D \)-vector spaces, and a representation \( V \) is irreducible if and only if \( \dim_k D \cdot V = 1 \). Thus, up to equivalence, the regular representation of \( D \) is the only irreducible representation of \( D \).

**Example 1.12** (Tautological representation of \( \text{End}_k(D(V)) \)). For any \( k \)-vector space \( V \neq 0 \), the representation of the algebra \( \text{End}_k(V) \) that is given by the identity map \( \text{End}_k(V) \to \text{End}_k(V) \) is irreducible. For, if \( u, v \in V \) are given, with \( u \neq 0 \), then there exists \( f \in \text{End}_k(V) \) such that \( f(u) = v \). Therefore, any nonzero subspace of \( V \) that is stable under all \( f \in \text{End}_k(V) \) must contain all of \( V \). The foregoing applies verbatim to any nonzero representation \( V \) of a division \( k \)-algebra \( D \): the embedding \( \text{End}_D(V) \hookrightarrow \text{End}_k(V) \) is an irreducible representation of the algebra \( \text{End}_D(V) \). If \( \dim_D V < \infty \), then this representation is in fact the only irreducible representation of \( \text{End}_D(V) \) up to equivalence; this is a consequence of Wedderburn’s Structure Theorem (§1.4.4).

**Example 1.13** (The standard representation of the Weyl algebra). Recall from Example 1.8 that the standard representation of \( A_1(k) \) is the algebra homomorphism \( A_1(k) \to \text{End}_k(V) \), with \( V = k[t] \), that is given by \( x_V = t \cdot \) and \( y_V = \frac{d}{dt} \). If \( \text{char} \ k = 0 \), then the standard representation is irreducible. To see this, let \( U \subseteq V \) be any nonzero subrepresentation of \( V \) and let \( 0 \neq f \in U \) be a polynomial of minimal degree among all nonzero polynomials in \( U \). Then \( f \in k^\times \); for, if \( \text{deg} f > 0 \), then \( 0 \neq \frac{d}{dt} f = y \cdot f \in U \) and \( \frac{d}{dt} f \) has smaller degree than \( f \). Therefore, \( k \subseteq U \) and repeated application of \( x_V \) gives that all \( k t^n \subseteq U \). This shows that \( U = V \), proving irreducibility of the standard representation for \( \text{char} \ k = 0 \). There are many more irreducible representations of \( A_1(k) \) in characteristic \( 0 \); see Block [16]. For irreducible representations of \( A_1(k) \) in positive characteristics, see Exercise 1.2.9.

\(^3\)Irreducible representations are also called \textit{simple} and they are often informally referred to as “irreps.”
One of the principal, albeit often unachievable, goals of representation theory is to provide, for a given $k$-algebra $A$, a good description of the following set:

$$\text{Irr } A \overset{\text{def}}{=} \text{ the set of equivalence classes of irreducible representations of } A$$

Of course, $\text{Irr } A$ can also be thought of as the set of isomorphism classes of irreducible left $A$-modules. We will generally use $\text{Irr } A$ to denote a full set of representatives of the equivalence classes and $S \in \text{Irr } A$ will indicate that $S$ is an irreducible representation of $A$. To see that $\text{Irr } A$ is indeed a set, we observe that every irreducible representation of $A$ is a homomorphic image of the regular representation $A_{\text{reg}}$. To wit:

**Lemma 1.14.** A full representative set for $\text{Irr } A$ is furnished by the non-equivalent factors $A_{\text{reg}}/L$, where $L$ is a maximal left ideal of $A$. In particular, $\dim_k S \leq \dim_k A$ for all $S \in \text{Irr } A$.

**Proof.** If $S$ is an irreducible representation of $A$, then any $0 \neq s \in S$ gives rise to a homomorphism of representations $f : A_{\text{reg}} \rightarrow S$, $a \mapsto as$. Since $\text{Im } f$ is a nonzero subrepresentation of $S$, it must be equal to $S$. Thus, $f$ is an epimorphism and $S \cong A_{\text{reg}}/L$ with $L = \text{Ker } f$; this is a maximal left ideal of $A$ by irreducibility of $S$. Conversely, all factors of the form $A_{\text{reg}}/L$, where $L$ is a maximal left ideal of $A$, are irreducible left $A$-modules, and hence we may select our equivalence classes of irreducible representations of $A$ from the set of these factors. The last assertion of the lemma is now clear. \hfill \square

**Example 1.15** (The polynomial algebra). By Example 1.7, a representation $V \in \text{Rep } k[t]$ corresponds to an endomorphisms $\tau = t_V \in \text{End}_k(V)$. Lemma 1.14 further tells us that irreducible representations of $k[t]$ have the form $V \cong k[t]_{\text{reg}}/L$, where $L$ is a maximal ideal of $k[t]$; so $L = (m(t))$ for a unique monic irreducible polynomial $m(t) \in k[t]$. Note that $m(t)$ is the characteristic polynomial of $\tau = t_V$. Thus, an irreducible representation of $k[t]$ is given by a finite-dimensional $V \in \text{Vect}_k$ and an endomorphism $\tau \in \text{End}_k(V)$ whose characteristic polynomial is irreducible. In particular, if $k$ is algebraically closed, then all irreducible representations of $k[t]$ are 1-dimensional.

In sharp contrast to the polynomial algebra, the Weyl algebra $A_1(k)$ has no nonzero finite-dimensional representations at all if $\text{char } k = 0$ (Example 1.8). In general, “most” irreducible representations of a typical infinite-dimensional non-commutative algebra will tend to be infinite dimensional, but the finite-dimensional ones, insofar as they exist, are of particular interest. Therefore, for any $k$-algebra $A$, we denote the full subcategory of $\text{Rep } A$ whose objects are the finite-dimensional representations of $A$ by

$$\text{Rep}_{\text{fin }} A$$
and we also define
\[
\text{Irr}_{\text{fin}} A \overset{\text{def}}{=} \{ S \in \text{Irr } A \mid \dim_k S < \infty \}
\]

1.2.4. Composition Series

Every nonzero \( V \in \text{Rep}_{\text{fin}} A \) can be assembled from irreducible pieces in the following way. To start, pick some irreducible subrepresentation \( V_1 \subseteq V \); any nonzero subrepresentation of minimal dimension will do. If \( V_1 \neq V \), then we may similarly choose an irreducible subrepresentation of \( V/V_1 \), which will have the form \( V_2/V_1 \) for some subrepresentation \( V_2 \subseteq V \). If \( V_2 \neq V \), then we continue in the same manner. Since \( V \) is finite dimensional, the process must stop after finitely many steps, resulting in a finite chain
\[
0 = V_0 \subset V_1 \subset \cdots \subset V_l = V
\]
of subrepresentations \( V_i \) such that all \( V_i/V_{i-1} \) are irreducible.

An analogous construction can sometimes be carried out even when the representation \( V \in \text{Rep } A \) is infinite dimensional (Exercises 1.2.10, 1.2.11). Any chain of the form (1.26), with irreducible factors \( \overline{V}_i = V_i/V_{i-1} \), is called a composition series of \( V \) and the number \( l \) is called the length of the series. If a composition series (1.26) is given and a \( k \)-basis of \( V \) is assembled from bases of the \( \overline{V}_i \), then the matrices of all operators \( a_V \) (\( a \in A \)) have block upper triangular form, with (possibly infinite) diagonal blocks coming from the irreducible representations \( \overline{V}_i \):

\[
\begin{pmatrix}
\alpha_{\overline{V}_1} & 0 & & \\
& \alpha_{\overline{V}_2} & & \\
& & & \ddots & \\
& & & & \alpha_{\overline{V}_l}
\end{pmatrix}
\]

**Example 1.16** (The polynomial algebra). Let \( V \in \text{Rep}_{\text{fin}} k[t] \) and assume that \( k \) is algebraically closed. Then, in view of Example 1.15, fixing a composition series for \( V \) amounts to the familiar process of choosing a \( k \)-basis of \( V \) such that the matrix of the endomorphism \( t_V \in \text{End}_k(V) \) is upper triangular. The eigenvalues of \( t_V \) occupy the diagonal of the matrix.

**Example 1.17** (Composition series need not exist). If \( A \) is any domain (not necessarily commutative) that is not a division algebra, then the regular representation \( A_{\text{reg}} \) does not have a composition series; in fact, \( A_{\text{reg}} \) does not even contain any irreducible subrepresentations. To see this, observe that subrepresentations of \( A_{\text{reg}} \) are the same as left ideals of \( A \). Moreover, if \( L \) is any nonzero left ideal of \( A \), then there exists some \( 0 \neq a \in L \) with \( a \notin A^X \). Then \( L \supseteq Aa \supsetneq Aa^2 \neq 0 \), showing that \( L \) is not irreducible.
The Jordan-Hölder Theorem

Representations that admit a composition series are said to be of finite length. The reason for this terminology will be clearer shortly. For now, we just remark that finite-length representations of a division algebra $D$ are the same as finite-dimensional left $D$-vector spaces (Example 1.11). For any algebra $A$, the class of all finite-length representations $V \in \text{Rep } A$ behaves quite well in several respects. Most importantly, all composition series of any such $V$ are very much alike; this is the content of the celebrated Jordan-Hölder Theorem, which is stated as part (b) of the theorem below. Part (a) shows that the property of having finite length also transfers well in short exact sequences in $\text{Rep } A$, that is, sequences of morphisms in $\text{Rep } A$ of the form

\[(1.27) \quad 0 \to U \xrightarrow{f} V \xrightarrow{g} W \to 0 \]

with $f$ injective, $g$ surjective, and $\text{Im } f = \text{Ker } g$ (as in §B.1.1).

**Theorem 1.18.**  
(a) Given a short exact sequence (1.27) in $\text{Rep } A$, the representation $V$ has finite length if and only if both $U$ and $W$ do.

(b) Let $0 = V_0 \subset V_1 \subset \cdots \subset V_l = V$ and $0 = V'_0 \subset V'_1 \subset \cdots \subset V'_l = V$ be two composition series of $V \in \text{Rep } A$. Then $l = l'$ and there exists a permutation $s$ of $\{1, \ldots, l\}$ such that $V_i/V_{i-1} \cong V'_{s(i)}/V'_{s(i)-1}$ for all $i$.

**Proof.** (a) First, assume that $U$ and $W$ have finite length and fix composition series $0 = U_0 \subset U_1 \subset \cdots \subset U_r = U$ and $0 = W_0 \subset W_1 \subset \cdots \subset W_s = W$. These series can be spliced together to obtain a composition series for $V$ as follows. Put $X_i = f(U_i)$ and $Y_j = g^{-1}(W_j)$. Then $X_i/X_{i-1} \cong U_i/U_{i-1}$ via $f$ and $Y_j/Y_{j-1} \cong W_j/W_{j-1}$ via $g$. Thus, the following is a composition series of $V$:

\[(1.28) \quad 0 = X_0 \subset X_1 \subset \cdots \subset X_r = Y_0 \subset Y_1 \subset \cdots \subset Y_s = V. \]

Conversely, assume that $V$ has a composition series (1.26). Put $U_i = f^{-1}(V_i)$ and observe that $U_i/U_{i-1} \hookrightarrow V_i/V_{i-1}$ via $f$; so each factor $U_i/U_{i-1}$ is either 0 or irreducible (in fact, isomorphic to $V_i/V_{i-1}$). Therefore, deleting repetitions from the chain $0 = U_0 \subset U_1 \subset \cdots \subset U_r = U$ if necessary, we obtain a composition series for $U$. Similarly, putting $W_i = g(V_i)$, each factor $W_i/W_{i-1}$ is a homomorphic image of $V_i/V_{i-1}$, and so we may again conclude that $W_i/W_{i-1}$ is either 0 or irreducible. Thus, we obtain the desired composition series of $W$ by deleting superfluous members from the chain $0 = W_0 \subset W_1 \subset \cdots \subset W_s = W$. This proves (a).

In preparation for the proof of (b), let us also observe that if $U \neq 0$ or, equivalently, $\text{Ker } g \neq 0$, then some factor $W_i/W_{i-1}$ will definitely be 0 in the above construction. Indeed, there is an $i$ such that $\text{Ker } g \nsubseteq V_i$ but $\text{Ker } g \nsubseteq V_{i-1}$. Irreducibility of $V_i/V_{i-1}$ forces $V_i = V_{i-1} + \text{Ker } g$ and so $W_i = W_{i-1}$. Therefore, $W$ has a composition series of shorter length than the given composition series of $V$. 

(b) We will argue by induction on \( \ell(V) \), which we define to be the minimum length of any composition series of \( V \). If \( \ell(V) = 0 \), then \( V = 0 \) and the theorem is clear. From now on assume that \( V \neq 0 \). For each subrepresentation \( 0 \neq U \subseteq V \), the factor \( V/U \) also has a composition series by part (a) and the observation in the preceding paragraph tells us that \( \ell(V/U) < \ell(V) \). Thus, by induction, the theorem holds for all factors \( V/U \) with \( U \neq 0 \).

Now consider two composition series as in the theorem. If \( V_1 = V'_1 \), then
\[
0 = V_1/V_1 \subset V_2/V_1 \subset \cdots \subset V_l/V_1 = V/V_1
\]
and
\[
0 = V'_1/V_1 \subset V'_2/V_1 \subset \cdots \subset V'_l/V_1 = V/V_1
\]
are two composition series of \( V/V_1 \) with factors isomorphic to \( V_i/V_{i-1} \) \( (i = 2, \ldots, l) \) and \( V'_j/V'_{j-1} \) \( (j = 2, \ldots, l') \), respectively. Thus the result follows in this case, because the theorem holds for \( V/V_1 \). So assume that \( V_1 \neq V'_1 \) and note that this implies \( V_1 \cap V'_1 = 0 \) by irreducibility of \( V_1 \) and \( V'_1 \). First, let us consider composition series for \( V/V_1 \). One is already provided by (1.29). To build another, put \( U = V_1 \oplus V'_1 \subseteq V \) and fix a composition series for \( V/U \), say \( 0 \subset U_1/U \subset \cdots \subset U_s/U = V/U \). Then we obtain the following composition series for \( V/V_1 \):
\[
0 \subset U/V_1 \subset U_1/V_1 \subset \cdots \subset U_s/V_1 = V/V_1
\]
The first factor of this series is \( U/V_1 \cong V'_1 \) and the remaining factors are isomorphic to \( U_i/U_{i-1} \) \( (i = 1, \ldots, s) \), with \( U_0 := U \). Since the theorem holds for \( V/V_1 \), the collections of factors in (1.29) and (1.30), with multiplicities, are the same up to isomorphism. Adding \( V_1 \) to both collections, we conclude that there is a bijective correspondence between the following two families of irreducible representations, with corresponding representations being isomorphic:
\[
V_i/V_{i-1} \quad (i = 1, \ldots, l) \quad \text{and} \quad V_1, V'_1, U_i/U_{i-1} \quad (i = 1, \ldots, s).
\]
Considering \( V/V'_1 \) in place of \( V/V_1 \), we similarly obtain a bijection between the family on the right and \( V'_j/V'_{j-1} \) \( (j = 1, \ldots, l') \), which implies the theorem. \( \square \)

**Length**

In light of the Jordan-Hölder Theorem, we may define the **length** of any finite-length representation \( V \in \text{Rep} A \) by

\[
\text{length } V \overset{\text{def}}{=} \text{the common length of all composition series of } V
\]
If \( V \) has no composition series, then we put \( \text{length } V = \infty \). Thus, \( \text{length } V = 0 \) means that \( V = 0 \) and \( \text{length } V = 1 \) says that \( V \) is irreducible. If \( A = D \) is a division algebra \( D \), then \( \text{length } V = \text{dim}_D V \). In general, for any short exact sequence
0 → U → V → W → 0 in \(\text{Rep} A\), we have the following generalization of a standard dimension formula for vector spaces (with the usual rules regarding \(\infty\)):

\[
\text{length } V = \text{length } U + \text{length } W.
\]

To see this, just recall that any two composition series of \(U\) and \(W\) can be spliced together to obtain a composition series for \(V\) as in (1.28).

The Jordan-Hölder Theorem also tells us that, up to isomorphism, the collection of factors \(V_i/V_{i-1} \in \text{Irr } A\) occurring in (1.26) is independent of the particular choice of composition series of \(V\). These factors are called the \textit{composition factors} of \(V\). The number of occurrences, again up to isomorphism, of a given \(S \in \text{Irr } A\) as a composition factor in any composition series of \(V\) is also independent of the choice of series; it is called the \textit{multiplicity} of \(S\) in \(V\). We will write

\[
\mu(S, V) \overset{\text{def}}{=} \text{multiplicity of } S \text{ in } V.
\]

For any finite-length representation \(V \in \text{Rep } A\), we evidently have

\[
\text{length } V = \sum_{S \in \text{Irr } A} \mu(S, V).
\]

Finally, by the same argument as above, (1.31) can be refined to the statement that multiplicities are additive in short exact sequences \(0 \rightarrow U \rightarrow V \rightarrow W \rightarrow 0\) in \(\text{Rep } A\): for every \(S \in \text{Irr } A\),

\[
\mu(S, V) = \mu(S, U) + \mu(S, W).
\]

### 1.2.5. Endomorphism Algebras and Schur’s Lemma

The following general lemma describes the endomorphism algebras of irreducible representations. Although very easy, it will be of great importance in the following.

**Schur’s Lemma.** Let \(S \in \text{Irr } A\). Then every nonzero morphism \(S \rightarrow V\) in \(\text{Rep } A\) is injective and every nonzero morphism \(V \rightarrow S\) is surjective. In particular, \(\text{End}_A(S)\) is a division \(k\)-algebra. If \(S \in \text{Irr}_{\text{fin}} A\), then \(\text{End}_A(S)\) is algebraic over \(k\).

**Proof.** If \(f : S \rightarrow V\) is nonzero, then \(\text{Ker } f\) is a subrepresentation of \(S\) with \(\text{Ker } f \neq S\). Since \(S\) is irreducible, it follows that \(\text{Ker } f = 0\) and so \(f\) is injective. Similarly, for any \(0 \neq f \in \text{Hom}_A(V, S)\), we must have \(\text{Im } f = S\), because \(\text{Im } f\) is a nonzero subrepresentation of \(S\). It follows that any nonzero morphism between irreducible representations of \(A\) is injective as well as surjective, and hence it is an isomorphism. In particular, all nonzero elements of the algebra \(\text{End}_A(S)\) have an inverse, proving that \(\text{End}_A(S)\) is a division \(k\)-algebra.

Finally, if \(S\) is finite dimensional over \(k\), then so is \(\text{End}_k(S)\). Hence, for each \(f \in \text{End}_k(S)\), the powers \(f^i (i \in \mathbb{Z}_+)\) are linearly dependent and so \(f\) satisfies a
nonzero polynomial over \( k \). Consequently, the division algebra \( \text{End}_A(S) \) is algebraic over \( k \). □

We will refer to \( \text{End}_A(S) \) as the **Schur division algebra** of the irreducible representation \( S \) and write

\[
D(S) \overset{\text{def}}{=} \text{End}_A(S)
\]

**The Weak Nullstellensatz.** Algebras \( A \) such that \( D(S) \) is algebraic over \( k \) for all \( S \in \text{Irr} \ A \) are said to satisfy the **weak Nullstellensatz**. See the discussion in §5.6.1 and in Appendix C for the origin of this terminology. Thus, finite-dimensional algebras certainly satisfy the weak Nullstellensatz, because all their irreducible representations are finite dimensional (Lemma 1.14). The weak Nullstellensatz will later also be established, in a more laborious manner, for certain infinite-dimensional algebras (Section 5.6). Exercise 1.2.12 discusses a “quick and dirty” way to obtain the weak Nullstellensatz under the assumption that the cardinality of the base field \( k \) is larger than \( \dim_k A \). In particular, if \( k \) is uncountable, then any affine \( k \)-algebra satisfies the weak Nullstellensatz.

**Splitting Fields.** We will say that the base field \( k \) of a \( k \)-algebra \( A \) is a **splitting field** for \( A \) if \( D(S) = k \) for all \( S \in \text{Irr}_f \ A \). By Schur’s Lemma, this certainly holds if \( k \) is algebraically closed, but often much less is required; see Corollary 4.16 below for an important example. We will elaborate on the significance of the condition \( D(S) = k \) in the next paragraph and again in Proposition 1.36 below.

**Centralizers and Double Centralizers.** The endomorphism algebra \( \text{End}_A(V) \) of an arbitrary \( V \in \text{Rep} \ A \) is the **centralizer** of \( A_V \) in \( \text{End}_k(V) \):

\[
\text{End}_A(V) = \{ f \in \text{End}_k(V) \mid a_V \circ f = f \circ a_V \text{ for all } a \in A \}
\]

\[
= C_{\text{End}_k(V)}(A_V).
\]

The centralizer of \( \text{End}_A(V) \) in \( \text{End}_k(V) \) is called the **bi-commutant** or **double centralizer** of the representation \( V \); it may also be described as the endomorphism algebra of \( V \), viewed as a representation of \( \text{End}_A(V) \) via the inclusion \( \text{End}_A(V) \hookrightarrow \text{End}_k(V) \). Thus, we define

\[
\text{BiEnd}_A(V) \overset{\text{def}}{=} C_{\text{End}_k(V)}(\text{End}_A(V)) = \text{End}_{\text{End}_A(V)}(V)
\]
1.2. Representations

Evidently, \( A_V \subseteq \text{BiEnd}_A(V) \); so we may think of any representation of \( A \) as an algebra map

\[
\rho: A \rightarrow \text{BiEnd}_A(V) \hookrightarrow \text{End}_k(V)
\]

Moreover, \( \text{BiEnd}_A(V) = \text{End}_k(V) \) if and only if \( \text{End}_A(V) \subseteq \mathcal{Z}(\text{End}_k(V)) \). Thus,

\[
(1.34) \quad \text{BiEnd}_A(V) = \text{End}_k(V) \iff \text{End}_A(V) = \mathbb{k} \text{Id}_V.
\]

Consequently, \( \mathbb{k} \) is a splitting field for \( A \) if and only if \( \text{BiEnd}_A(S) = \text{End}_k(S) \) for all \( S \in \text{Irr}_{\text{fin}} A \).

1.2.6. Indecomposable Representations

A nonzero \( V \in \text{Rep} A \) is said to be **indecomposable** if \( V \) cannot be written as a direct sum of nonzero subrepresentations. Irreducible representations are evidently indecomposable, but the converse is far from true. For example, \( A_{\text{reg}} \) is indecomposable for any commutative domain \( A \), because any two nonzero subrepresentations (ideals) of \( A \) intersect nontrivially; the same holds for the field of fractions of \( A \), because its nonzero subrepresentations have nonzero intersection with \( A \).

**Example 1.19** (The polynomial algebra). The Structure Theorem for Modules over PIDs ([64, Chapter 12] or [111, Chapter 3]) yields all indecomposable representation of \( \mathbb{k}[t] \) that are finitely generated: the only infinite-dimensional such representation, up to isomorphism, is \( \mathbb{k}[t]_{\text{reg}} \) and every finite-dimensional such \( V \) is isomorphic to \( \mathbb{k}[t]_{\text{reg}}/(p^r) \) for a unique monic irreducible polynomial \( p \in \mathbb{k}[t] \). The Structure Theorem also tells us that an arbitrary \( V \in \text{Rep}_{\text{fin}} \mathbb{k}[t] \) is the direct sum of indecomposable subrepresentations corresponding to the elementary divisors \( p^r \) of \( V \). This decomposition of \( V \) is unique up to the isomorphism type of the summands and their order in the sum. In this subsection, we will see this holds for any \( \mathbb{k} \)-algebra.

To start with, it is clear that any \( V \in \text{Rep}_{\text{fin}} A \) can be decomposed into a finite direct sum of indecomposable subrepresentations. Indeed, \( V = 0 \) is a direct sum with zero indecomposable summands; and any \( 0 \neq V \in \text{Rep} A \) is either already indecomposable or else \( V = V_1 \oplus V_2 \) for nonzero subrepresentations \( V_i \) which both have a decomposition of the desired form by induction on the dimension. More interestingly, the decomposition of \( V \) thus obtained is essentially unique. This is the content of the following classical theorem, which is usually attributed to Krull and Schmidt. Various generalizations of the theorem also have the names of Remak and/or Azumaya attached, but we shall focus on the case of finite-dimensional representations.
Krull-Schmidt Theorem. Any finite-dimensional representation of an algebra can be decomposed into a finite direct sum of indecomposable subrepresentations and this decomposition is unique up to the order of the summands and up to isomorphism.

More explicitly, the uniqueness statement asserts that if } \bigoplus_{i=1}^{r} V_i \cong \bigoplus_{j=1}^{s} W_j \text{ for indecomposable } V_i, W_j \in \text{Rep}_{\text{fin}} A, \text{ then } r = s \text{ and there is a permutation } s \text{ of the indices such that } V_i \cong W_{s(i)} \text{ for all } i. \text{ The proof will depend on the following lemma.}

Lemma 1.20. Let } V \in \text{Rep}_{\text{fin}} A \text{ be indecomposable. Then each } \phi \in \text{End}_A(V) \text{ is either an automorphism or nilpotent. Furthermore, the nilpotent endomorphisms form an ideal of } \text{End}_A(V).

Proof. Viewing } V \text{ as a representation of the polynomial algebra } \mathbb{K}[t] \text{ with } t_V = \phi, \text{ we know from the Structure Theorem for Modules over PIDs that } V \text{ is the direct sum of its primary components,}

\[ V(p) = \{ v \in V \mid p(\phi)^r(v) = 0 \text{ for some } r \in \mathbb{Z}_+ \}, \]

where } p \in \mathbb{K}[t] \text{ runs over the monic irreducible factors of the minimal polynomial of } \phi. \text{ Each } V(p) \text{ is an } A\text{-subrepresentation of } V. \text{ Since } V \text{ is assumed indecomposable, there can only be one nonzero component. Thus, } p(\phi)^r = 0 \text{ for some monic irreducible } p \in \mathbb{K}[t] \text{ and some } r \in \mathbb{Z}_+. \text{ If } p = t \text{ then } \phi^r = 0. \text{ Otherwise, } 1 = ta + p^r b \text{ for suitable } a, b \in \mathbb{K}[t] \text{ and it follows that } a(\phi) = \phi^{-1}. \text{ This proves the first assertion.}

For the second assertion, consider } \phi, \psi \in \text{End}_A(V). \text{ If } \phi \circ \psi \text{ is bijective, then so are both } \phi \text{ and } \psi. \text{ Thus, we only need to show that if } \phi, \psi \text{ are nilpotent, then } \theta = \phi + \psi \text{ is nilpotent as well. But otherwise } \theta \text{ is an automorphism and } \text{Id}_V - \theta^{-1} \circ \phi = \theta^{-1} \circ \psi. \text{ The right hand side is nilpotent whereas the left hand side has inverse } \sum_{i \geq 0} (\theta^{-1} \circ \phi)^i, \text{ giving the desired contradiction.} \square

The ideal } N = \{ \phi \in \text{End}_A(V) \mid \phi \text{ is nilpotent} \} \text{ in Lemma 1.20 clearly contains all proper left and right ideals of } \text{End}_A(V) \text{ and } \text{End}_A(V)/N \text{ is a division algebra. Thus, the algebra } \text{End}_A(V) \text{ is local.}

Proof of the Krull-Schmidt Theorem. Only uniqueness remains to be addressed. So let } V := \bigoplus_{i=1}^{r} V_i \text{ and } W := \bigoplus_{j=1}^{s} W_j \text{ be given with indecomposable } V_i, W_j \in \text{Rep}_{\text{fin}} A \text{ and assume that } \phi: V \rightarrow W \text{ is an isomorphism. Let } \mu_i: V_i \hookrightarrow V \text{ and } \pi^*_i: V \twoheadrightarrow V_i \text{ be the standard embedding and projection maps as in §1.1.4; so } \pi^*_j \circ \mu_i = \delta_{i,j} \text{Id}_{V_i} \text{ and } \sum_i \pi^*_i \circ \mu_i = \text{Id}_V. \text{ Similarly, we also have } \mu'_j: W_j \hookrightarrow W \text{ and } \pi'_j: W \twoheadrightarrow W_j. \text{ The maps}

\[ \alpha_j := \pi'_j \circ \phi \circ \mu_1 : V_1 \rightarrow W_j \quad \text{and} \quad \beta_j := \pi_1 \circ \phi^{-1} \circ \mu'_j : W_j \rightarrow V_1 \]

satisfy } \sum_j \beta_j \circ \alpha_j = \text{Id}_{V_1}. \text{ It follows from Lemma 1.20 that some } \beta_j \circ \alpha_j \text{ must be an automorphism of } V_1; \text{ after renumbering if necessary, we may assume that } j = 1.
Since $W_1$ is indecomposable, it further follows that $\alpha_1$ and $\beta_1$ are isomorphisms (Exercise 1.1.2); so $V_i \cong W_1$. Finally, consider the map

$$\psi : V_{>1} := \bigoplus_{i > 1} V_i \xleftarrow{\mu_{>1}} V \xrightarrow{\sim} W \xrightarrow{\pi_{>1}} W_{>1} := \bigoplus_{j > 1} W_j,$$

where $\mu_{>1}$ and $\pi_{>1}$ again are the standard embedding and projection maps. It suffices to show that $\psi$ is surjective. For, then $\psi$ must be an isomorphism for dimension reasons and an induction finishes the proof. So let $v \in \text{Ker} \psi$. Then $\phi \circ \mu_{>1}(v) = \mu_1'(w)$ for some $w \in W_1$ and $\beta_1(w) = \pi_1 \circ \phi^{-1} \circ \phi \circ \mu_{>1}(v) = 0$. Since $\beta_1$ is mono, it follows that $w = 0$, and since $\phi \circ \mu_{>1}$ is mono as well, it further follows that $v = 0$ as desired.

### Exercises for Section 1.2

Unless mentioned otherwise, $A \in \text{Alg}_k$ is arbitrary in these exercises.

#### 1.2.1 (Kernels) Given a map $\phi : A \to B$ in $\text{Alg}_k$, consider the functors $\phi^* : \text{Rep} B \to \text{Rep} A$ and $\phi_* = \text{Ind}^B_A : \text{Rep} A \to \text{Rep} B$ (§1.2.2). Show:

(a) $\text{Ker}(\phi^* V) = \phi^{-1}(\text{Ker} V)$ for $V \in \text{Rep} B$.

(b) Assume that $B$ is free as right $A$-module via $\phi$. Then, for any $W \in \text{Rep} A$, $\text{Ker}(\phi, W) = \{b \in B \mid bB \subseteq B\phi(\text{Ker} W)\}$, the largest ideal of $B$ that is contained in the left ideal $B\phi(\text{Ker} W)$.

#### 1.2.2 (Faithfulness) Let $V \in \text{Rep} A$ be such that $\text{Res}^B_A V$ is finitely generated. Show that $V$ is faithful if and only if $A\text{reg}$ embeds into $V^{\otimes n}$ for some $n \in \mathbb{N}$.

#### 1.2.3 (Twisting representations) For $V \in \text{Rep} A$ and $\alpha \in \text{Aut}_{\text{Alg}_k}(A)$, consider the twisted representation $\alpha V$ as in (1.24). Show:

(a) $\alpha(\beta V) \cong \alpha \circ \beta V$ for all $\alpha, \beta \in \text{Aut}_{\text{Alg}_k}(A)$.

(b) If $\alpha$ is an inner automorphism, that is, $\alpha(a) = uau^{-1}$ for some $u \in A^X$, then $\alpha V \cong V$ in $\text{Rep} A$ via $\alpha_v \leftrightarrow u_v$.

(c) The map $\alpha : V \xrightarrow{\sim} \alpha V$ yields a bijection between the subrepresentations of $V$ and $\alpha V$. In particular, $\alpha V$ is irreducible, completely reducible, has finite length etc. if and only if this holds for $V$.

(d) $\alpha A\text{reg} \cong A\text{reg}$ via $\alpha a \leftrightarrow \alpha(a)$.

#### 1.2.4 (Extension of scalars for homomorphisms) For given representations $V, W \in \text{Rep} A$ and a given field extension $K/\mathbb{k}$, show that the $K$-linear map (B.27) restricts to a $K$-linear map $\text{K} \otimes \text{Hom}_A(V, W) \to \text{Hom}_{K \otimes A}(K \otimes V, K \otimes W)$, $\lambda \otimes f \mapsto \rho_{\text{reg}}(\lambda) \otimes f$. Use the facts stated in §B.3.4 to show that this map is always injective and that it is bijective if $V$ is finite dimensional or the field extension $K/\mathbb{k}$ is finite.

#### 1.2.5 (Noether-Deuring Theorem) Let $V, W \in \text{Rep}_{\text{fin}} A$ and let $K/\mathbb{k}$ be a field extension. The Noether-Deuring Theorem states that $K \otimes V \cong K \otimes W$ in $\text{Rep}(K \otimes A)$
if and only if $V \cong W$ in $\text{Rep} A$. To prove the nontrivial direction, assume that $K \otimes V \cong K \otimes W$ in $\text{Rep} (K \otimes A)$ and complete the following steps.

(a) Fix a $k$-basis $(\phi_i)_{i=1}^r$ of $\text{Hom}_A(V, W)$ and identify $\text{Hom}_K(K \otimes V, K \otimes W)$ with $\bigoplus_{i=1}^r K \otimes \phi_i$ (Exercise 1.2.4). Observe that $\det(\sum_i \lambda_i \otimes \phi_i) = \Lambda^n(\sum_i \lambda_i \otimes \phi_i)$ is a homogeneous polynomial $f(\lambda_1, \ldots, \lambda_r)$ of degree $n = \dim_V V = \dim_K W$ over $k$ and $f(\lambda_1, \ldots, \lambda_r) \neq 0$ for some $(\lambda_i) \in K^r$.

(b) If $|k| \geq n$, conclude that $f(\lambda_1', \ldots, \lambda_r') \neq 0$ for some $(\lambda_i') \in \mathbb{R}^r$ (Exercise C.3.2). Deduce that $\sum_i \lambda_i' \phi_i \in \text{Hom}_A(V, W)$ is an isomorphism.

(c) If $|k| < n$, then choose some finite field extension $F/k$ with $|F| > n$ and elements $\mu_i \in F$ with $f(\mu_1, \ldots, \mu_r) \neq 0$ to obtain $F \otimes V \cong F \otimes W$. Conclude that $V^\otimes d \cong W^\otimes d$ in $\text{Rep} A$, with $d = [F : k]$. Invoke the Krull-Schmidt Theorem (§1.2.6) to further conclude that $V \cong W$ in $\text{Rep} A$.

1.2.6 (Reynolds operators). Let $A \in \text{Alg}_k$ be arbitrary and let $B$ be a subalgebra of $A$. A Reynolds operator for the extension $B \subseteq A$, by definition, is a map $\pi : A \to B$ in $\text{BMod}_B$ such that $\pi|_B = \text{Id}_B$. Assuming that such a map $\pi$ exists, prove:

(a) If $A$ is left (or right) noetherian, then so is $B$; likewise for artinian.

(b) Let $W \in \text{Rep} B$. The composite of $\pi \otimes_B \text{Id}_W : \text{Ind}_B^A W \to \text{Ind}_B^A W$ with the canonical isomorphism $\text{Ind}_B^A W \cong W$ is an epimorphism $\text{Res}_B^A \text{Ind}_B^A W \to W$ in $\text{Rep} B$ that is split by the map $\sigma : W \to \text{Res}_B^A \text{Ind}_B^A W, w \mapsto 1 \otimes w$. (See Exercise 1.1.2.) Conclude that $W$ is isomorphic to a direct summand of $\text{Res}_B^A \text{Ind}_B^A W$.

(c) Similarly, the composite of the canonical isomorphism $W \cong \text{Coind}_B^A W$ with $\pi^* : \text{Coind}_B^A W \to \text{Coind}_B^A W$ is a monomorphism $\psi_\pi : W \hookrightarrow \text{Res}_B^A \text{Coind}_B^A W$ in $\text{Rep} B$ that splits the map $\tau : \text{Res}_B^A \text{Coind}_B^A W \to W, f \mapsto f(1)$. Thus, $W$ is isomorphic to a direct summand of $\text{Res}_B^A \text{Coind}_B^A W$. The unique lift of $\psi_\pi$ to a map $\Psi_\pi : \text{Ind}_B^A W \to \text{Coind}_B^A W$ in $\text{Rep} A$ (Proposition 1.9) satisfies $\tau \circ \Psi_\pi \circ \sigma = \text{Id}_W$.

1.2.7 (Cofinite subalgebras). Let $A$ be a $k$-algebra and let $B$ be a subalgebra such that $A$ is finitely generated as a left $B$-module, say $A = Ba_1 + \cdots + Ba_m$.

(a) Show that, for any $W \in \text{Rep} B$, there is a $k$-linear embedding $\text{Coind}_B^A W \hookrightarrow W^\otimes m$ given by $f \mapsto (f(a_1))$.

(b) Let $0 \neq V \in \text{Rep} A$ be finitely generated. Use Exercise 1.1.3(a) to show that, for some $W \in \text{Irr} B$, there is an epimorphism $\text{Res}_B^A V \to W$ in $\text{Rep} B$.

(c) Conclude from (a), (b) and Proposition 1.9 that, for every $V \in \text{Irr} A$, there exists some $W \in \text{Irr} B$ such that $V$ embeds into $W^\otimes m$ as a $k$-vector space.

1.2.8 (Commutative cofinite subalgebras). Let $A$ be an affine $k$-algebra having a commutative subalgebra $B \subseteq A$ such that $A$ is finitely generated as left $B$-module. Use the weak Nullstellensatz (Section C.1), the Artin-Tate Lemma (Exercise 1.1.7) and Exercise 1.2.7(c) to show that all $V \in \text{Irr} A$ are finite dimensional.

---

*Reynolds operators are also referred to as conditional expectations in the theory of operator algebras.*
1.2.9 (Representations of the Weyl algebra). Let $A = A_1(\mathbb{k})$ denote the Weyl algebra and let $V = \mathbb{k}[t]$ be the standard representation of $A$ (Examples 1.8 and 1.13).

(a) Show that $V \not\cong A_{\text{reg}}$ in $\text{Rep} \ A$. Show also that $V$ is faithful if $\text{char} \mathbb{k} = 0$ (and recall from Example 1.13 that $V$ is also irreducible in this case), but $V$ is neither irreducible nor faithful if $\text{char} \mathbb{k} = p > 0$; determine $\ker V$ in this case.

(b) Assuming $\mathbb{k}$ to be algebraically closed with $\text{char} \mathbb{k} = p > 0$, show that all $S \in \text{Irr} \ A$ have degree $p$. (Use Exercises 1.1.14(c) and 1.2.8.)

1.2.10 (Finite length and chain conditions). Before tackling this exercise, it may be useful to review Exercise 1.1.4. A representation $V \in \text{Rep} \ A$ is said to be artinian if $V$ satisfies the Descending Chain Condition (DCC): if $U_1 \supseteq U_2 \supseteq U_3 \supseteq \ldots$ are subrepresentations of $V$, then $U_n = U_{n+1} = \ldots$ for some $n$.

(a) Show that DCC is equivalent to the Minimum Condition: Every nonempty collection of subrepresentations of $V$ has at least one minimal member.

(b) Given a short exact sequence $0 \to U \to V \to W \to 0$ in $\text{Rep} \ A$, show that $V$ is artinian if and only if both $U$ and $W$ are artinian.

(c) Show that $V \in \text{Rep} \ A$ has finite length if and only if $V$ is artinian and noetherian. Give another proof of Theorem 1.18(a) using this fact in conjunction with (b) and Exercise 1.1.4(b).

1.2.11 (Finite length and filtrations). A filtration of length $l$ of $V \in \text{Rep} \ A$, by definition, is any chain of subrepresentations $\mathcal{F}: 0 = V_0 \subsetneq V_1 \subsetneq \ldots \subsetneq V_l = V$. If all $V_i$ also occur in another filtration of $V$, then the latter filtration is called a refinement of $\mathcal{F}$; the refinement is said to be proper if it has larger length than $\mathcal{F}$. Thus, a composition series of $V$ is the same as a filtration of finite length that admits no proper refinement. Prove:

(a) If $V$ has finite length, then any filtration $\mathcal{F}$ can be refined to a composition series of $V$.

(b) $V$ has finite length if and only if there is a bound on the lengths of all finite-length filtrations of $V$.

1.2.12 (Weak Nullstellensatz for large base fields). Consider the Schur division algebra $D(S)$ for $S \in \text{Irr} \ A$.

(a) Show that $\dim_\mathbb{k} D(S) \leq \dim_\mathbb{k} S \leq \dim_\mathbb{k} A$.

(b) Show that, for any division $\mathbb{k}$-algebra $D$ and any $d \in D$ that is not algebraic over $\mathbb{k}$, the set $\{ (d - \lambda)^{-1} \mid \lambda \in \mathbb{k} \}$ is linearly independent over $\mathbb{k}$.

(c) Conclude from (a) and (b) that if the cardinality $|\mathbb{k}|$ is strictly larger than $\dim_\mathbb{k} A$, then $D(S)$ is algebraic over $\mathbb{k}$.
1.3. Primitive Ideals

The investigation of the set \( \text{Irr} A \) of irreducible representations of a given algebra \( A \), in many cases of interest, benefits from an ideal theoretic perspective. The link between representations and ideals of \( A \) is provided by the notion of the kernel of a representation \( V \in \text{Rep} A \),

\[
\text{Ker} V = \{ a \in A \mid a.v = 0 \text{ for all } v \in V \}.
\]

Isomorphic representations evidently have the same kernel, but the converse is generally far from true. For example, the standard representation and the regular representation of the Weyl algebra are are not isomorphic (in any characteristic), even though they are both faithful in characteristic 0 (Exercise 1.2.9).

The kernels of irreducible representations of \( A \) are called the primitive ideals\(^5\) of \( A \). If \( S \in \text{Irr} A \) is written in the form \( S \cong A/L \) for some maximal left ideal \( L \) of \( A \) as in Lemma 1.14, then \( \text{Ker} S = \{ a \in A \mid aA \subseteq L \} \); this set can also be described as the largest ideal of \( A \) that is contained in \( L \). We shall denote the collection of all primitive ideals of \( A \) by

\[ \text{Prim} A. \]

Thus, there always is the surjection

\[
\begin{array}{ccc}
\text{Irr} A & \longrightarrow & \text{Prim} A \\
\downarrow & & \downarrow \\
S & \longrightarrow & \text{Ker} S
\end{array}
\]

\[ (1.35) \]

While this map is not bijective in general, its fibers do at least afford us a rough classification of the irreducible representations of \( A \).

1.3.1. Degree-1 Representations

Representations of degree 1 of any algebra \( A \) are clearly irreducible. They are given by homomorphisms \( \phi \in \text{Hom}_{\text{Alg}}(A, \mathbb{k}) \), since \( \text{End}_{\mathbb{k}}(V) = \mathbb{k} \) if \( \dim_{\mathbb{k}} V = 1 \). For any such \( \phi \), we will use the notation

\[ \mathbb{k}_\phi \]

to denote the field \( \mathbb{k} \) with \( A \)-action \( a.\lambda = \phi(a).\lambda \) for \( a \in A \) and \( \lambda \in \mathbb{k} \). The primitive ideal that is associated to the irreducible representation \( \mathbb{k}_\phi \) is \( \text{Ker} \mathbb{k}_\phi = \text{Ker} \phi \); this is an ideal of codimension 1 in \( A \) and all codimension-1 ideals have the form \( \text{Ker} \phi \) with \( \phi \in \text{Hom}_{\text{Alg}}(A, \mathbb{k}) \). Assuming \( A \neq 0 \) (otherwise \( \text{Irr} A = \emptyset \)) and viewing \( \mathbb{k} \subseteq A \) via the unit map, we have \( A = \mathbb{k} \oplus \text{Ker} \phi \) and \( \phi(a) \) is the projection of \( a \in A \) onto

---

\(^5\)Strictly speaking, primitive ideals should be called left primitive, since irreducible representations are irreducible left modules. Right primitive ideals, defined as the annihilators of irreducible right modules, do not always coincide with primitive ideals in the above sense [13].
the first summand. Thus, we can recover \( \phi \) from \( \text{Ker } \phi \). Consequently, restricting (1.35) to degree-1 representations, we obtain bijections of sets

\[
\text{Irr } A \quad \triangleright \quad \text{Hom}_{\text{Alg}}(A, \mathbb{k}) \quad \triangleright \quad \{ \text{codimension-1 ideals of } A \} \quad \triangleright \quad \{ \text{equivalence classes of degree-1 representations of } A \}
\]

\[
\phi \quad \triangleright \quad \text{Ker } \phi \quad \triangleright \quad \mathbb{k}_\phi
\]

1.3.2. Commutative Algebras

If \( A \) is a commutative \( \mathbb{k} \)-algebra, then maximal left ideals are the same as maximal ideals of \( A \). Thus, denoting the collection of all maximal ideals of \( A \) by

\[ \text{MaxSpec } A, \]

we know from Lemma 1.14 that each \( S \in \text{Irr } A \) has the form \( S \equiv A/P \) for some \( P \in \text{MaxSpec } A \). Since \( \text{Ker}(A/I) = I \) holds for every ideal \( I \) of \( A \), we obtain that \( P = \text{Ker } S \) and \( S \equiv A/\text{Ker } S \). This shows that the primitive ideals of \( A \) are exactly the maximal ideals and that (1.35) is a bijection for commutative \( A \):

\[
\text{Irr } A \quad \triangleright \quad \text{Prim } A = \text{MaxSpec } A
\]

Thus, for commutative \( A \), the problem of describing \( \text{Irr } A \) reduces to the description of \( \text{MaxSpec } A \).

Now assume that \( A \) is affine commutative and that the base field \( \mathbb{k} \) is algebraically closed. Then all irreducible representations \( A/P \) are 1-dimensional by Hilbert’s Nullstellensatz (Section C.1). Hence, for any ideal \( P \) of \( A \), the following are equivalent:

\[
P \text{ is primitive } \iff P \text{ is maximal } \iff A/P = \mathbb{k}.
\]

In view of (1.36) we obtain a bijection of sets

\[
\text{Irr } A \quad \triangleright \quad \text{Hom}_{\text{Alg}}(A, \mathbb{k})
\]

With this identification, \( \text{Irr } A \) can be thought of geometrically as the set of closed points of an affine algebraic variety over \( \mathbb{k} \). For example, (1.7) tells us that, for the polynomial algebra \( \mathbb{k}[x_1, x_2, \ldots, x_n] \), the variety in question is affine \( n \)-space \( \mathbb{k}^n \):

\[
\text{Irr } \mathbb{k}[x_1, x_2, \ldots, x_n] \quad \triangleright \quad \mathbb{k}^n.
\]
The pullback of an irreducible representation of $A$ along a $\mathbb{k}$-algebra map $\phi: B \to A$ (§1.2.2) is a degree-1 representation of $B$; so we obtain a map $\phi^* = \text{Res}_A^B: \text{Irr} A \to \text{Irr} B$. If $B$ is also affine commutative, then this is a morphism of affine algebraic varieties [97]. These remarks place the study of irreducible representations of affine commutative algebras over an algebraically closed base field within the realm of algebraic geometry which is outside the scope of this book. Nonetheless, the geometric context sketched above does provide the original background for some of the material on primitive ideals to be discussed later in this section. We end our excursion on commutative algebras with a simple example.

**Example 1.21.** As was mentioned, the irreducible representations of the polynomial algebra $A = \mathbb{k}[x, y]$ over an algebraically closed field $\mathbb{k}$ correspond to the points of the affine plane $\mathbb{k}^2$. Let us consider the subalgebra $B = \mathbb{k}[x^2, y^2, xy]$ and let $\phi: B \hookrightarrow A$ denote the inclusion map. It is not hard to see that $B \cong \mathbb{k}[x_1, x_2, x_3]/(x_1 x_2 - x_3)$; so the irreducible representations of $B$ correspond to the points of the cone $x_3 = x_1 x_2$ in $\mathbb{k}^3$. The following picture illustrates the restriction map $\phi^* = \text{Res}_A^B: \text{Irr} A \to \text{Irr} B$; this map is easily seen to be surjective.

![Diagram](image)

**1.3.3. Connections with Prime and Maximal Ideals**

For a general $A \in \text{Alg}_k$, primitive ideals are sandwiched between maximal and prime ideals of $A$:

\[
\text{MaxSpec } A \subseteq \text{Prim } A \subseteq \text{Spec } A
\]

(1.40)

Here, $\text{MaxSpec } A$ is the set of all maximal ideals of $A$ as in §1.3.2 and $\text{Spec } A$ denotes the set of all prime ideals of $A$. Recall that an ideal $P$ of $A$ is **prime** if $P \neq A$ and $IJ \subseteq P$ for ideals $I, J$ of $A$ implies that $I \subseteq P$ or $J \subseteq P$. To see that primitive ideals are prime, assume that $P = \text{Ker } S$ for $S \in \text{Irr } A$ and let $I, J$ be ideals of $A$ such that $I \nsubseteq P$ and $J \nsubseteq P$. Then $I.S = S = J.S$ by irreducibility, and hence $IJ.S = S$. Therefore, $IJ \nsubseteq P$ as desired. For the first inclusion in (1.40), let $P \in \text{MaxSpec } A$ and let $L$ be any maximal left ideal of $A$ containing $P$. Then $A/L$ is irreducible and $\text{Ker}(A/L) = P$. Thus, all maximal ideals of $A$ are primitive, thereby establishing the inclusions in (1.40). As we shall see, these inclusions are...
in fact equalities if the algebra $A$ is finite dimensional (Theorem 1.38). However, in general, all inclusions in (1.40) are strict; see Example 1.24 below and many others later on in this book.

We also remind the reader that an ideal $I$ of $A$ is called **semiprime** if, for any ideal $J$ of $A$ and any non-negative integer $n$, the inclusion $J^n \subseteq I$ implies that $J \subseteq I$. Prime ideals are clearly semiprime and intersections of semiprime ideals are evidently semiprime again. Thus, the intersection of any collection of primes is a semiprime ideal. It is a standard ring theoretic fact that all semiprime ideals arise in this manner: semiprime ideals are exactly the intersections of collections of primes (e.g., [125, 10.11]).

### 1.3.4. The Jacobson-Zariski Topology

The set of $\text{Spec } A$ of all prime ideals of an arbitrary algebra $A$ carries a useful topology, the **Jacobson-Zariski topology**. This topology is defined by declaring the subsets of the form

$$\mathcal{V}(I) \overset{\text{def}}{=} \{ P \in \text{Spec } A \mid P \supseteq I \}$$

to be closed, where $I$ can be any subset of $A$. Evidently, $\mathcal{V}(\emptyset) = \text{Spec } A$, $\mathcal{V}(\{1\}) = \emptyset$, and $\mathcal{V}(\bigcup_{\alpha} I_{\alpha}) = \bigcap_{\alpha} \mathcal{V}(I_{\alpha})$ for any collection $\{I_{\alpha}\}$ of subsets of $A$. Moreover, we may clearly replace a subset $I \subseteq A$ by the ideal of $A$ that is generated by $I$ without changing $\mathcal{V}(I)$. Thus, the closed subsets of Spec $A$ can also be described as the sets of the form $\mathcal{V}(I)$, where $I$ is an ideal of $A$. The defining property of prime ideals implies that $\mathcal{V}(I) \cup \mathcal{V}(J) = \mathcal{V}(IJ)$ for ideals $I$ and $J$. Thus, finite unions of closed sets are again closed, thereby verifying the topology axioms. The Jacobson-Zariski topology on Spec $A$ induces a topology on the subset $\text{Prim } A$, the closed subsets being those of the form $\mathcal{V}(I) \cap \text{Prim } A$; likewise for MaxSpec $A$.

The Jacobson-Zariski topology is related to the standard Zariski topology on a finite-dimensional $\mathbb{k}$-vector space $V$ (Section C.3). Indeed, let

$$O(V) = \text{Sym } V^*$$

denote the algebra of polynomial functions on $V$. If $\mathbb{k}$ is algebraically closed, then the weak Nullstellensatz (Section C.1) yields a bijection

$$
\begin{array}{ccc}
V & \overset{\sim}{\longrightarrow} & \text{MaxSpec } O(V) \\
\downarrow \quad \quad \downarrow & & \downarrow \\
\nu & \mapsto & \mathfrak{m}_\nu := \{ f \in O(V) \mid f(\nu) = 0 \}
\end{array}
$$

Viewing this as an identification, the Zariski topology on $V$ is readily seen to coincide with the Jacobson-Zariski topology on MaxSpec $O(V)$.

In comparison with the more familiar topological spaces from analysis, say, the topological space Spec $A$ generally has rather bad separation properties. Indeed, a “point” $P \in \text{Spec } A$ is closed exactly if the prime ideal $P$ is in fact maximal.
Exercise 1.3.1 explores the Jacobson-Zariski topology in more detail. Here, we content ourselves by illustrating it with three examples. Further examples will follow later.

**Example 1.22** (The polynomial algebra \( k[x] \)). Starting with \( k[x] \), we have

\[
\text{Spec } k[x] = \{(0)\} \cup \text{MaxSpec } k[x] = \{(0)\} \cup \{(f) \mid f \in k[x] \text{ irreducible}\}.
\]

If \( k \) is algebraically closed, then MaxSpec \( k[x] \) is in bijection with \( k \) via \((x - \lambda) \leftrightarrow \lambda \). Therefore, one often visualizes Spec \( k[x] \) as a “line,” the points on the line corresponding to the maximal ideals and the line itself corresponding to the ideal \((0)\). The latter ideal is a **generic point** for the topological space Spec \( A \): the closure of \((0)\) is all of Spec \( A \). Figure 1.1 renders Spec \( k[x] \) in three ways, with red dots representing maximal or, equivalently, primitive ideals in each case. The solid gray lines in the top picture represent inclusions. The large black area in the other two pictures represents the generic point \((0)\). The third picture also aims to convey the fact that \((0)\) is the determined by the maximal ideals, being their intersection, and that the topological space Spec \( k[x] \) is quasi-compact (Exercise 1.3.1).

![Figure 1.1. Spec k[x]](image)

**Example 1.23** (The polynomial algebra \( k[x, y] \)). The topology for \( k[x, y] \) is slightly more difficult to visualize than for \( k[x] \). As a set,

\[
\text{Spec } k[x, y] = \{(0)\} \cup \{(f) \mid f \in k[x, y] \text{ irreducible}\} \cup \text{MaxSpec } k[x, y].
\]

Assuming \( k \) to be algebraically closed, maximal ideals of \( k[x, y] \) are in bijection with points of the plane \( \mathbb{A}^2 \) via \((x - \lambda, y - \mu) \leftrightarrow (\lambda, \mu) \). Figure 1.2 depicts the topological space Spec \( k[x, y] \), the generic point \((0)\) again being represented by a large black region. The two curves in the plane are representative for the infinitely many primes that are generated by irreducible polynomials \( f \in k[x, y] \); and finally, we have sprinkled a few red points throughout the plane to represent MaxSpec \( k[x, y] \). A point lies on a curve exactly if the corresponding maximal ideal contains the principal ideal \((f)\) giving the curve.

**Example 1.24** (The quantum plane). Fix a scalar \( q \in \mathbb{K} \) that is not a root of unity. The **quantum plane** is the algebra

\[
A = O_q(\mathbb{K}^2) \overset{\text{def}}{=} k(x, y)/(xy - qyx).
\]
Our goal is to describe $\text{Spec } A$, paying attention to which primes are primitive or maximal. First note that the zero ideal of $A$ is certainly prime, because $A$ is a domain (Exercise 1.1.15). It remains to describe the nonzero primes of $A$. We refer to Exercise 1.1.15 for the fact that every nonzero ideal of $A$ contains some standard monomial $x^i y^j$. Observe that both $x$ and $y$ are normal elements of $A$ in the sense that $(x) = xA = Ax$ and likewise for $y$. Therefore, if $x^i y^j \in P$ for some $P \in \text{Spec } A$, then $x^i y^j A = (x)^i (y)^j \subseteq P$, and hence $x \in P$ or $y \in P$. In the former case, $P/(x)$ is a prime ideal of $A/(x) \equiv \mathbb{k}[y]$, and hence $P/(x)$ is either the zero ideal of $\mathbb{k}[y]$ or else $P/(x)$ is generated by some irreducible polynomial $g(y)$. Thus, if $x \in P$ either $P = (x)$ or $P = (x, g(y))$, which is maximal. Similarly, if $y \in P$ either $P = (y)$ or $P$ is the maximal ideal $(y, f(x))$ for some irreducible $f \in \mathbb{k}[x]$. Only $(x, y)$ occurs in both collections of primes, corresponding to $g(y) = y$ or $f(x) = x$. Therefore $\text{Spec } A$ can be pictured as shown in Figure 1.3. Solid gray lines represent inclusions, and primitive ideals are marked in red. The maximal ideals on top of the diagram in Figure 1.3 are all primitive by (1.40). On the other hand, neither $(x)$ nor $(y)$ are primitive by (1.37), because they correspond to non-maximal ideals of commutative (in fact, polynomial) algebras. It is less clear, why the zero ideal should be primitive. The reader is asked to verify this in Exercise 1.3.4, but we will later see (Exercise 5.6.5) that primitivity of $(0)$ also follows from the fact that the intersection of all nonzero primes is nonzero, which is clear from Figure 1.3: $(x) \cap (y) \neq (0)$. Note that, in this example, all inclusions in (1.40) are strict.
We finish our discussion of the quantum plane by offering another visualization of \( \text{Spec} \, A \), which emphasizes the fact that \( (0) \) is a generic point for the topological space \( \text{Spec} \, A \) — this point is represented by the large red area in the picture on the right. We also assume \( k \) to be algebraically closed. The maximal ideals \( (x, g(y)) = (x, y - \eta) \) with \( \eta \in k \) are represented by points on the \( y \)-axis, the axis itself being the generic point \( (x) \), and similarly for the \( x \)-axis with generic point \( (y) \).

1.3.5. The Jacobson Radical

The intersection of all primitive ideals of an arbitrary algebra \( A \) will play an important role in the following; it is called the \textbf{Jacobson radical} of \( A \):

\[
\text{rad} \, A \overset{\text{def}}{=} \bigcap_{P \in \text{Prim} \, A} P = \{ a \in A \mid a.S = 0 \text{ for all } S \in \text{Irr} \, A \}
\]

Being an intersection of primes, the Jacobson radical is a semiprime ideal of \( A \). Algebras with vanishing Jacobson radical are called \textit{semiprimitive}. We put

\[
A^{\text{s.p.}} \overset{\text{def}}{=} A/\text{rad} \, A
\]

Since \( (\text{rad} \, A).S = 0 \) holds for all \( S \in \text{Irr} \, A \), inflation along the canonical surjection \( A \to A^{\text{s.p.}} \) as in §1.2.2 yields a bijection

\[
(1.41) \quad \text{Irr} \, A^{\text{s.p.}} \overset{\sim}{\to} \text{Irr} \, A
\]
and $P \leftrightarrow P / \text{rad } A$ gives a bijection $\text{Prim } A \cong \text{Prim } A^{s.p.}$. Thus, $A^{s.p.}$ is semiprimitive:

(1.42) $\text{rad } A^{s.p.} = 0$.

In describing $\text{Irr } A$, we may therefore assume that $A$ is semiprimitive.

We finish this section by giving, for a finite-dimensional algebra $A$, a purely ring theoretic description of the Jacobson radical that makes no mention of representations: $\text{rad } A$ is the largest nilpotent ideal of $A$. Recall that an ideal $I$ of an algebra $A$ is called nilpotent if $I^n = 0$ for some $n$; likewise for left or right ideals.

**Proposition 1.25.** The Jacobson radical $\text{rad } A$ of any algebra $A$ contains all nilpotent left and right ideals of $A$. Moreover, for each finite-length $V \in \text{Rep } A$,

$$(\text{rad } A)^{\text{length } V} V = 0.$$

If $A$ is finite dimensional, then $\text{rad } A$ is itself nilpotent.

**Proof.** We have already pointed out that $\text{rad } A$ is a semiprime. Now, any semiprime ideal $I$ of $A$ contains all left ideals $L$ of $A$ such that $L^n \subseteq I$. To see this, note that $LA$ is an ideal of $A$ that satisfies $(LA)^n = L^n A \subseteq I$. By the defining property of semiprime ideals, it follows that $LA \subseteq I$ and hence $L \subseteq I$. A similar argument applies to right ideals. In particular, every semiprime ideal contains all nilpotent left and right ideals of $A$. This proves the first statement.

Now assume that $0 = V_0 \subset V_1 \subset \cdots \subset V_l = V$ is a composition series of $V$. Since $\text{rad } A$ annihilates all irreducible factors $V_i / V_{i-1}$, it follows that $(\text{rad } A)^i V_i \subseteq V_{i-1}$ and so $(\text{rad } A)^i V = 0$. For finite-dimensional $A$, this applies to the regular representation $V = A_{\text{reg}}$, giving $(\text{rad } A)^i A_{\text{reg}} = 0$ and hence $(\text{rad } A)^i = 0$. □

**Exercises for Section 1.3**

In these exercises, $A$ denotes an arbitrary $k$-algebra unless otherwise specified.

1.3.1 (Jacobson-Zariski topology). For any subset $X \subseteq \text{Spec } A$ and any ideal $I$ of $A$, put

$$\mathcal{I}(X) \overset{\text{def}}{=} \bigcap_{P \in X} P \quad \text{and} \quad \sqrt{I} \overset{\text{def}}{=} \mathcal{I}(\mathcal{V}(I)) = \bigcap_{P \in \text{Spec } A} P.$$

These are semiprime ideals of $A$ and all semiprime ideals are of this form. The ideal $\sqrt{I}$ is called the semiprime radical of $I$; it is clearly the smallest semiprime ideal of $A$ containing $I$. Consider the Jacobson-Zariski topology of Spec $A$.

(a) Show that the closure of $X$ is given by $\overline{X} = \mathcal{V}(\mathcal{I}(X))$.

(b) Conclude that the following are inclusion reversing bijections that are inverse to each other:
\[ \{ \text{closed subsets of Spec } A \} \xleftrightarrow[\mathcal{J}(\cdot)]{\mathcal{V}(\cdot)} \{ \text{semiprime ideals of } A \} \ . \]

Thus, the Jacobson-Zariski topology on Spec \( A \) determines all semiprime ideals of \( A \) and their inclusion relations among each other.

(c) Show that \( \mathcal{J}(\bar{X}) = \mathcal{J}(X) \) and \( \mathcal{V}(\sqrt{I}) = \mathcal{V}(I) \).

(d) A topological space is said to be \textit{irreducible} if it cannot be written as the union of two proper closed subsets. Show that, under the bijection in (b), the irreducible closed subsets of Spec \( A \) correspond to the prime ideals of \( A \).

(e) Show that Spec \( A \) is quasi-compact: if Spec \( A = \bigcup_{l \in L} U_l \) with open subsets of \( U_l \), then Spec \( A = \bigcup_{l \in L'} U_l \) for some finite \( L' \subseteq L \).

\subsection*{1.3.2 (Maximum condition on semiprime ideals)} Assume that \( A \) satisfies the maximum condition on semiprime ideals:

\[ \text{MAX}_{\text{semi-prime}} : \text{ Every nonempty collection of semiprime ideals of } A \text{ has at least one maximal member.} \]

Clearly, every right or left noetherian algebra satisfies this condition. Furthermore, affine PI-algebras are also known to satisfy \( \text{MAX}_{\text{semi-prime}} \) (e.g., Rowen [178, 6.3.36']).

(a) Show that every semiprime ideal of \( A \) is an intersection of \text{finitely many} primes of \( A \).

(b) Conclude that every closed subset of Spec \( A \), for the Jacobson-Zariski topology, is a finite union of irreducible closed sets (Exercise 1.3.1). Moreover, the topology of Spec \( A \) is determined by the inclusion relations among the primes of \( A \).

\subsection*{1.3.3 (Characterization of the Jacobson radical)} Show that the following subsets of \( A \) are all equal to \( \text{rad } A \): (i) the intersection of all maximal left ideals of \( A \), (ii) the intersection of all maximal right ideals of \( A \), (iii) the set \( \{ a \in A \mid 1 + xay \in A^\times \text{ for all } x, y \in A \} \).

\subsection*{1.3.4 (Quantum plane)} Let \( A = O_q(\mathbb{R}^2) \) be the quantum plane, with \( q \in \mathbb{k}^\times \) not a root of unity.

(a) Show that \( V = A/A(xy - 1) \) is a faithful irreducible representation of \( A \). Thus, \( (0) \) is a primitive ideal of \( A \).

(b) Assuming \( \mathbb{k} \) to be algebraically closed, show that the following account for all closed subsets of Spec \( A \): all finite subsets of MaxSpec \( A \) (including \( \emptyset \)), \( \mathcal{V}(x) \cup X \) for any finite subset \( X \subseteq \{(x - \xi, y) \mid \xi \in \mathbb{k}^\times \} \), \( \mathcal{V}(y) \cup Y \) for any finite subset \( Y \subseteq \{(x, y - \eta) \mid \eta \in \mathbb{k}^\times \} \), \( \mathcal{V}(x) \cup \mathcal{V}(y) \), and Spec \( A \). Here, we have written \( \mathcal{V}(f) = \mathcal{V}(\{f\}) \) for \( f \in A \).

\subsection*{1.3.5 (Centralizing homomorphisms)} An algebra map \( \phi : A \to B \) is said to be \textit{centralizing} if \( \phi(A) \) and the centralizer \( C_B(\phi(A)) = \{ b \in B \mid b\phi(a) = \phi(b)a b \forall a \in A \} \) together generate the algebra \( B \). Surjective algebra maps are clearly centralizing, but there are many others, e.g., the standard embedding \( A \hookrightarrow A[x] \).
(a) Show that composites of centralizing homomorphisms are centralizing.

(b) Let $\phi: A \to B$ be centralizing. Show that $\phi(\mathcal{Z}A) \subseteq \mathcal{Z}B$. For every ideal $I$ of $A$, show that $B\phi(I) = \phi(I)B$. Deduce the existence of a map $\text{Spec } B \to \text{Spec } A$, $P \mapsto \phi^{-1}(P)$.

1.4. Semisimplicity

In some circumstances, a given representation of an algebra can broken down into irreducible building blocks in a better way than choosing a composition series, namely as a direct sum of irreducible subrepresentations. Representations allowing such a decomposition are called completely reducible. It turns out that completely reducible representations share some useful features with vector spaces, notably the existence of complements for subrepresentations. In this section, we give several equivalent characterizations of complete reducibility (Theorem 1.28); we describe a standard decomposition of completely reducible representations, the decomposition into homogeneous components (§1.4.2); and we determine the structure of the algebras $A$ having the property that all $V \in \text{Rep } A$ are completely reducible (Wedderburn’s Structure Theorem). Algebras with this property are called semisimple.

Unless explicitly stipulated otherwise, $A$ will continue to denote an arbitrary $\mathbb{k}$-algebra in this section.

1.4.1. Completely Reducible Representations

Recall that $V \in \text{Rep } A$ is said to be completely reducible if

$$ V = \bigoplus_{i \in I} S_i $$

with irreducible subrepresentations $S_i \subseteq V$. Thus, each $v \in V$ can be uniquely written as a sum $v = \sum_{i \in I} v_i$ with $v_i \in S_i$ and $v_i = 0$ for all but finitely many $i \in I$. The case $V = 0$ is included here, corresponding to the empty sum.

**Example 1.26 (Division algebras).** Every representation $V$ of a division algebra is completely reducible. Indeed, any choice of basis for $V$ yields a decomposition of $V$ as a direct sum of irreducible subrepresentations.

**Example 1.27 (Polynomial algebras).** By (1.6) representations of the polynomial algebra $A = \mathbb{k}[x_1, x_2, \ldots, x_n]$ are given by a $\mathbb{k}$-vector space $V$ and a collection of $n$ pairwise commuting operators $\xi_i = (x_i)_V \in \text{End}_\mathbb{k}(V)$. Assuming $\mathbb{k}$ to be algebraically closed, $V$ is irreducible if and only if $\dim_{\mathbb{k}} V = 1$ by (1.39). A completely reducible representation $V$ of $A$ is thus given by $n$ simultaneously diagonalizable operators $\xi_i \in \text{End}_\mathbb{k}(V)$; in other words, $V$ has a $\mathbb{k}$-basis consisting

---

*This fails for the standard embedding of $A$ into the power series algebra $A[[x]]$: there are examples, due to G. Bergman, of primes $P \in \text{Spec } A[[x]]$ such that $P \cap A$ is not even semiprime [164, Example 4.2].

Completely reducible representations are also referred to as semisimple.
of eigenvectors for all $\xi_i$. If $V$ is finite dimensional, then we know from linear algebra that such a basis exists if and only if the operators $\xi_i$ commute pairwise and the minimal polynomial of each $\xi_i$ is separable, that is, it has no multiple roots.

**Characterizations of Complete Reducibility**

Recall the following familiar facts from linear algebra: all bases of a vector space have the same cardinality; every generating set of a vector space contains a basis; and every subspace of a vector space has a complement. The theorem below extends these facts to completely reducible representations of arbitrary algebras. Given a representation $V$ and a subrepresentation $U \subseteq V$, a *complement* for $U$ in $V$ is a subrepresentation $C \subseteq V$ such that $V = U \oplus C$.

**Theorem 1.28.** (a) Let $V \in \text{Rep} A$ be completely reducible, say $V = \bigoplus_{i \in I} S_i$ for irreducible subrepresentations $S_i$. Then the following are equivalent:

(i) $I$ is finite;  
(ii) $V$ has finite length;  
(iii) $V$ is finitely generated.

In this case, $|I| = \text{length } V$ (as in §1.2.4). In general, if $\bigoplus_{i \in I} S_i = \bigoplus_{j \in J} T_j$ with irreducible $S_i, T_j \in \text{Rep } A$, then $|I| = |J|$.

(b) The following are equivalent for any $V \in \text{Rep } A$:

(i) $V$ is completely reducible;
(ii) $V$ is a sum (not necessarily direct) of irreducible subrepresentations;
(iii) Every subrepresentation $U \subseteq V$ has a complement.

**Proof.** (a) First assume that $I$ is finite, say $I = \{1, 2, \ldots, l\}$, and put $V_i = \bigoplus_{j \leq i} S_j$. Then $0 = V_0 \subset V_1 \subset \cdots \subset V_l = V$ is a composition series of $V$, with factors $V_i/V_{i-1} \cong S_i$. Thus length $V = l$, proving the implication (i) $\Rightarrow$ (ii). Furthermore, since reducible modules are finitely generated (in fact, cyclic), (ii) always implies (iii), even when $V$ is not necessarily completely reducible (Exercise 1.1.3). Now assume that $V$ is finitely generated, say $V = Av_1 + Av_2 + \cdots + Av_r$. For each $j$, we have $v_j \in \bigoplus_{i \in I_j} S_i$ for some finite subset $I_j \subseteq I$. It follows that $I = \bigcup_{j=1}^r I_j$ is finite. This proves the equivalence of (i) – (iii) as well as the equality $|I| = \text{length } V$ for finite $I$. Since property (ii) and the value of length $V$ only depend on the isomorphism type of $V$ and not on the given decomposition $V = \bigoplus_{i \in I} S_i$, we also obtain that $|I| = |J|$ if $I$ is finite and $V \cong \bigoplus_{j \in J} T_j$ with irreducible $T_j \in \text{Rep } A$.

It remains to show that $|I| = |J|$ also holds if $I$ and $J$ are infinite. Replacing each $T_j$ by its image in $V$ under the isomorphism $V \cong \bigoplus_{j \in J} T_j$, we may assume that $T_j \subseteq V$ and $V = \bigoplus_{j \in J} T_j$. Select elements $0 \neq v_j \in T_j$. Then $T_j = Av_j$ and so $\{v_j\}_{j \in J}$ is a generating set for $V$. Exactly as above, we obtain that $I = \bigcup_{j \in J} I_j$ for suitable finite subsets $I_j \subseteq I$. Since $J$ is infinite, the union $\bigcup_{j \in J} I_j$ has cardinality at most $|J|$; see [25, Cor. 3 on p. E III.49]. Therefore, $|I| \leq |J|$. By symmetry, equality must hold.

(b) The implication (i) $\Rightarrow$ (ii) being trivial, let us assume (ii) and write $V = \bigoplus_{i \in I} S_i$ with irreducible subrepresentations $S_i \subseteq V$. 

Claim. Given a subrepresentation $U \subseteq V$, there exists a subset $J \subseteq I$ such that $V = U \oplus \bigoplus_{i \in J} S_i$.

This will prove (iii), with $C = \bigoplus_{i \in J} S_i$, and the case $U = 0$ also gives (i). To prove the claim, choose a subset $J \subseteq I$ that is maximal with respect to the property that the sum $V' := U + \sum_{i \in J} S_i$ is direct. The existence of $J$ is clear if $I$ is finite; in general, it follows by a straightforward Zorn’s Lemma argument. We aim to show that $V' = V$. If not, then $S_k \not\subseteq V'$ for some $k \in I$. Since $S_k$ is irreducible, this forces $S_k \cap V' = 0$, which in turn implies that the sum $V' + S_k = U + \sum_{i \in J \cup \{k\}} S_i$ is direct, contradicting maximality of $J$. Therefore, $V' = V$, proving the claim.

Finally, let us derive (ii) from (iii). To this end, let $S$ denote the sum of all irreducible subrepresentations of $V$. Our goal is to show that $S = V$. If not, then $V = S \oplus C$ for some nonzero subrepresentation $C \subseteq V$ by virtue of (iii). To reach a contradiction, it suffices to show that every nonzero subrepresentation $C \subseteq V$ contains an irreducible subrepresentation. For this, we may clearly replace $C$ by $Av$ for any $0 \neq v \in C$, and hence we may assume that $C$ is cyclic. Thus, there is some subrepresentation $D \subseteq C$ such that $C/D$ is irreducible (Exercise 1.1.3). Using (iii) to write $V = D \oplus E$, we obtain $C = D \oplus (E \cap C)$. Hence, $E \cap C \cong C/D$ is the desired irreducible subrepresentation of $C$. This proves (ii), finishing the proof of the theorem.

Corollary 1.29. Subrepresentations and homomorphic images of completely reducible representations are completely reducible. More precisely, if $V = \bigoplus_{i \in I} S_i$ for irreducible subrepresentations $S_i \subseteq V$, then all subrepresentations and all homomorphic images of $V$ are equivalent to direct sums $\bigoplus_{i \in J} S_i$ for suitable $J \subseteq I$.

Proof. Consider an epimorphism $f : V = \bigoplus_{i \in I} S_i \twoheadrightarrow W$ in $\text{Rep } A$. The claim in the proof of Theorem 1.28 gives $V = \text{Ker } f \oplus \bigoplus_{i \in J} S_i$ for some $J \subseteq I$. Hence, $W \cong V/\text{Ker } f \cong \bigoplus_{i \in J} S_i$, proving the statement about homomorphic images. Every subrepresentation $U \subseteq V$ is in fact also a homomorphic image of $V$: choosing a complement $C$ for $U$ in $V$, we obtain a projection map $V = U \oplus C \twoheadrightarrow U$. 

Theorem 1.28(a) allows us to define the length of any completely reducible representation $V = \bigoplus_{i \in I} S_i$ by

$$\text{length } V \overset{\text{def}}{=} |I|$$

In the finite case, this agrees with our definition of length in §1.2.4, and it refines the earlier definition for completely reducible representations $V$ of infinite length. We shall mostly be interested in completely reducible representations $V$ having finite length. The set theoretic calisthenics in the proof of Theorem 1.28 are then unnecessary.
1.4.2. Socle and Homogeneous Components

The sum of all irreducible subrepresentations of an arbitrary \( V \in \text{Rep} A \), which already featured in the proof of Theorem 1.28, is called the socle of \( V \). For a fixed \( S \in \text{Irr} A \), we will also consider the sum of all subrepresentations of \( V \) that are equivalent to \( S \); this sum is called the \( S \)-homogeneous component of \( V \):

\[
\text{soc } V \overset{\text{def}}{=} \text{the sum of all irreducible subrepresentations of } V
\]

\[
V(S) \overset{\text{def}}{=} \text{the sum of all subrepresentations } U \subseteq V \text{ such that } U \cong S
\]

Thus, \( V \) is completely reducible if and only if \( V = \text{soc } V \). In general, \( \text{soc } V \) is the unique largest completely reducible subrepresentation of \( V \), and it is the sum of the various homogeneous components \( V(S) \). We will see below that this sum is in fact direct. Of course, it may happen that \( \text{soc } V = 0 \); for example, this holds for the regular representation of any domain that is not a division algebra (Example 1.17).

Example 1.30 (Weight spaces and eigenspaces). Let \( S = \mathbb{k}_\phi \) be a degree-1 representation of an algebra \( A \), with \( \phi \in \text{Hom}_{\text{Alg}}(A, \mathbb{k}) \) as in §1.3.1. Then the \( S \)-homogeneous component \( V(\mathbb{k}_\phi) \) will be written as \( V_\phi \). Explicitly,

\[
V_\phi \overset{\text{def}}{=} \{ v \in V \mid a.v = \phi(a)v \text{ for all } a \in A \}
\]

If \( V_\phi \neq 0 \), then \( \phi \) is called a weight of the representation \( V \) and \( V_\phi \) is called the corresponding weight space. In the special case where \( A = \mathbb{k}[t] \) is the polynomial algebra, the map \( \phi \) is determined by the scalar \( \lambda = \phi(t) \in \mathbb{k} \) and \( V_\phi \) is the usual eigenspace for the eigenvalue \( \lambda \) of the endomorphism \( t_V \in \text{End}_\mathbb{k}(V) \).

The following proposition generalizes some familiar facts about eigenspaces.

Proposition 1.31. Let \( V \in \text{Rep} A \). Then:

(a) \( \text{soc } V = \bigoplus_{S \in \text{Irr } A} V(S) \).

(b) If \( f : V \to W \) is a map in \( \text{Rep} A \), then \( f(V(S)) \subseteq W(S) \) for all \( S \in \text{Irr } A \).

(c) For any subrepresentation \( U \subseteq V \), we have \( \text{soc } U = U \cap \text{soc } V \) and \( U(S) = U \cap V(S) \) for all \( S \in \text{Irr } A \).

Proof. (a) We only need to show that the sum of all homogeneous components is direct, that is,

\[
V(S) \cap \bigoplus_{T \neq S} V(T) = 0
\]

for all \( S \in \text{Irr } A \). Denoting the intersection on the left by \( X \), we know by Corollary 1.29 that \( X \) is completely reducible and that each irreducible subrepresentation
of \(X\) is equivalent to \(S\) and also to one of the representations \(T \in \text{Irr} \ A\) with \(T \neq S\). Since there are no such irreducible representations, we must have \(X = 0\). This proves (a).

For (b) and (c), note that Corollary 1.29 also tells us that \(f(V(S))\) and \(U \cap V(S)\) are both equivalent to direct sums of copies of the representation \(S\), which implies the inclusions \(f(V(S)) \subseteq W(S)\) and \(U \cap V(S) \subseteq U(S)\). Since the inclusion \(U(S) \subseteq U \cap V(S)\) is obvious, the proposition is proved. \(\square\)

**Multiplicities.** For any \(V \in \text{Rep} \ A\) and any \(S \in \text{Irr} \ A\), we put

\[
m(S, V) \overset{\text{def}}{=} \text{length } V(S)
\]

Thus,

\[
(1.43) \quad V(S) \cong S^{\oplus m(S, V)},
\]

where the right hand side denotes the direct sum of \(m(S, V)\) many copies of \(S\), and Proposition 1.31(a) implies that

\[
(1.44) \quad \text{soc } V \cong \bigoplus_{S \in \text{Irr} \ A} S^{\oplus m(S, V)}.
\]

The foregoing will be most important in the case of a completely reducible representation \(V\). In this case, (1.44) shows that \(V\) is determined, up to equivalence, by the cardinalities \(m(S, V)\). Any \(S \in \text{Irr} \ A\) such that \(m(S, V) \neq 0\) is called an irredicible constituent of \(V\). If \(V\) is completely reducible of finite length, then each \(m(S, V)\) is identical to the multiplicity \(\mu(S, V)\) of \(S\) in \(V\) as defined in §1.2.4. Therefore, \(m(S, V)\) is also referred to as the multiplicity of \(S\) in \(V\), even when \(V\) has infinite length. If \(V\) is a finite-length representation, not necessarily completely reducible, then \(m(S, V) \leq \mu(S, V)\) for all \(S \in \text{Irr} \ A\).

The following proposition expresses multiplicities as dimensions. Recall that, for any \(V, W \in \text{Rep} \ A\), the vector space \(\text{Hom}_A(V, W)\) is a \((\text{End}_A(W), \text{End}_A(V))\)-bimodule via composition (Example 1.3 or §B.2.1). In particular, for any \(S \in \text{Irr} \ A\), we may regard \(\text{Hom}_A(V, S)\) as a left vector space over the Schur division algebra \(D(S)\) and \(\text{Hom}_A(S, V)\) is a right vector space over \(D(S)\).

**Proposition 1.32.** Let \(V \in \text{Rep} \ A\) be completely reducible of finite length. Then, for any \(S \in \text{Irr} \ A\),

\[
m(S, V) = \dim_{D(S)} \text{Hom}_A(V, S) = \dim_{D(S)} \text{Hom}_A(S, V).
\]

**Proof.** The functors \(\text{Hom}_A(\cdot, S)\) and \(\text{Hom}_A(S, \cdot)\) commute with finite direct sums; see (B.14). Moreover, by Schur’s Lemma, \(\text{Hom}_A(V(T), S) = 0\) for distinct \(T, S \in \text{Irr} \ A\). Therefore,

\[
\text{Hom}_A(V, S) \cong \text{Hom}_A(V(S), S) \overset{(1.43)}{=} \text{Hom}_A(S^{\oplus m(S, V)}, S) \cong D(S)^{\oplus m(S, V)}.
\]
Consequently, \( m(S, V) = \dim_{D(S)} \text{Hom}_A(V, S) \). The verification of the second equality is analogous; an explicit isomorphism is given by

\[
\Hom_A(S, V) \otimes_{D(S)} S \xrightarrow{\sim} V(S)
\]

\[
f \otimes s \xrightarrow{\sim} f(s)
\]

**1.4.3. Endomorphism Algebras**

In this subsection, we give a description of the endomorphism algebra \( \text{End}_A(V) \) of any completely reducible \( V \in \text{Rep} A \) having finite length. Recall from (1.44) that \( V \) can be uniquely written as

\[
V \cong S_1^{\oplus m_1} \oplus S_2^{\oplus m_2} \oplus \cdots \oplus S_t^{\oplus m_t}
\]

for pairwise distinct \( S_i \in \text{Irr} A \) and positive integers \( m_i \). The description of \( \text{End}_A(V) \) will be in terms of the direct product of matrix algebras over the Schur division algebras \( D(S_i) \). Here, the **direct product** of algebras \( A_1, \ldots, A_t \) is the cartesian product,

\[
\prod_{i=1}^t A_i = A_1 \times \cdots \times A_t,
\]

with the componentwise algebra structure; e.g., multiplication is given by

\[
(x_1, x_2, \ldots, x_t)(y_1, y_2, \ldots, y_t) = (x_1 y_1, x_2 y_2, \ldots, x_t y_t).
\]

**Proposition 1.33.** Let \( V \) be as in (1.45) and let \( D(S_i) = \text{End}_A(S_i) \) denote the Schur division algebra of \( S_i \). Then

\[
\text{End}_A(V) \cong \prod_{i=1}^t \text{Mat}_{m_i}(D(S_i)).
\]

**Proof.** By Schur’s Lemma, \( \text{Hom}_A(S_i, S_j) = 0 \) for \( i \neq j \). Therefore, putting \( D_i = D(S_i) \), Lemma 1.4(a) gives an algebra isomorphism

\[
\text{End}_A(V) \cong \begin{pmatrix} \text{Mat}_{m_1}(D_1) & & \\
 & \text{Mat}_{m_2}(D_2) & \\
 & & \ddots \\
 & & & \text{Mat}_{m_t}(D_t) \end{pmatrix}.
\]

This is exactly what the proposition asserts. \( \square \)
1.4.4. Semisimple Algebras

The algebra $A$ is called **semisimple** if the following equivalent conditions are satisfied:

(i) The regular representation $A_{\text{reg}}$ is completely reducible;

(ii) All $V \in \text{Rep} A$ are completely reducible.

Condition (ii) certainly implies (i). For the converse, note that every $V \in \text{Rep} A$ is a homomorphic image of a suitable direct sum of copies of the regular representation $A_{\text{reg}}$: any family $(v_i)_{i \in I}$ of generators of $V$ gives rise to an epimorphism $A_{\text{reg}}^{\oplus I} \twoheadrightarrow V$, $(a_i) \mapsto \sum_i a_i v_i$. Now $A_{\text{reg}}^{\oplus I}$ is completely reducible by (i), being a direct sum of completely reducible representations, and it follows from Corollary 1.29 that $V$ is completely reducible as well; in fact, $V$ is isomorphic to a direct sum of certain irreducible constituents of $A_{\text{reg}}$, possibly with multiplicities greater than 1. Thus, (i) and (ii) are indeed equivalent. Since property (ii) evidently passes to homomorphic images of $A$, we obtain in particular that all homomorphic images of semisimple algebras are again semisimple.

As we have seen, division algebras are semisimple (Example 1.26). The main result of this section, Wedderburn’s Structure Theorem, gives a complete description of all semisimple algebras: they are exactly the finite direct products of matrix algebras over various division algebras. In detail:

**Wedderburn’s Structure Theorem.** The $k$-algebra $A$ is semisimple if and only if

$$A \cong \prod_{i=1}^t \text{Mat}_{m_i}(D_i)$$

for division $k$-algebras $D_i$. The data on the right are determined by $A$ as follows:

- $t = \# \text{Irr } A$, say $\text{Irr } A = \{S_1, S_2, \ldots, S_t\}$;
- $D_i \cong D(S_i)^{\text{op}}$;
- $m_i = m(S_i, A_{\text{reg}}) = \dim_{D(S_i)} S_i$;
- $\text{Mat}_{m_i}(D_i) \cong \text{BiEnd}_A(S_i)$.

**Proof.** First assume that $A$ is semisimple. Since the regular representation $A_{\text{reg}}$ has finite length, being generated by the identity element of $A$, it follows that

$$A_{\text{reg}} \cong S_1^{\oplus m_1} \oplus S_2^{\oplus m_2} \oplus \cdots \oplus S_t^{\oplus m_t}.$$
with pairwise distinct \( S_i \in \text{Irr} A \) and positive integers \( m_i \) as in (1.45). We obtain algebra isomorphisms,

\[
A \cong \text{End}_A(A_{\text{reg}})^{\text{op}} \\
\cong \prod_{i=1}^t \text{Mat}_{m_i}(D(S_i))^{\text{op}} \\
\cong \text{Lemma 1.5(a)} \prod_{i=1}^t \text{Mat}_{m_i}(D(S_i))
\]

with \( D_i = D(S_i)^{\text{op}} \). Here, we have tacitly used the obvious isomorphism \((A^{\text{op}})^{\text{op}} \cong A\) and that \( ^{\text{op}} \) commutes with direct products. Since opposite algebras of division algebras are clearly division algebras as well, it follows that all \( D_i \) are division algebras (Schur’s Lemma), and so we have proved the asserted structure of \( A \).

Conversely, assume that \( A \cong \prod_{i=1}^t A_i \) with \( A_i = \text{Mat}_{m_i}(D_i) \) for division \( k \)-algebras \( D_i \) and positive integers \( m_i \). The direct product structure of \( A \) implies that

\[
A_{\text{reg}} \cong \bigoplus_{i=1}^t \pi_i^*(A_i)_{\text{reg}},
\]

where \( \pi_i^* : \text{Rep} A_i \to \text{Rep} A \) denotes the inflation functor along the standard projection \( \pi_i : A \to A_i \) (§1.2.2). Inflation injects each \( \text{Irr} A_i \) into \( \text{Irr} A \) and yields a bijection \( \text{Irr} A \cong \bigsqcup_{i=1}^t \text{Irr} A_i \). Viewing any \( V \in \text{Rep} A_i \) as a representation of \( A \) by inflation, the subalgebras \( A_V \) and \( (A_i)_V \) of \( \text{End}_A(V) \) coincide. In particular, \( V \) is completely reducible in \( \text{Rep} A \) if and only if this holds in \( \text{Rep} A_i \). Moreover, \( \text{End}_A(V) = \text{End}_A(V) \) and \( \text{BiEnd}_A(V) = \text{BiEnd}_A(V) \).

In light of these remarks, we may assume that \( A = \text{Mat}_m(D) \) for some division \( k \)-algebra \( D \) and we must show: \( A_{\text{reg}} \) is completely reducible; \# \( \text{Irr} A = 1 \), say \( \text{Irr} A = \{ S \} \); \( m = m(S, A_{\text{reg}}) = \dim_{D(S)} S ; D \cong D(S)^{\text{op}} ; \) and \( A \cong \text{BiEnd}_A(S) \). Let \( L_j \subseteq A \) denote the collection of all matrices such that nonzero matrix entries can only occur in the \( j \)th column. Then each \( L_j \) is a left ideal of \( A \) and \( A_{\text{reg}} = \bigoplus_j L_j \). Moreover, \( L_j \equiv S := D^{\otimes m}_{\text{reg}} \) as left module over \( A \), with \( A \) acting on \( D^{\otimes m}_{\text{reg}} \) by matrix multiplication as in Lemma 1.4(b). Therefore, \( A_{\text{reg}} \cong S^{\otimes m} \). Moreover, since the regular representation of \( D \) is irreducible, it is easy to see that \( S \in \text{Irr} A \) (Exercise 1.4.2). This shows that \( A_{\text{reg}} \) is completely reducible, with \( m = m(S, A_{\text{reg}}) \), and so \( A \) is semisimple. Since every representation of a semisimple algebra is isomorphic to a direct sum of irreducible constituents of the regular representation, as we have remarked at the beginning of this subsection, we also obtain that \( \text{Irr} A = \{ S \} \). As for the Schur division algebra of \( S \), we have

\[
D(S) = \text{End}_A(D^{\otimes m}_{\text{reg}}) \cong \text{End}_D(D_{\text{reg}}) \cong D^{\text{op}}.
\]
Consequently, \( D \cong D(S)^{\text{op}} \) and \( \dim_{D(S)} S \) is equal to the dimension of \( D^{\otimes m} \) as a right vector space over \( D \), which is \( m \). Finally, Lemma 1.5 implies that \( \text{End}_{D^\text{op}}(D) \cong D^m \) and so we obtain
\[
\text{BiEnd}_A(S) = \text{End}_{D(S)}(S) \cong \text{End}_{D^\text{op}}(D^{\otimes m}) \cong \text{Mat}_m(\text{End}_{D^\text{op}}(D)) \cong \text{Mat}_m(D) = A.
\]
This finishes the proof of Wedderburn’s Structure Theorem. \( \square \)

1.4.5. Consequences of Wedderburn’s Structure Theorem

First, for future reference, let us restate the isomorphism in Wedderburn’s Structure Theorem.

Corollary 1.34. If \( A \) is semisimple, then there is an isomorphism of algebras
\[
A \cong \prod_{S \in \text{Irr} A} \text{BiEnd}_A(S)
\]
\[
\begin{array}{c}
a \\
\mapsto \\
(\{a_S\})
\end{array}
\]

Split Semisimple Algebras. A semisimple algebra \( A \) is called split semisimple if \( A \) is finite dimensional and \( k \) is a splitting field for \( A \). Then \( \text{BiEnd}_A(S) = \text{End}_k(S) \) for all \( S \in \text{Irr} A \) by (1.34), and so the isomorphism in Corollary 1.34 takes the form
\[
A \cong \prod_{S \in \text{Irr} A} \text{End}_k(S) \cong \prod_{S \in \text{Irr} A} \text{Mat}_{\dim_k S}(k).
\]

Our next corollary records some important numerology resulting from this isomorphism. For any algebra \( A \) and any \( a, b \in A \), the expression
\[
[a, b] := ab - ba
\]
will be called a Lie commutator and the \( k \)-subspace of \( A \) that is generated by the Lie commutators will be denoted by \([A, A]\).

Corollary 1.35. Let \( A \) be a split semisimple \( k \)-algebra. Then:

(a) \# \text{Irr} A = \dim_k A/[A, A];

(b) \( \dim_k A = \sum_{S \in \text{Irr} A} (\dim_k S)^2 \);

(c) \( m(S, A_{\text{reg}}) = \dim_k S \text{ for all } S \in \text{Irr} A \).

Proof. Under the isomorphism (1.46), the subspace \([A, A] \subseteq A \) corresponds to \( \prod_{S \in \text{Irr} A} [\text{Mat}_{d_S}(k), \text{Mat}_{d_S}(k)] \), where we have put \( d_S = \dim_k S \). Each of the subspaces \([\text{Mat}_{d_S}(k), \text{Mat}_{d_S}(k)]\) coincides with the kernel of the matrix trace, \( \text{Mat}_{d_S}(k) \to k \). Indeed, any Lie commutator of matrices has trace zero and, on the other hand, using the matrices \( e_{i,j} \) having a 1 in position \((i, j)\) and 0s elsewhere, we can form the Lie commutators \([e_{i,i}, e_{i,j}] = e_{i,j} (i \neq j)\) and \([e_{i,i+1}, e_{i+1,i}] = e_{i,i} - e_{i+1,i+1}\), which together span a subspace of codimension 1 in \( \text{Mat}_{d_S}(k) \). Thus,
each \([\text{Mat}_d(k), \text{Mat}_d(k)]\) has codimension 1, and hence \(\dim_k A/[A, A]\) is equal to the number of matrix components in (1.46), which in turn equals \(#\text{Irr} A\) by Wedderburn’s Structure Theorem. This proves (a).

Part (b) is clear from (1.46). Finally, (c) follows from (1.46) and the statement about multiplicities in Wedderburn’s Structure Theorem. □

**Primitive Central Idempotents.** For any semisimple algebra \(A\), we let \(e(S) \in A\) denote the element corresponding to \((0, \ldots, 0, \text{Id}_S, 0, \ldots, 0) \in \prod_{S \in \text{Irr} A} \text{BiEnd}_A(S)\) under the isomorphism of Corollary 1.34. Thus,

\[
e(S)_{S'} = \delta_{S,S'} \text{Id}_S
\]

for \(S, S' \in \text{Irr} A\). All \(e(S)\) belong to the center \(Z_A\) and they satisfy

\[
e(S)e(S') = \delta_{S,S'} e(S) \quad \text{and} \quad \sum_{S \in \text{Irr} A} e(S) = 1.
\]

The elements \(e(S)\) are called the **primitive central idempotents** of \(A\). For any \(V \in \text{Rep} A\), it follows from (1.47) that the operator \(e(S)_V\) is the identity on the \(S\)-homogeneous component \(V(S)\) and it annihilates all other homogeneous components of \(V\). Thus, the idempotent \(e(S)_V\) is the projection of \(V = \bigoplus_{S' \in \text{Irr} A} V(S')\) onto \(V(S)\):

\[
e(S)_V : V = \bigoplus_{S' \in \text{Irr} A} V(S') \xrightarrow{\text{proj}} V(S).
\]

### 1.4.6. Finite-Dimensional Irreducible Representations

In this subsection, we record some applications of the foregoing to representations of arbitrary algebras \(A\), not necessarily semisimple. Recall from (1.33) that the image \(A_V = \rho(A)\) of every representation \(\rho : A \to \text{End}_k(V)\) is contained in the double centralizer \(\text{BiEnd}_A(V) \subseteq \text{End}_k(V)\). Our focus will be on finite-dimensional representations, \(V \in \text{Rep}_{\text{fin}} A\).

**Burnside’s Theorem.** Let \(A \in \text{Alg}_k\) and let \(V \in \text{Rep}_{\text{fin}} A\). Then \(V\) is irreducible if and only if \(\text{End}_A(V)\) is a division algebra and \(A_V = \text{BiEnd}_A(V)\). In this case, \(A_V\) is isomorphic to a matrix algebra over the division algebra \(D(V)^{op}\).

**Proof.** First, assume that \(V\) is irreducible. Then \(V\) is a finite-dimensional left vector space over the Schur division algebra \(D(V) = \text{End}_A(V)\). Therefore, Lemma 1.5 implies that \(\text{BiEnd}_A(V) = \text{End}_{D(V)}(V)\) is a matrix algebra over \(D(V)^{op}\). In order to show that \(A_V = \text{BiEnd}_A(V)\), we may replace \(A\) by \(\overline{A} = A/\text{Ker} V\), because \(A_V = \overline{A}_V\) and \(\text{BiEnd}_A(V) = \text{BiEnd}_{\overline{A}}(V)\). It suffices to show that \(\overline{A}\) is semisimple; for, then
Corollary 1.34 will tell us that $\overline{A}_V = \text{BiEnd}_{\overline{A}}(V)$. Fix a $k$-basis $(v_i)_i$ of $V$. Then
\[ \{a \in \overline{A} \mid a.v_i = 0 \text{ for all } i\} = \text{Ker}_{\overline{A}}(V) = 0, \]
and hence we have an embedding
\[ \overline{A}_{\text{reg}} \hookrightarrow V^\oplus n \]
\[ a \mapsto (a.v_i)_i \]
Since $V^\oplus n$ is completely reducible, it follows from Corollary 1.29 that $\overline{A}_{\text{reg}}$ is completely reducible as well, proving that $\overline{A}$ is semisimple as desired.

Conversely, assume that $D = \text{End}_A(V)$ is a division algebra and that $A_V = \text{BiEnd}_A(V)$. Recall that $\text{BiEnd}_A(V) = \text{End}_D(V)$. Thus, $A_V = \text{End}_D(V)$ and it follows from Example 1.12 that $V$ is an irreducible representation of $A_V$. Hence $V \in \text{Irr} A$, completing the proof of Burnside’s Theorem.

\section*{Absolute Irreducibility}

Recall that the base field $k$ is said to be a splitting field for the $k$-algebra $A$ if $D(S) = k$ for all $S \in \text{Irr}_{\text{fin}} A$. We now discuss the relevance of this condition in connection with extensions of the base field (§1.2.2). Specifically, the representation $V \in \text{Rep} A$ is called \textit{absolutely irreducible} if $K \otimes V$ is an irreducible representation of $K \otimes A$ for every field extension $K/k$. Note that irreducibility of $K \otimes V$ for even one given field extension $K/k$ certainly forces $V$ to be irreducible, because any subrepresentation $0 \subset U \subset V$ would give rise to a subrepresentation $0 \subset K \otimes U \subset K \otimes V$.

\textbf{Proposition 1.36.} Let $A \in \text{Alg}_k$ and let $S \in \text{Irr}_{\text{fin}} A$. Then $S$ is absolutely irreducible if and only $D(S) = k$.

\textbf{Proof.} First assume that $D(S) = k$. Then $A_S = \text{End}_k(S)$ by Burnside’s Theorem. For any field extension $K/k$, the canonical map $K \otimes \text{End}_k(S) \to \text{End}_K(K \otimes S)$ is surjective; in fact, it is an isomorphism by (B.27). Hence, $K \otimes \rho$ maps $K \otimes A$ onto $\text{End}_K(K \otimes S)$ and so $K \otimes S$ is irreducible by Example 1.12.

Conversely, if $S$ is absolutely irreducible and $\overline{k}$ is an algebraic closure of $k$, then $\overline{k} \otimes S$ is a finite-dimensional irreducible representation of $\overline{k} \otimes A$. Hence Schur’s Lemma implies that $D(\overline{k} \otimes S) = \overline{k}$. Since $D(\overline{k} \otimes S) \cong \overline{k} \otimes D(S)$ (Exercise 1.2.4), we conclude that $D(S) = k$.

\textbf{Corollary 1.37 (Frobenius Reciprocity).} Let $\phi: A \to B$ be a homomorphism of semisimple $k$-algebras and let $S \in \text{Irr}_{\text{fin}} A$ and $T \in \text{Irr}_{\text{fin}} B$ be absolutely irreducible. Then:
\[ m(S, \text{Res}_B^A T) = m(T, \text{Ind}_B^A S). \]

\textbf{Proof.} Observe that $V := \text{Ind}_A^B S = \phi_* S \in \text{Rep} B$ is completely reducible of finite length, being a finitely generated representation of a semisimple algebra. Thus, Proposition 1.32 gives $m(T, V) = \dim_k \text{Hom}_B(V, T)$, because $D(T) = k$. Similarly,
putting $W = \text{Res}_B^A T = \phi^* T$, we obtain $m(S, W) = \dim_k \text{Hom}_A(S, W)$. Finally, $\text{Hom}_B(V, T) \cong \text{Hom}_A(S, W)$ as $k$-vector spaces by Proposition 1.9.

□

Kernels

An ideal $I$ of an arbitrary $k$-algebra $A$ will be called cofinite if $\dim_k A/I < \infty$. We put

$$\text{Spec}_{\text{cofin}} A \overset{\text{def}}{=} \{ P \in \text{Spec} A \mid \dim_k A/P < \infty \}$$

and likewise for $\text{Prim}_{\text{cofin}} A$ and $\text{MaxSpec}_{\text{cofin}} A$. The next theorem shows that all three sets coincide and that they are in bijection with $\text{Irr}_{\text{fin}} A$. Of course, if $A$ is finite dimensional, then $\text{Irr}_{\text{fin}} A = \text{Irr} A$, $\text{Spec}_{\text{cofin}} A = \text{Spec} A$, etc.

**Theorem 1.38.** All cofinite prime ideals of any $A \in \text{Alg}_k$ are maximal; so

$$\text{MaxSpec}_{\text{cofin}} A = \text{Prim}_{\text{cofin}} A = \text{Spec}_{\text{cofin}} A.$$

Moreover, there is a bijection

$$\begin{align*}
\text{Irr}_{\text{fin}} A & \longrightarrow \text{Spec}_{\text{cofin}} A \\
\uparrow & \quad \uparrow \\
S & \longmapsto \text{Ker} S
\end{align*}$$

**Proof.** In view of the general inclusions (1.40), the equalities $\text{MaxSpec}_{\text{cofin}} A = \text{Prim}_{\text{cofin}} A = \text{Spec}_{\text{cofin}} A$ will follow if we can show that any $P \in \text{Spec}_{\text{cofin}} A$ is in fact maximal. For this, after replacing $A$ by $A/P$, we may assume that $A$ is a finite-dimensional algebra that is prime (i.e., the product of any two nonzero ideals of $A$ is nonzero) and we must show that $A$ is simple. Choose a minimal nonzero left ideal $L \subseteq A$. Then $L \in \text{Irr}_{\text{fin}} A$. Furthermore, since $A$ is prime and $I = LA$ is a nonzero ideal of $A$ with $(\text{Ker} L) I = 0$, we must have $\text{Ker} L = 0$. Therefore, $A \cong A_L$ and so Burnside’s Theorem implies that $A$ is isomorphic to a matrix algebra over some division algebra, whence $A$ is simple (Exercise 1.1.13).

For the asserted bijection with $\text{Irr}_{\text{fin}} A$, note that an irreducible representation $S$ is finite dimensional if and only if $\text{Ker} S$ is cofinite. Therefore, the surjection $\text{Irr} A \twoheadrightarrow \text{Prim} A$ in (1.35) restricts to a surjection $\text{Irr}_{\text{fin}} A \twoheadrightarrow \text{Spec}_{\text{cofin}} A$. In order to show that this map is also injective, let $S, S' \in \text{Irr}_{\text{fin}} A$ be such that $\text{Ker} S = \text{Ker} S'$ and let $\overline{A}$ denote the quotient of $A$ modulo this ideal. Then $S, S' \in \text{Irr} \overline{A}$ and $\overline{A}$ is isomorphic to a matrix algebra over some division algebra by Burnside’s Theorem. Since such algebras have only one irreducible representation up to equivalence by Wedderburn’s Structure Theorem, we must have $S \cong S'$ in $\text{Rep} \overline{A}$ and hence in $\text{Rep} A$ as well.

□

1.4.7. Finite-Dimensional Algebras

The following theorem gives an ideal theoretic characterization of semisimplicity for finite-dimensional algebras.
**Theorem 1.39.** The following are equivalent for a finite-dimensional algebra $A$:

(i) $A$ is semisimple;

(ii) $\text{rad } A = 0$;

(iii) $A$ has no nonzero nilpotent right or left ideals.

**Proof.** If $A$ is semisimple, then $A_{\text{reg}}$ is a sum of irreducible representations. Since $(\text{rad } A).S = 0$ for all $S \in \text{Irr } A$, it follows that $(\text{rad } A).A_{\text{reg}} = 0$ and so $\text{rad } A = 0$. Thus (i) implies (ii). In view of Proposition 1.25, (ii) and (iii) are equivalent. It remains to show that (ii) implies (i). So assume that $\text{rad } A = \bigcap_{S \in \text{Irr } A} \text{Ker } S = 0$. Since $A$ is finite dimensional, some finite intersection $\bigcap_{i=1}^r \text{Ker } S_i$ must be 0, the $\text{Ker } S_i$ being pairwise distinct maximal ideals of $A$ (Theorem 1.38). The Chinese Remainder Theorem (e.g., [24, Proposition 9 on p. A.1.104]) yields an isomorphism of algebras $A \cong \prod_{i=1}^r A/\text{Ker } S_i$ and Burnside’s Theorem (§1.4.6) further tells us that each $A/\text{Ker } S_i$ is a matrix algebra over a division algebra. Semisimplicity of $A$ now follows from Wedderburn’s Structure Theorem. This proves the theorem. □

Condition (iii), for an arbitrary algebra $A$, is equivalent to semiprimeness of the zero ideal of $A$; see the proof of Proposition 1.25. Such algebras are called semiprime. Similarly, any algebra whose zero ideal is prime is called prime and likewise for “primitive.” For a finite-dimensional algebra $A$, the properties of being prime, primitive, or simple are all equivalent by Theorem 1.38 and Theorem 1.39 gives the same conclusion for the properties of being semiprime, semiprimitive, or semisimple. We will refer to the algebra $A^{s.p.} = A/\text{rad } A$, which is always semiprimitive by (1.42), as the **semisimplification** of $A$ when $A$ is finite dimensional.

**Exercises for Section 1.4**

*In these exercises, $A$ denotes an $k$-algebra.*

**1.4.1** (Radical of a representation). The **radical** of a representation $V \in \text{Rep } A$ is defined by

$$\text{rad } V \overset{\text{def}}{=} \bigcap \{ \text{all maximal subrepresentations of } V \}.$$  

Here, a subrepresentation $M \subseteq V$ is called maximal if $V/M$ is irreducible. The empty intersection is understood to be equal to $V$. Thus, the Jacobson radical $\text{rad } A$ is the same as $\text{rad } A_{\text{reg}}$; see Exercise 1.3.3. Prove:

(a) If $V$ is finitely generated, then $\text{rad } V \subseteq V$. (Use Exercise 1.1.3.)

(b) If $V$ is completely reducible, then $\text{rad } V = 0$. Give an example showing that the converse need not hold.

(c) If $U \subseteq V$ is a subrepresentation such that $V/U$ is completely reducible, then $U \supseteq \text{rad } V$. If $V$ is artinian (see Exercise 1.2.10), then the converse holds.

(d) $(\text{rad } A).V \subseteq \text{rad } V$; equality holds if $A$ is finite dimensional.
1.4.2 (Matrix algebras). Let $S \in \text{Irr} A$. Viewing $S^\otimes n$ as a representation of $\text{Mat}_n(A)$ as in Lemma 1.4(b), prove that $S^\otimes n$ is irreducible.

1.4.3 (Semisimple algebras). (a) Show that the algebra $A$ is semisimple if and only if all short exact sequences in $\text{Rep} A$ are split. (See Exercise 1.1.2).

(b) Let $A \cong \prod_{i=1}^t \text{Mat}_{m_i}(D_i)$ be semisimple. Describe the ideals $I$ of $A$ and the factors $A/I$. Show that $A$ has exactly $t$ prime ideals and that there is a bijection $\text{Irr} A \longleftrightarrow \text{Spec} A, S \leftrightarrow \text{Ker} S$. Moreover, all ideals $I$ of $A$ are idempotent: $I^2 = I$.

1.4.4 (Faithful completely reducible representations). Assume that $V \in \text{Rep} A$ is faithful and completely reducible. Show:

(a) The algebra $A$ is semiprime. In particular, if $A$ is finite dimensional, then $A$ is semisimple.

(b) If $V$ is finite dimensional, then $A$ is finite dimensional and semisimple. Moreover, $\text{Irr} A$ is the set of distinct irreducible constituents of $V$.

(c) The conclusion of (b) fails if $V$ is not finite dimensional: $A$ need not be semisimple.

1.4.5 (Galois descent). Let $K$ be a field and let $\Gamma$ be a subgroup of $\text{Aut}(K)$. Let $K = K^\Gamma$ be the field of $\Gamma$-invariants in $K$. For a given $V \in \text{Vec}_K$, consider the action $\Gamma \subseteq K \otimes V$ with $\gamma \in \Gamma$ acting by $\gamma(\lambda \otimes v) = (\gamma \otimes \text{Id}_V)(\lambda \otimes v) = \gamma(\lambda) \otimes v$. View $V$ as a $K$-subspace of $K \otimes V$. \[ \text{Irr}(V) \otimes \Gamma = \{ x \in K \otimes V \mid \gamma(x) = x \text{ for all } \gamma \in \Gamma \}, \]

(a) Show that $V = (K \otimes V)^\Gamma = \{ x \in K \otimes V \mid \gamma(x) = x \text{ for all } \gamma \in \Gamma \}$, the space of $\Gamma$-invariants in $K \otimes V$.

(b) Let $W \subseteq K \otimes V$ be any $\Gamma$-stable $K$-subspace. Show that $W = K \otimes W^\Gamma$, where $W^\Gamma = W \cap V$ is the space of $\Gamma$-invariants in $W$.

1.4.6 (Galois action). Let $K/\mathbb{k}$ be a finite Galois extension with Galois group $\Gamma = \text{Gal}(K/\mathbb{k})$. Show:

(a) The kernel of any $f \in \text{Hom}_{\mathbb{k}}(A, K)$ belongs to $\text{MaxSpec} A$.

(b) For $f, f' \in \text{Hom}_{\mathbb{k}}(A, K)$, we have $\text{Ker} f = \text{Ker} f'$ if and only if $f' = \gamma \circ f$ for some $\gamma \in \Gamma$. (You can use Exercise 1.4.5 for this.)

1.4.7 (Extension of scalars and complete reducibility). Let $V \in \text{Rep} A$. For a given field extension $K/\mathbb{k}$, consider the representation $K \otimes V \in \text{Rep} (K \otimes A)$ ($\S 1.2.2$). Prove (using Exercise 1.4.5):

(a) If $K \otimes V$ is completely reducible, then so is $V$.

(b) If $V$ is irreducible and $K/\mathbb{k}$ is finite separable, then $K \otimes V$ is completely reducible of finite length. (Reduce to the case where $K/\mathbb{k}$ is Galois and use Exercise 1.4.5.) Give an example showing that $K \otimes V$ need not be irreducible.

(c) If the field $\mathbb{k}$ is perfect and $V$ is finite dimensional and completely reducible, then $K \otimes V$ is completely reducible for every field extension $K/\mathbb{k}$.
1.4.8 (Extension of scalars and composition factors). Let \( V, W \in \text{Rep}_{\text{fin}} A \) and let \( K/\mathbb{k} \) be a field extension. Prove:

(a) If \( V \) and \( W \) are irreducible and non-equivalent, then the representations \( K \otimes V \) and \( K \otimes W \) of \( K \otimes A \) have no common composition factor.

(b) Conclude from (a) that, in general, \( K \otimes V \) and \( K \otimes W \) have a common composition factor if and only if \( V \) and \( W \) have a common composition factor.

1.4.9 (Splitting fields). Assume that \( A \) is finite-dimensional and defined over some subfield \( \mathbb{k}_0 \subseteq \mathbb{k} \), that is, \( A \cong \mathbb{k} \otimes_{\mathbb{k}_0} A_0 \) for some \( \mathbb{k}_0 \)-algebra \( A_0 \). Assume further that \( \mathbb{k} \) is a splitting field for \( A \). For a given field \( F \) with \( \mathbb{k}_0 \subseteq F \subseteq \mathbb{k} \), consider the \( F \)-algebra \( B = F \otimes_{\mathbb{k}_0} A_0 \). Show that \( F \) is a splitting field for \( B \) if and only if each \( S \in \text{Irr} A \) is defined over \( F \), that is, \( S \cong \mathbb{k} \otimes_F T \) for some \( T \in \text{Irr} B \).

1.4.10 (Separable algebras). The algebra \( A \) is called \textit{separable} if there exists an element \( e = \sum_i x_i \otimes y_i \in A \otimes A \) satisfying \( m(e) = 1 \) and \( ae = ea \) for all \( a \in A \). Here, \( m : A \otimes A \to A \) is the multiplication map and \( A \otimes A \) is viewed as \((A, A)\)-bimodule using left multiplication on the first factor and right multiplication on the second; so the conditions on \( e \) are: \( \sum x_i y_i = 1 \) and \( \sum_i a x_i \otimes y_i = \sum_i x_i \otimes y_i a \).

(a) Assuming \( A \) to be separable, show that \( K \otimes A \) is separable for every field extension \( K/\mathbb{k} \). Conversely, if \( K \otimes A \) is separable for some \( K/\mathbb{k} \), then \( A \) is separable.

(b) With \( e \) as above, show that the products \( x_i y_j \) generate \( A \) as \( \mathbb{k} \)-vector space. Thus, separable \( \mathbb{k} \)-algebras are finite dimensional.

(c) For \( V, W \in \text{Rep} A \), consider \( \text{Hom}_\mathbb{k}(V, W) \in \text{Rep} (A \otimes A^{op}) \) as in Example 1.3 and view \( e \in A \otimes A^{op} \). Show that \( e f \in \text{Hom}_A(V, W) \) for all \( f \in \text{Hom}_\mathbb{k}(V, W) \). Furthermore, assuming \( W \) to be a subrepresentation of \( V \) and \( f|_W = \text{Id}_W \), show that \( (e f)|_W = \text{Id}_W \). Conclude that separable algebras are semisimple.

(d) Conclude from (a)–(c) that \( A \) is separable if and only if \( A \) is finite dimensional and \( K \otimes A \) is semisimple for every field extension \( K/\mathbb{k} \).

1.5. Characters

The focus in this section continues to be on the finite-dimensional representations of a given \( \mathbb{k} \)-algebra \( A \). Analyzing a typical \( V \in \text{Rep}_{\text{fin}} A \) can be a daunting task, especially if the dimension of \( V \) is large. In this case, explicit computations with the operators \( a_V \ (a \in A) \) involve prohibitively complex matrix operations. Fortunately, it turns out that a surprising amount of information can be gathered from just the traces of all \( a_V \); these traces form the so-called character of \( V \). For example, we shall see that if \( V \) is completely reducible and \( \text{char} \ \mathbb{k} = 0 \), then the representation \( V \) is determined up to equivalence by its character (Theorem 1.45). Before proceeding, the reader may wish to have a quick look at Appendix B for the basics concerning traces of linear operators on a finite-dimensional vector space.

\textit{Throughout this section,} \( A \) \textit{denotes an arbitrary} \( \mathbb{k} \)-\textit{algebra.}
1.5.1. Definition and Basic Properties

The **character** of $V \in \text{Rep}_{\text{fin}} A$ is the linear form on $A$ that is defined by

$$\chi_V : A \rightarrow \mathbb{k}$$

$$a \mapsto \text{trace}(a_V)$$

(1.50)

Characters tend to be most useful if $\text{char}\mathbb{k} = 0$; the following example gives a first illustration of this fact.

**Example 1.40** (The regular character). If the algebra $A$ is finite dimensional, then we can consider the character of the regular representation $A_{\text{reg}}$.

$$\chi_{\text{reg}} \overset{\text{def}}{=} \chi_{A_{\text{reg}}} : A \rightarrow \mathbb{k}$$

The regular character $\chi_{\text{reg}}$ is also denoted by $T_{A/\mathbb{k}}$ when $A$ or $\mathbb{k}$ need to be made explicit. For the matrix algebra $A = \text{Mat}_n(\mathbb{k})$, one readily checks (Exercise 1.5.2) that

$$\chi_{\text{reg}} = n \text{ trace} : \text{Mat}_n(\mathbb{k}) \rightarrow \mathbb{k}.$$ 

In particular, the regular character $\chi_{\text{reg}}$ of $\text{Mat}_n(\mathbb{k})$ vanishes if $\text{char}\mathbb{k}$ divides $n$. If $K/\mathbb{k}$ is a finite field extension, then we may view $K$ as a finite-dimensional $\mathbb{k}$-algebra. All finite-dimensional representations of $K$ are equivalent to $K^{\otimes n}_{\text{reg}}$ for suitable $n$. It is a standard fact from field theory that $\chi_{\text{reg}} \neq 0$ if and only if the extension $K/\mathbb{k}$ is separable (Exercise 1.5.5). Thus, by the lemma below, all characters $\chi_V$ of the $\mathbb{k}$-algebra $K$ vanish if $K/\mathbb{k}$ is not separable.

**Additivity.** The following lemma states a basic property of characters: additivity on short exact sequences of representations.

**Lemma 1.41.** If $0 \rightarrow U \rightarrow V \rightarrow W \rightarrow 0$ is a short exact sequence in $\text{Rep}_{\text{fin}} A$, then $\chi_V = \chi_U + \chi_W$.

**Proof.** First note that if $f : U \cong V$ is an isomorphism of finite-dimensional representations, then $\chi_V = \chi_U$. Indeed, $a_V = f \circ a_U \circ f^{-1}$ holds for all $a \in A$ by (1.22), and hence $\text{trace}(a_V) = \text{trace}(f \circ a_U \circ f^{-1}) = \text{trace}(a_U \circ f^{-1} \circ f) = \text{trace}(a_U)$.

Now let $0 \rightarrow U \xrightarrow{f} V \xrightarrow{g} W \rightarrow 0$ be a short exact sequence of finite-dimensional representations. Thus, $\text{Im} f$ is an $A$-submodule of $V$ such that $U \cong \text{Im} f$ and $W \cong V/\text{Im} f$ as $A$-modules. In view of the first paragraph, we may assume that $U$ is an $A$-submodule of $V$ and $W = V/U$. Extending a $\mathbb{k}$-basis of $U$ to
a $k$-basis of $V$, the matrix of each $a_V$ has block upper triangular form:

$$
\begin{pmatrix}
  a_U & * \\
  0 & a_{V/U}
\end{pmatrix}.
$$

Taking traces, we obtain $\text{trace}(a_V) = \text{trace}(a_U) + \text{trace}(a_{V/U})$ as desired. \hfill \Box

**Multiplicativity.** Let $V \in \text{Rep} A$ and $W \in \text{Rep} B$; so we have algebra maps $A \to \text{End}_k(V)$ and $B \to \text{End}_k(W)$. By bi-functoriality of the tensor product of algebras (Exercise 1.1.10), we obtain an algebra map $A \otimes B \to \text{End}_k(V) \otimes \text{End}_k(W)$. Composing this map with the canonical map $\text{End}_k(V) \otimes \text{End}_k(W) \to \text{End}_k(V \otimes W)$ in (B.17), which is also a map in $\text{Alg}_k$, we obtain the algebra map

$$
A \otimes B \longrightarrow \text{End}_k(V \otimes W)
$$

(1.51)

$$
a \otimes b \longmapsto a_V \otimes b_W
$$

This makes $V \otimes W$ a representation of $A \otimes B$, called the **outer tensor product** of $V$ and $W$; it is sometimes denoted by $V \boxtimes W$. If $V$ and $W$ are finite dimensional, then (B.26) gives

$$
\chi_{V \otimes W}(a \otimes b) = \chi_V(a) \chi_W(b).
$$

(1.52)

**1.5.2. Spaces of Trace Forms**

Each character $\chi_V$ ($V \in \text{Rep}_{\text{fin}} A$) is a linear form on $A$, but more can be said:

(i) By virtue of the standard trace identity, $\text{trace}(f \circ g) = \text{trace}(g \circ f)$ for $f, g \in \text{End}_k(V)$, all characters vanish on the $k$-subspace $[A, A] \subseteq A$ that is spanned by the Lie commutators $[a, a'] = aa' - a'a$ with $a, a' \in A$.

(ii) The character $\chi_V$ certainly also vanishes on $\text{Ker} V$; note that this is a cofinite ideal of $A$.

(iii) In fact, $\chi_V$ vanishes on the semiprime radical $\sqrt{\text{Ker} V}$ (as defined in Exercise 1.3.1). To see this, note that $\sqrt{\text{Ker} V}$ coincides with the preimage of $\text{rad}(A/\text{Ker} V)$ in $A$ by Theorem 1.38 and so some power of $\sqrt{\text{Ker} V}$ is contained in $\text{Ker} V$ by Proposition 1.25. Therefore, all elements of $\sqrt{\text{Ker} V}$ act as nilpotent endomorphisms on $V$. Since nilpotent endomorphisms have trace 0, it follows that $\sqrt{\text{Ker} V} \subseteq \text{Ker} \chi_V$.

Below, we will formalize these observations.
Universal Trace and Trace Forms. By (i), each character factors through the canonical map

\[
\text{Tr}: A \rightarrow A / \{[A, A] \}
\]

(1.53)

The map \(\text{Tr}\) satisfies the trace identity \(\text{Tr}(aa') = \text{Tr}(a'a)\) for \(a, a' \in A\); it will be referred to as the universal trace of \(A\). Note that \(\text{Tr}\) gives a functor \(\text{Alg}_k \rightarrow \text{Vect}_k\), because any \(k\)-algebra map \(\phi: A \rightarrow B\) satisfies \(\phi([A, A]) \subseteq [B, B]\), and hence \(\phi\) passes down to a \(k\)-linear homomorphism \(\text{Tr} \phi: \text{Tr} A \rightarrow \text{Tr} B\).

We will identify the linear dual \((\text{Tr} A)^*\) with the subspace of \(A^*\) consisting of all linear forms on \(A\) that vanish on \([A, A]\). This subspace will be referred to as the space of trace forms on \(A\) and denoted by \(A^*_{\text{trace}}\).

Finite Trace Forms. By (ii) and (iii), all characters belong to the following subspaces of \(A^*_{\text{trace}}\):

\[
A_{\text{trace}}^\circ \overset{\text{def}}{=} \{ f \in A^*_{\text{trace}} \mid f \text{ vanishes on some cofinite ideal of } A \}
\]

\[
C(A) \overset{\text{def}}{=} \{ t \in A^*_{\text{trace}} \mid t \text{ vanishes on some cofinite semiprime ideal of } A \}
\]

To see that these are indeed subspaces of \(A^*_{\text{trace}}\), observe that the intersection of any two cofinite (semiprime) ideals is again cofinite (semiprime). We will call \(A_{\text{trace}}^\circ\) the space of finite trace forms on \(A\) in reference to the so-called finite dual, \(A^\circ = \{ f \in A^* \mid f \text{ vanishes on some cofinite ideal of } A \}\), which will play a prominent role in Part IV. As above, it is easy to see that \(A^\circ\) and \(A_{\text{trace}}^\circ\) are contravariant functors \(\text{Alg}_k \rightarrow \text{Vect}_k\).

To summarize, for any \(V \in \text{Rep}_\text{fin} A\), we have

\[
\chi_V \in C(A) \subseteq A^\circ_{\text{trace}} \subseteq A^*_{\text{trace}} \subseteq A^*. 
\]

(1.54)

We will see shortly that, if \(k\) is a splitting field for \(A\), then \(C(A)\) is the subspace of \(A^*\) that is spanned by the characters of all finite-dimensional representations of \(A\) (Theorem 1.44).
The Case of Finite-Dimensional Algebras. Now assume that $A$ is finite dimensional. Then all ideals are cofinite and the Jacobson radical $\text{rad} A$ is the unique smallest semiprime ideal of $A$ (Proposition 1.25). Therefore, $C(A)$ is exactly the subspace of $A^*_\text{trace} = A^*_\text{trace}$ consisting of all trace forms that vanish on $\text{rad} A$. The latter space may be identified with $(A^{s.p.})^*_\text{trace}$, where $A^{s.p.} = A/\text{rad} A$ is the semisimplification of $A$. So

$$C(A) \cong (A/[A, A] + \text{rad} A)^* \cong (\text{Tr} A^{s.p.})^* \cong (A^{s.p.})^*_\text{trace}$$

1.5.3. Algebras in Positive Characteristics

This subsection focuses on the case where $\text{char} \mathbb{k} = p > 0$. Part (a) of the following lemma shows that the $p^\text{th}$ power map gives an additive endomorphism of $\text{Tr} A = A/[A, A]$, often called the Frobenius endomorphism. For part (b), we put

$$T(A) \overset{\text{def}}{=} \{ a \in A \mid a^{p^n} \in [A, A] \text{ for some } n \in \mathbb{Z}_+ \}.$$  

Lemma 1.42. Assume that $\text{char} \mathbb{k} = p > 0$. Then:

(a) $(a + b)^p \equiv a^p + b^p \mod [A, A]$ for all $a, b \in A$. Furthermore, $a \in [A, A]$ implies $a^p \in [A, A]$.

(b) $T(A)$ is a $\mathbb{k}$-subspace of $A$ containing $[A, A]$. If $I$ is a nilpotent ideal of $A$, then $T(A)/I \cong T(A/I)$ via the canonical map $A \to A/I$.

Proof. (a) Expanding $(a + b)^p$, we obtain the sum of all $2^p$ products of length $p$ with factors equal to $a$ or $b$. The cyclic group $C_p$ acts on the set of these products by cyclic permutations of the factors. There are two fixed points, namely the products $a^p$ and $b^p$; all other orbits have size $p$. Modulo $[A, A]$, the elements of each orbit are identical, because $x_1x_2\ldots x_s \equiv x_sx_1\ldots x_{s-1} \mod [A, A]$. Since $\text{char} \mathbb{k} = p$, the orbits of size $p$ contribute 0 to the sum modulo $[A, A]$ and we are left with $(a + b)^p \equiv a^p + b^p \mod [A, A]$. This proves the first assertion of (a). Next, we show that $a \in [A, A]$ implies $a^p \in [A, A]$. By the foregoing, we may assume that $a = xy - yx$ for some $x, y \in A$. Calculating modulo $[A, A]$, we have $a^p \equiv (xy)^p - (yx)^p = [x, z]$ with $z = y(xy)^{p-1}$; so $a^p \equiv 0 \mod [A, A]$ as desired.

(b) Let $\phi = ,^p$ denote the Frobenius endomorphism of $\text{Tr} A$ as provided by part (a). Then $0 = \text{Ker} \phi^0 \subseteq \text{Ker} \phi \subseteq \cdots \subseteq \text{Ker} \phi^n \subseteq \cdots$ is a chain of subgroups of $\text{Tr} A$ and each $\text{Ker} \phi^n$ is also stable under scalar multiplication, because $\phi^n(\lambda x) = \lambda^p \phi^n(x)$ for $x \in \text{Tr} A$ and $\lambda \in \mathbb{k}$. Therefore, the union of all $\text{Ker} \phi^n$ is a $\mathbb{k}$-subspace of $\text{Tr} A$ and the preimage of this subspace under the canonical map $A \to \text{Tr} A$ is a $\mathbb{k}$-subspace of $A$. This subspace is $T(A)$. Finally, let $I$ be a nilpotent ideal of $A$. Then $I$ is clearly contained in $T(A)$ and $T(A)$ maps to $T(A/I)$ under the canonical map. For surjectivity, let $z = a + I \in T(A/I)$. Then $z^{p^n} = a^{p^n} + I \in [A/I, A/I]$ for some $n$ and so $a^{p^n} \in [A, A] + I$. Part (a) gives
Suppose that

\[(A, A) + I P^n \subseteq [A, A] + I P^m \text{ for all } m \in \mathbb{Z}_+.\]

Choosing \(m\) large enough, we have \(I P^n = 0\). Hence \(a P^n \in [A, A]\), proving that \(a \in T(A)\) as desired. \(\square\)

Returning to our discussion of trace forms, we now give a description of \(C(A)\) for a finite-dimensional algebra \(A\) over a field \(k\) with \(\text{char } k = p > 0\). We also assume that \(k\) is a splitting field for \(A\) in the sense of §1.2.5.

**Proposition 1.43.** Let \(A\) be finite-dimensional and assume that \(\text{char } k = p > 0\) and that \(k\) is a splitting field for \(A\). Then \(C(A) \cong (A/T(A))^\circ\).

**Proof.** In light of (1.55), we need to show that \(A/T(A) \cong \text{Tr } A^{k_p}\). But \(\text{rad } A\) is nilpotent (Proposition 1.25), and so \((A/T(A))/\text{rad } A \cong T(A^{k_p})\) by Lemma 1.42. Therefore, it suffices to show that \(T(A^{k_p}) = [A^{k_p}, A^{k_p}]\). By assumption of \(k\), the Wedderburn decomposition \(A^{k_p} \cong \prod_{S \in \text{Irr } A} A_S\) has components \(A_S\) that are matrix algebras over \(k\). Clearly, \(T(A^{k_p}) \cong \prod_{S} T(A_S)\) and \([A^{k_p}, A^{k_p}] \cong \prod_{S} [A_S, A_S]\). Thus, it suffices to consider a matrix algebra \(\text{Mat}_d(k)\). We have seen in the proof of Corollary 1.35 that the space \([\text{Mat}_d(k), \text{Mat}_d(k)]\) consists of all trace-0 matrices and hence has codimension 1 in \(\text{Mat}_d(k)\). Since the idempotent matrix \(e_{1,1}\) does not belong to \(T(\text{Mat}_d(k))\), it follows that \(T(\text{Mat}_d(k)) = [\text{Mat}_d(k), \text{Mat}_d(k)]\) finishing the proof. \(\square\)

### 1.5.4. Irreducible Characters

Characters of finite-dimensional irreducible representations are referred to as **irreducible characters**. Since every finite-dimensional representation has a composition series, all characters are sums of irreducible characters by Lemma 1.41. The following theorem lists some important properties of the collection of irreducible characters of an arbitrary algebra \(A \in \text{Alg}_k\).

**Theorem 1.44.**

(a) The irreducible characters \(\chi_S\) for \(S \in \text{Irr}_\text{fin} A\) such that \(\text{char } k\) does not divide \(\dim_k D(S)\) are linearly independent.

(b) \# \text{Irr}_\text{fin} A \leq \dim_k C(A).

(c) If \(k\) is a splitting field for \(A\), then the irreducible characters of \(A\) form a \(k\)-basis of \(C(A)\). In particular, \# \text{Irr}_\text{fin} A = \dim_k C(A)\) holds in this case.

**Proof.** (a) Suppose that \(\sum_{i=1}^l \lambda_i \chi_{S_i} = 0\) with \(\lambda_i \in k\) and distinct \(S_i \in \text{Irr}_\text{fin} A\) such that \(\text{char } k\) does not divide \(\dim_k D(S_i)\). We need to show that all \(\lambda_i = 0\). By Theorem 1.38, the annihilators \(\text{Ker } S_i\) are distinct maximal ideals of \(A\), and \(A/\text{Ker } S_i \cong B_i := \text{Mat}_{m_i}(D_i)\) by Burnside’s Theorem (§1.4.6), with \(D_i = D(S_i)^{\text{op}}\).
The Chinese Remainder Theorem yields a surjective homomorphism of algebras,

\[
A \longrightarrow A/\bigcap_{i=1}^{r} \text{Ker} S_i \cong B := \prod_{i=1}^{r} B_i
\]

Let \( e_j \in B \) be the element corresponding to the \( t \)-tuple with the matrix \( e_{1,1} \in \text{Mat}_{m_i}(D_i) \) in the \( j \)-th component and the 0-matrix in all other components. It is easy to see (Exercise 1.5.2(a)) that \( \chi_{S_i}(e_j) = \dim_k(D_j) \delta_{i,j} \), because \( A \) acts on each \( S_i \) via the standard action of \( \text{Mat}_{m_i}(D_i) \) on \( D_i^{\otimes m_i} \). We conclude that 

\[
0 = \sum_j \lambda_j \dim_k(D_j) \delta_{i,j} \quad \text{for all } j.
\]

Finally, our hypothesis implies that \( \dim_k(D_j) \delta_{i,j} \neq 0 \), giving the desired conclusion \( \lambda_j = 0 \).

(b) Let \( S_1, \ldots, S_r \) be nonequivalent finite-dimensional irreducible representations of \( A \) and consider the epimorphism (1.56). Since \( B \) is finite dimensional and semiprime, we have \( B_{\text{trace}} \hookrightarrow C(A) \). Moreover, clearly, \( B_{\text{trace}} \cong \bigoplus_i (B_i)_{\text{trace}} \). Thus, it suffices to show that every finite-dimensional \( \overline{k} \)-algebra \( B \) has a nonzero trace form. To see this, consider the algebra \( \overline{B} = B \otimes \overline{k} \), where \( \overline{k} \) is an algebraic closure of \( k \), and fix some \( \overline{S} \in \text{Irr} \overline{B} \). The character \( \chi_{\overline{S}} \) is nonzero by (a) and its restriction to \( B \) is nonzero as well, because the canonical image of \( B \) generates \( \overline{B} \) as \( \overline{k} \)-vector space. Composing \( \chi_{\overline{S}} \) with a suitable \( k \)-linear projection of \( \overline{k} \) onto \( k \) yields the desired trace form.

(c) Since \( D(S) = k \) for all \( S \in \text{Irr}_{\text{fin}} \) by hypothesis, the irreducible characters \( \chi_S \) are linearly independent by (a). It suffices to show that the \( \chi_S \) span \( C(A) \). But any \( t \in C(A) \) vanishes some cofinite semiprime ideal \( I \) of \( A \). The algebra \( A/I \) is split semisimple by Theorem 1.39 and the trace form \( t \) can be viewed as an element of \( C(A/I) = (A/I)_{\text{trace}}^* \). Since \( \# \text{Irr} A/I = \dim_k (A/I)_{\text{trace}}^* \) by Corollary 1.35(a), we know that \( t \) is a linear combination of irreducible characters of \( A/I \). Viewing these characters as (irreducible) characters of \( A \) by inflation, we have written \( t \) as a linear combination of irreducible characters of \( A \). This completes the proof. \( \square \)

Characters of Completely Reducible Representations. It is a fundamental fact of representation theory that, under some restrictions on \( \text{char } k \), finite-dimensional completely reducible representations are determined up to equivalence by their character. In detail:

**Theorem 1.45.** Let \( V, W \in \text{Rep}_{\text{fin}} A \) be completely reducible and assume that \( \text{char } k = 0 \) or \( \text{char } k > \max \{ \dim_k V, \dim_k W \} \). Then \( V \cong W \) if and only if \( \chi_V = \chi_W \).

**Proof.** Since \( V \cong W \) clearly always implies \( \chi_V = \chi_W \) (Lemma 1.41), let us assume \( \chi_V = \chi_W \) and prove that \( V \cong W \). To this end, write \( V \cong \bigoplus_{S \in \text{Irr}_{\text{fin}}} A^{\otimes m_S} \) and \( W \cong \bigoplus_{S \in \text{Irr}_{\text{fin}}} A^{\otimes n_S} \) with \( m_S = m(S, V) \) and \( n_S = m(S, W) \). Lemma 1.41 gives \( \chi_V = \sum_S m_S \chi_S \) and \( \chi_W = \sum_S n_S \chi_S \), and we need to show that \( m_S = n_S \) for all \( S \).
Certain aspects of the representation theory of a given \( k \)-algebra \( A \), especially those related to characters, can be conveniently packaged with the aid of the Grothendieck group of finite-dimensional representations of \( A \). This group will be denoted\(^8\) by \( \mathcal{R}(A) \).

By definition, \( \mathcal{R}(A) \) is the abelian group with generators \([V]\) for \( V \in \text{Rep}_{\text{fin}} A \) and with relations

\[
[V] = [U] + [W]
\]

for each short exact sequence \( 0 \to U \to V \to W \to 0 \) in \( \text{Rep}_{\text{fin}} A \). Formally, \( \mathcal{R}(A) \) is the factor of the free abelian group on the set of all isomorphism classes \( (V) \) of finite-dimensional representations \( V \) of \( A \) — these isomorphism classes do indeed form a set—modulo the subgroup that is generated by the elements \( (V) - (U) - (W) \) arising from short exact sequences \( 0 \to U \to V \to W \to 0 \) in \( \text{Rep}_{\text{fin}} A \). The generator \([V]\) is the image of \((V)\) in \( \mathcal{R}(A) \). The point of this construction is as follows. Suppose we have a rule assigning to each \( V \in \text{Rep}_{\text{fin}} A \) a value \( f(V) \) in some abelian group \((\mathcal{G}, +)\) in such a way that the assignment is additive on short exact sequences in \( \text{Rep}_{\text{fin}} A \) in the sense that exactness of \( 0 \to U \to V \to W \to 0 \) implies that \( f(V) = f(U) + f(W) \) holds in \( \mathcal{G} \). Then we obtain a well-defined group homomorphism

\[
f : \mathcal{R}(A) \to \mathcal{G}, \quad [V] \mapsto f(V).
\]

We will use the above construction for certain classes of representations other than the objects of \( \text{Rep}_{\text{fin}} A \) in \$2.1.3\; ; for a discussion of Grothendieck groups in great generality, the reader may wish to consult \cite{Swan1969, Bourbaki1958}.

**Group Theoretical Structure.** If \( 0 = V_0 \subseteq V_1 \subseteq \cdots \subseteq V_i = V \) is any chain of finite-dimensional representations of \( A \), then the relations of \( \mathcal{R}(A) \) and a straightforward induction imply that

\[
[V] = \sum_{i=1}^{l} [V_i/V_{i-1}].
\]

In particular, taking a composition series for \( V \), we see that \( \mathcal{R}(A) \) is generated by the elements \([S]\) with \( S \in \text{Irr}_{\text{fin}} A \). In fact, these generators are \( \mathbb{Z} \)-independent:

\(^8\)Other notations are also used in the literature; for example, \( \mathcal{R}(A) \) is denoted by \( G_0^A(\text{Rep}_A) \) in Swan \cite{Swan1969} and by \( R_k(A) \) in Bourbaki \cite{Bourbaki1958}.
1.5. Characters

**Proposition 1.46.** The group $\mathcal{R}(A)$ is isomorphic to the free abelian group with basis the set $\text{Irr}_{\text{fin}} A$ of isomorphism classes of finite-dimensional irreducible representations of $A$. An explicit isomorphism is given by multiplicities,

$$
\mu: \mathcal{R}(A) \longrightarrow \mathbb{Z}^{\oplus \text{Irr}_{\text{fin}} A}
$$

In particular, for $V, W \in \text{Rep}_{\text{fin}} A$, the equality $[V] = [W]$ holds in $\mathcal{R}(A)$ if and only if $\mu(S, V) = \mu(S, W)$ for all $S \in \text{Irr}_{\text{fin}} A$.

**Proof.** The map $\mu$ yields a well-defined group homomorphism by virtue of the fact that multiplicities are additive on short exact sequences by (1.32). For $S \in \text{Irr}_{\text{fin}} A$, one has $\mu([S]) = (\delta_{S,T})_T$. These elements form the standard $\mathbb{Z}$-basis of $\mathbb{Z}^{\oplus \text{Irr}_{\text{fin}} A}$. Therefore, the generators $[S]$ are $\mathbb{Z}$-independent and $\mu$ is an isomorphism. \(\square\)

**Functoriality.** Pulling back representations along a given algebra map $\phi: A \rightarrow B$ (§1.2.2) turns short exact sequences in $\text{Rep}_{\text{fin}} B$ into short exact sequences in $\text{Rep}_{\text{fin}} A$. Therefore, we obtain a group homomorphism

$$
\mathcal{R}(\phi): \mathcal{R}(B) \longrightarrow \mathcal{R}(A)
$$

In this way, we may view $\mathcal{R}$ as a contravariant functor from $\text{Alg}_k$ to the category $\text{AbGroups} \equiv \mathbb{Z}\text{Mod}$ of all abelian groups.

**Lemma 1.47.** Let $\phi: A \rightarrow B$ be a surjective algebra map. Then $\mathcal{R}(\phi)$ is a split injection coming from the inclusion $\phi^*: \text{Irr}_{\text{fin}} B \hookrightarrow \text{Irr}_{\text{fin}} A$:

$$
\mathcal{R}(B) \xhookrightarrow{\mathcal{R}(\phi)} \mathcal{R}(A)
$$

If $\text{Ker} \phi \subseteq \text{rad} A$, then $\mathcal{R}(\phi)$ is an isomorphism.

**Proof.** The first assertion is immediate from Proposition 1.46, since inflation $\phi^*$ clearly gives inclusions $\text{Irr} B \hookrightarrow \text{Irr} A$ and $\text{Irr}_{\text{fin}} B \hookrightarrow \text{Irr}_{\text{fin}} A$. If $\text{Ker} \phi \subseteq \text{rad} A$ then these inclusions are in fact bijections by (1.41). \(\square\)
Extension of the Base Field. Let $K/k$ be a field extension. For any $A \in \text{Alg}_k$ and any $V \in \text{Rep}_{\text{fin}} A$, consider algebra $K \otimes A \in \text{Alg}_K$ and the representation $K \otimes V \in \text{Rep}_{\text{fin}} (K \otimes A)$ as in (1.25). By exactness of the scalar extension functor $K \otimes \cdot$, this process leads to a well-defined group homomorphism
\begin{equation}
K \otimes \cdot : \mathcal{R}(A) \to \mathcal{R}(K \otimes A)
\end{equation}
(1.57)

$[V] \mapsto [K \otimes V]$

Lemma 1.48. The scalar extension map (1.57) is injective.

Proof. In view of Proposition 1.46, it suffices to show that, for $S \neq T \in \text{Irr}_{\text{fin}} A$, the representations $K \otimes S, K \otimes T \in \text{Rep}_{\text{fin}} (K \otimes A)$ have no common composition factor. To prove this, we may replace $A$ by $A/\text{Ker} S \cap \text{Ker} T$, thereby reducing to the case where the algebra $A$ is semisimple. The central primitive idempotent $e(S) \in A$ acts as the identity on $S$ and as 0 on $T$; see (1.47). Viewed as an element of $K \otimes A$, we have the same actions of $e(S)$ on $K \otimes S$ and on $K \otimes T$, whence these representations cannot have a common composition factor. □

The Character Map. Since characters are additive on short exact sequences in $\text{Rep}_{\text{fin}} A$ by Lemma 1.41, they give rise to a well-defined group homomorphism
\begin{equation}
\chi : \mathcal{R}(A) \to C(A) \leftarrow A^\circ_{\text{trace}}
\end{equation}
(1.58)

$[V] \mapsto \chi_V$

with $C(A) \subseteq A^\circ_{\text{trace}}$ as in (1.54). The character map $\chi$ is natural in $A$, that is, for any morphism $\phi : A \to B$ in $\text{Alg}_k$, the following diagram clearly commutes:

\begin{equation}
\begin{array}{ccc}
\mathcal{R}(B) & \xrightarrow{\chi} & B^\circ_{\text{trace}} \\
\mathcal{R}(\phi) \downarrow & & \downarrow \phi^\circ_{\text{trace}} \\
\mathcal{R}(A) & \xrightarrow{\chi} & A^\circ_{\text{trace}}
\end{array}
\end{equation}
(1.59)

Since $C(A)$ is a $k$-vector space, $\chi$ lifts uniquely to a $k$-linear map
\[ \chi_k : \mathcal{R}_k(A) \overset{\text{def}}{=} \mathcal{R}(A) \otimes \mathbb{Z} k \to C(A). \]

Proposition 1.49. The map $\chi_k$ is injective. If $k$ is a splitting field for $A$, then $\chi_k$ is an isomorphism.

Proof. First assume that $k$ is a splitting field for $A$. By Proposition 1.46, the classes $[S] \otimes 1$ with $S \in \text{Irr}_{\text{fin}} A$ form a $k$-basis of $\mathcal{R}_k(A)$, and by Theorem 1.44,
the irreducible characters $\chi_S$ form a $\kappa$-basis of $C(A)$. Since $\chi_\kappa([S] \otimes 1) = \chi_S$, the proposition follows in this case.

If $\kappa$ is arbitrary, fix an algebraic closure $\overline{\kappa}$ and consider the algebra $\overline{A} = \overline{\kappa} \otimes A$. Every trace form $A \to \kappa$ extends uniquely to a trace form $\overline{A} \to \overline{\kappa}$, giving a map $A^*_\text{trace} \to \overline{A}^*_\text{trace}$ (which is in fact an embedding). The following diagram is evidently commutative:

\[
\begin{array}{ccc}
\mathcal{R}_\kappa(A) & \xrightarrow{\chi_\kappa} & A^*_\text{trace} \\
\mathbb{K} \otimes \cdot & \downarrow & \downarrow \\
\mathcal{R}_\overline{\kappa}(\overline{A}) & \xleftarrow{\chi_{\overline{\kappa}}} & \overline{A}^*_\text{trace}
\end{array}
\]

(1.60)

Here, $\mathbb{K} \otimes \cdot$ and $\chi_{\overline{\kappa}}$ are injective by Lemma 1.48 and the first paragraph of this proof, respectively, whence $\chi_\kappa$ must be injective as well. $\square$

Positive Structure and Dimension Augmentation. The following subset is called the positive cone of $\mathcal{R}(A)$:

$$\mathcal{R}(A)_+ \overset{\text{def}}{=} \{ [V] \in \mathcal{R}(A) \mid V \in \text{Rep}_{\text{fin}} A \}.$$ 

This is a submonoid of the group $\mathcal{R}(A)$, because $0 = [0] \in \mathcal{R}(A)_+$ and $[V] + [V'] = [V \oplus V'] \in \mathcal{R}(A)_+$ for $V, V'$ in $\text{Rep}_{\text{fin}} A$. Under the isomorphism $\mathcal{R}(A) \cong \mathbb{Z}^\oplus \text{Irr}_{\text{fin}} A$ (Proposition 1.46), $\mathcal{R}(A)_+$ corresponds to $\mathbb{Z}^\oplus_{\text{Irr}_{\text{fin}} A}$. Thus, every element of $\mathcal{R}(A)$ is a difference of two elements of $\mathcal{R}(A)_+$ and $x = 0$ is the only element of $\mathcal{R}(A)_+$ such that $-x \in \mathcal{R}(A)_+$. This also follows from the fact that $\mathcal{R}(A)$ is equipped with a group homomorphism, called the dimension augmentation,

$$\dim: \mathcal{R}(A) \longrightarrow (\mathbb{Z}, +)$$

$$\begin{array}{ccc}
[V] & \downarrow \downarrow & \dim_\kappa V \\
\psi & \psi & \\
\end{array}$$

and $\dim x > 0$ for $0 \neq x \in \mathcal{R}(A)_+$. Defining

$$x \leq y \overset{\text{def}}{\iff} y - x \in \mathcal{R}(A)_+$$

we obtain a translation invariant partial order on the group $\mathcal{R}(A)$.

**Exercises for Section 1.5**

Unless otherwise specified, we consider an arbitrary $A \in \text{Alg}_\kappa$ below.

1.5.1 (Idempotents). Let $e = e^2 \in A$ be an idempotent and let $V \in \text{Rep} A$. Show:

(a) $U \cap e.V = e.U$ holds for every subrepresentation $U \subseteq V$.

(b) If $V$ is finite dimensional, then $\chi_V(e) = (\dim_\kappa e.V)1_\kappa$.
1.5.2 (Matrices). (a) Let $V \in \text{Rep}_{\text{fin}} A$. Viewing $V^{\oplus n}$ as a representation of $\text{Mat}_n(A)$ as in Lemma 1.4(b), show that $\chi_{V^{\oplus n}}(a) = \sum_i \chi_V(a_{i,i})$ for $a = (a_{i,j}) \in \text{Mat}_n(A)$.

(b) Let trace: $\text{Mat}_n(k) \to k$ be the ordinary matrix trace. Show that all characters of $\text{Mat}_n(k)$ are of the form $k \in \mathbb{Z}_+$ and that $\chi_{\text{reg}} = n \text{ trace}$.

(c) Show that $\text{Mat}_n(A)/[\text{Mat}_n(A), \text{Mat}_n(A)] \cong A/[A, A]$ in $\text{ Vect}_k$ via the map $\text{Mat}_n(A) \to A/[A, A]$, $(a_{i,j}) \mapsto \sum_i \text{Tr}(a_{i,i})$, where $\text{Tr}$ is the universal trace (1.53).

1.5.3 (Regular character and field extensions). Assume that $\dim_k A < \infty$ and let $K$ a field with $k \subseteq K \subseteq \mathcal{E}(A)$. Thus, we may also view $A \in \text{ Alg}_K$. Let $T_{A/k}$, $T_{A/K}$ and $T_{K/k}$ denote the regular characters of $A \in \text{ Alg}_k$, $A \in \text{ Alg}_K$ and $K \in \text{ Alg}_k$, respectively. Show that $T_{A/k} = T_{K/k} \circ T_{A/K}$.

1.5.4 (Finite-dimensional central simple algebras). This exercise uses definitions and results from Exercises 1.1.13 and 1.4.10. We assume $\dim_k A < \infty$.

(a) Show that $A$ is central simple iff $\overline{k} \otimes A \cong \text{Mat}_n(\overline{k})$, where $\overline{k}$ denotes an algebraic closure of $k$ and $n = \sqrt{\dim_k A}$. Consequently, finite-dimensional central simple algebras are separable.

(b) If $A$ is central simple, show that the regular character $\chi_{\text{reg}} = T_{A/k}$ vanishes if and only if $\text{char} k$ divides $\dim_k A$.

1.5.5 (Separable field extensions). Let $K/k$ be a finite field extension. Recall from field theory that $K/k$ is separable iff there are $[K : k]$ distinct $k$-algebra embeddings $\sigma_i : K \to \overline{k}$, where $\overline{k}$ is an algebraic closure of $k$ (e.g., [127, V.4 and V.6]). If $K/k$ is not separable, then $\text{char} k = p > 0$ and there is an intermediate field $k \subseteq F \subseteq K$ such that the minimal polynomial of every $a \in K$ over $F$ has the form $x^p - a$ for some $r \in \mathbb{N}$ and $a \in F$. Show that the following are equivalent:

(i) $K/k$ is a separable field extension;

(ii) $K \otimes k \cong [K:k]$ as $k$-algebras;

(iii) $K$ is a separable $k$-algebra in the sense of Exercise 1.4.10;

(iv) the regular character $T_{K/k}$ (Exercise 1.5.3) is nonzero.

1.5.6 (Separable algebras, again). (a) Using Exercise 1.4.10 and the preceding two exercises, show that $A$ is separable if and only if $A$ is finite dimensional, semisimple, and $\mathcal{E}(D(S))/k$ is a separable field extension for all $S \in \text{ Irr}_A$.

(b) Assume $\dim_k A < \infty$ and that $A \cong k \otimes_{k_0} A_0$ for some perfect subfield $k_0 \subseteq k$ and some $A_0 \in \text{ Alg}_{k_0}$. Show that $\text{rad} A \cong k \otimes_{k_0} \text{rad} A_0$.

1.5.7 (Independence of nonzero irreducible characters). For $S \in \text{ Irr}_{\text{fin}} A$, put $K(S) = \mathcal{E}(D(S))$. Show:

(a) $\chi_S \neq 0$ if and only if $\text{char} k \nmid \dim_{K(S)} D(S)$ and $K(S)/k$ is separable.

(b) The nonzero irreducible characters $\chi_S$ are $k$-linearly independent.
1.5.8 (Algebras that are defined over finite fields). Assume $\dim_k A < \infty$ and that $A \cong k \otimes_{k_0} A_0$ for some finite subfield $k_0 \subseteq k$ and some $A_0 \in \text{Alg}_{k_0}$.

(a) Let $S \in \text{Irr} A$ be absolutely irreducible, let $F(S)$ be the subfield of $k$ that is generated by $k_0$ and $\chi_S(A_0)$, and let $B(S) = F(S) \otimes_{k_0} A_0$. Show that $F(S)$ is finite and that $S \cong k \otimes_{F(S)} T$ for some $T \in \text{Irr} B(S)$. (Use Wedderburn’s Theorem on finite division algebras.)

(b) Assume that $k$ is a splitting field for $A$ and let $F$ be the (finite) subfield of $k$ that is generated by $k_0$ and the subfields $F(S)$ with $S \in \text{Irr} A$. Show that $F$ is a splitting field for the algebra $B = F \otimes_{k_0} A_0$. (Use Exercise 1.4.9.)

1.5.9 (Irreducible representations of tensor products). (a) Let $S \in \text{Rep}_{\text{fin}} A$ and $T \in \text{Rep}_{\text{fin}} B$ be absolutely irreducible. Use Burnside’s Theorem (§1.4.6) to show that the outer tensor product $S \boxtimes T$ is an absolutely irreducible representation of the algebra $A \otimes B$.

(b) Assuming $A$ and $B$ to be split semisimple, show that $A \otimes B$ is split semisimple as well and all its irreducible representation arise as in (a).
Further Topics on Algebras

This short chapter concludes our coverage of general algebras and their representations by introducing two topics: projective modules and Frobenius algebras. Neither will be needed in Part III, but the special case of symmetric algebras will play a role in Part II and our treatment of Hopf algebras in Part IV will make significant use of Frobenius algebras and also, to a lesser degree, of projective modules. Therefore, at a first pass through this book, this chapter can be skipped and later referred to as the need arises.

To the reader wishing to learn more about the representation theory of algebras, we recommend the classic [51] by Curtis and Reiner along with its extensive updates [52], [53] and the monograph [9] by Auslander, Reiten, and Smalø, which focusses on artinian algebras.

2.1. Projectives

So far in this book, the focus has been on irreducible and completely reducible representations. For algebras that are not necessarily semisimple, another class of representations also plays an important role: the projective modules. We will also refer to them as the projectives of the algebra in question, for short; we shall however refrain from calling them “projective representations,” since this term has a different specific meaning in group representation theory, being reserved for group homomorphisms of the form $G \to \text{PGL}(V)$ for some $V \in \text{Vec}_k$.

In this section, with the exception of §2.1.4, $A$ denotes an arbitrary $k$-algebra.
2.1.1. Definition and Basic Properties

A module \( P \in \text{A}\text{Mod} \) is called \textit{projective} if \( P \) is isomorphic to a direct summand of some free module: \( P' \oplus Q = A_{\text{reg}}^I \) for suitable \( P', Q \in \text{A}\text{Mod} \) with \( P' \cong P \) and some set \( I \). Projective modules, like free modules, can be thought of as approximate “vector spaces over \( A \),” but projectives are a much more ample and stable class of modules than free modules.

**Proposition 2.1.** The following are equivalent for \( P \in \text{A}\text{Mod} \).

(i) \( P \) is projective.

(ii) Given an epimorphism \( f: M \rightarrow N \) and an arbitrary \( p: P \rightarrow N \) in \( \text{A}\text{Mod} \), there exists a “lift” \( \tilde{p}: P \rightarrow M \) in \( \text{A}\text{Mod} \) such that \( f \circ \tilde{p} = p \):

\[
\begin{array}{ccc}
P & \xrightarrow{\exists \tilde{p}} & M \\
\downarrow{p} & & \downarrow{f} \\
N & & \\
\end{array}
\]

(iii) Every epimorphism \( f: M \rightarrow P \) in \( \text{A}\text{Mod} \) splits: there exists a homomorphism \( s: P \rightarrow M \) in \( \text{A}\text{Mod} \) such that \( f \circ s = \text{Id}_P \).

**Proof.** (i)\(\Rightarrow\)(ii): Say \( P' \oplus Q = F \) with \( F = A_{\text{reg}}^I \) and \( P' \cong P \). Identifying \( P \) with \( P' \), as we may, the embedding \( \mu: P \hookrightarrow F \) and the projection \( \pi: F \twoheadrightarrow P \) along \( Q \) satisfy

\[ \pi \circ \mu = \text{Id}_P. \]

Consider the module map

\[ q = p \circ \pi : F \rightarrow N. \]

In order to construct a lift for \( q \), fix an \( A \)-basis \( (f_i)_{i \in I} \) for \( F \). Since the map \( f \) is surjective, we may choose elements \( m_i \in M \) such that \( f(m_i) = q(f_i) \) for all \( i \in I \). Now define the desired lift \( \overline{q}: F \rightarrow M \) by \( \overline{q}(f_i) = m_i \); this determines \( \overline{q} \) unambiguously on \( F \) and we have

\[ f \circ \overline{q} = q, \]

as one checks by evaluating both functions on the basis \( (f_i)_{i \in I} \). Putting \( \overline{p} = \overline{q} \circ \mu: P \rightarrow M \), we obtain the desired equality,

\[ f \circ \overline{p} = f \circ \overline{q} \circ \mu = q \circ \mu = p \circ \pi \circ \mu = p \circ \text{Id}_P = p. \]

(ii)\(\Rightarrow\)(iii): Taking \( N = P \) and \( p = \text{Id}_P \), the lift \( \tilde{p}: P \rightarrow M \) from (ii) will serve as the desired splitting map \( s \).

(iii)\(\Rightarrow\)(i): As we have observed before (§1.4.4), any generating family \( (x_i)_{i \in I} \) of \( P \) gives rise to an epimorphism \( f: F = A_{\text{reg}}^I \rightarrow P, (a_i) \mapsto \sum_i a_i x_i \). If \( s \) is the splitting provided by (iii), then \( F = P' \oplus Q \) with \( P' = \text{Im} s \cong P \) and \( Q = \text{Ker} f \) (Exercise 1.1.2), proving (i). \( \square \)
2.1. Projectives

Description in Terms of Matrices. Any finitely generated projective $P \in _A\text{Mod}$ can be realized by means of an idempotent matrix over $A$. Indeed, there are $A$-module maps $P \xrightarrow{\pi} A_{\text{reg}}^\otimes_n$ for some $n$ with $\pi \circ \mu = \text{Id}_P$. Then $\mu \circ \pi \in \text{End}_A(A_{\text{reg}}^\otimes) \cong \text{Mat}_n(A)^{\text{op}}$ (Lemma 1.5) gives an idempotent matrix $e = e^2 \in \text{Mat}_n(A)$ with
\[
P \cong A_{\text{reg}}^\otimes e.
\]

Functorial Aspects. By §B.2.1, each $M \in _A\text{Mod}$ gives rise to a functor
\[
\text{Hom}_A(M, \cdot) : _A\text{Mod} \rightarrow \text{Vect}_k.
\]
This functor is left exact: if $0 \rightarrow X \xrightarrow{f} Y \xrightarrow{g} Z \rightarrow 0$ is a short exact sequence in $_A\text{Mod}$, then $0 \rightarrow \text{Hom}_A(M, X) \xrightarrow{f_*} \text{Hom}_A(M, Y) \xrightarrow{g_*} \text{Hom}_A(M, Z)$ is exact in $\text{Vect}_k$. However, $\text{Hom}_A(M, \cdot)$ is generally not exact, because $g_*$ need not be epi. Characterization (ii) in Proposition 2.1 can be reformulated as follows:
\[
M \in _A\text{Mod} \text{ is projective if and only if } \text{Hom}_A(M, \cdot) \text{ is exact.}
\]
In lieu of $\text{Hom}_A(M, \cdot)$, we can of course equally well consider the (contravariant) functor $\text{Hom}_A(\cdot, M) : _A\text{Mod} \rightarrow \text{Vect}_k$. The $A$-modules $M$ for which the latter functor is exact are called injective; see Exercise 2.1.1.

Dual Bases. Consider the functor $\cdot^\vee = \text{Hom}_A(\cdot, A_{\text{reg}}) : _A\text{Mod} \rightarrow \text{Vect}_k$, and, for any $M \in _A\text{Mod}$, let
\[
\langle \cdot, \cdot \rangle : M^\vee \times M \rightarrow A
\]
denote the evaluation pairing: $\langle f, m \rangle = f(m)$. A family $(x_i, x^i)_{i \in I}$, with each $(x_i, x^i) \in M \times M^\vee$, is said to form a pair of dual bases for $M$ if, for each $x \in M$,
\begin{enumerate}
  \item $\langle x^i, x \rangle = 0$ for almost all $i \in I$, and
  \item $x = \sum_{i \in I} \langle x^i, x \rangle x_i$.
\end{enumerate}
We equip $M^\vee$ with a right $A$-module structure by defining $\langle fa, m \rangle = \langle f, m \rangle a$ for $f \in M^\vee$, $a \in A$ and $m \in M$. Then we have the following canonical homomorphism in $\text{Vect}_k$:
\[
\begin{aligned}
M^\vee \otimes_A M & \longrightarrow \text{End}_A(M) \\
\text{ev} & \\
\psi & \\
f \otimes m & \longmapsto \langle x \mapsto \langle f, x \rangle m \rangle
\end{aligned}
\]
For $A = \mathbb{k}$, part (b) of the following lemma reduces to the standard isomorphism $\text{End}_k(V) \cong V \otimes V^* \cong V^* \otimes V$ for a finite-dimensional $V \in \text{Vect}_k$; see (B.19).

Lemma 2.2. Let $M \in _A\text{Mod}$.
\begin{enumerate}
  \item A pair of dual bases $(x_i, x^i)_{i \in I}$ for $M$ exists if and only if $M$ is projective. In this case, the family $(x_i)_{i \in I}$ generates $M$, and any generating family of $M$ can be chosen for $(x_i)_{i \in I}$.
\end{enumerate}
The map (2.3) is an isomorphism if and only if \( M \) is finitely generated projective.

**Proof.** (a) For given generators \((x_i)_{i \in I}\) of \( M \), consider the epimorphism \( A_\text{reg}^{\oplus I} \to M \), \((a_i)_i \mapsto \sum_i a_i x_i\). If \( M \) is projective, then we may fix a splitting \( s: M \to A_\text{reg}^{\oplus I} \). Now let \( \pi_i: A_\text{reg} \to A \), \((a_i)_i \mapsto a_i\), be the projection onto the \( i \)th component and define \( x^i = \pi_i \circ s \in \text{Hom}_A(M, A_\text{reg}) \) to obtain the desired pair of dual bases. Conversely, if \((x_i, x^i)_{i \in I}\) are dual bases for \( M \), then condition (ii) implies that \((x_i)_{i \in I}\) generates \( M \) and the map \( M \to A_\text{reg}^{\oplus I}, x \mapsto ((x^i, x)_i) \), splits the epimorphism \( A_\text{reg}^{\oplus I} \to M \), \((a_i)_i \mapsto \sum_i a_i x_i\). Therefore, \( M \) is isomorphic to a direct summand of \( A_\text{reg}^{\oplus I} \) and hence \( M \) is projective.

(b) Let \( \mu = \mu_M \) denote the map (2.3) and note that \( \phi \circ \mu(f \otimes m) = \mu(f \otimes \phi(m)) \) and \( \mu(f \otimes m) \circ \phi = \mu((f \circ \phi) \otimes m) \) for \( \phi \in \text{End}_A(M) \). Therefore, \( \text{Im} \mu \) is an ideal of \( \text{End}_A(M) \), and hence surjectivity of \( \mu \) is equivalent to the condition \( \text{Id}_M \in \text{Im} \mu \). Furthermore, \( \mu(\sum_{i=1}^n f_i \otimes m_i) = \text{Id}_M \) says exactly that \((f_i)_{i}, (m_i)_{i}\) are dual bases for \( M \). Hence we know by (a) that \( \text{Id}_M \in \text{Im} \mu \) if and only if \( M \) is finitely generated projective. Consequently, \( \mu \) is surjective if and only if \( M \) is finitely generated projective. It remains to show that \( \mu \) is also injective in this case. Fixing \( A \)-module maps \( M \xrightarrow{\mu_F} F = A^{\oplus n} \) with \( \pi \circ \mu = \text{Id}_M \), we obtain the commutative diagram

\[
\begin{array}{ccc}
M^\vee \otimes_A M & \xrightarrow{\mu_M} & \text{End}_A(M) \\
\pi^\vee \otimes \mu \downarrow & & \downarrow \mu \circ \cdot \pi \\
F^\vee \otimes_A F & \xrightarrow{\mu_F} & \text{End}_A(F)
\end{array}
\]

The vertical maps are injective, because they have left inverses \( \pi^\vee \otimes \pi \) and \( \pi \circ \cdot \circ \mu \). Let \( x^i \in F \) be the \( i \)th coordinate map and let \( x_i \in F \) be the standard \( i \)th basis element. Then \( F^\vee = \bigoplus_i x^i A \), \( F = \bigoplus_i Ax_i \) and

\[
\mu_F(\sum_{i,j} x^i a_{ij} \otimes x_j) = ((a_i)_i \mapsto \sum_{i,j} a_{ij})
\]

Therefore, \( \mu_F \) is a bijection by Lemma 1.5(b), and hence \( \mu_M \) is injective. \( \square \)

**Categories of Projectives.** We shall consider the following (full) subcategories of \( \textit{Mod}_A \): the categories consisting of all projectives, the finitely generated projectives, and the finite-dimensional projectives of \( A \); they will respectively be denoted by

\[
\text{Proj}_A, \text{proj}_A \text{ and Proj}_{\text{fin}} A.
\]

Our primary concern will be with the latter two. Induction along an algebra homomorphism \( \alpha: A \to B \) gives a functor

\[
\alpha_* = \text{Ind}_{B_A}^{B}: \text{Proj}_A \to \text{Proj}_B
\]
Indeed, for any $P \in \mathcal{A}^\text{proj}$, the functor $\text{Hom}_A(P, \text{Res}^B_A) : B\text{Mod} \to \text{Vect}_k$ is exact, being the composite of exact functors by (2.2), and $\text{Hom}_A(P, \text{Res}^B_A) \cong \text{Hom}_B(\text{Ind}^B_A P, \cdot)$ by Proposition 1.9. Thus, $\text{Hom}_B(\text{Ind}^B_A P, \cdot)$ is exact, showing that $\text{Ind}^B_A P$ is projective. Alternatively, $P$ is isomorphic to a direct summand of $A_{\text{reg}}$ for some set $I$, and so $\text{Ind}^B_A P = B \otimes_A P$ is isomorphic to a direct summand of $B_{\text{reg}} = \text{Ind}^B_A A_{\text{reg}}$ because $\text{Ind}^B_A$ commutes with direct sums. If $P$ is finitely generated, then $I$ may be chosen finite. Thus, induction restricts to a functor $\text{Ind}^B_A : \mathcal{A}^\text{proj} \to \mathcal{B}^\text{proj}$. Using the functor $\text{Mat}_n : \text{Alg}_k \to \text{Alg}_k$ and (2.1), we obtain

$$\text{Ind}^B_A P \cong B_{\text{reg}} \text{Mat}_n(\alpha)(e).$$

### 2.1.2. Hattori-Stallings Traces

For $P \in \mathcal{A}^\text{proj}$, the Dual Basis Lemma (Lemma 2.2) allows us to consider the following map\(^1\):

\[
\begin{align*}
\text{Tr}_P & : \text{End}_A(P) \cong P^\vee \otimes_A P \overset{\cong}{\longrightarrow} \text{Tr} A = A/[A, A] \\
& \overset{\psi}{\Rightarrow} \overset{\psi}{\Rightarrow} \\
& f \otimes p \overset{\psi}{\Rightarrow} (f, p) + [A, A]
\end{align*}
\]

(2.4)

Observe that this map is $k$-linear in both $f$ and $p$ and, for any $a \in A$,

$$\langle fa, p \rangle = \langle f, p \rangle a \equiv a\langle f, p \rangle = (fa, ap) \mod [A, A].$$

Therefore, $\text{Tr}_P$ is a well-defined $k$-linear map, which generalizes the standard trace map (B.23) for $A = k$. Extending the notion of dimension over $k$, the rank of $P$ is defined by

\[
\begin{align*}
\text{rank } P & \overset{\text{def}}{=} \text{Tr}_P(\text{Id}_P) \in \text{Tr} A \\
& \overset{\text{def}}{=} \text{Tr}_P(\text{Id}_P) \in \text{Tr} A
\end{align*}
\]

(2.5)

Using dual bases $(x_i, x^i)^n_{i=1}$ for $P$, the rank of $P$ may be expressed as follows:

\[
\text{rank } P = \sum_i \langle x^i, x_i \rangle + [A, A].
\]

(2.6)

Alternatively, writing $P \cong A_{\text{reg}}^{\oplus n}$ for some idempotent matrix $e = e^2 \in \text{Mat}_n(A)$ as in (2.1), one obtains $\text{rank } P = \sum_i e_{ii} + [A, A]$ (Exercise 2.1.5). In particular, $\text{rank } A_{\text{reg}}^{\oplus n} = n + [A, A]$ and so (2.5) extends the standard definition of the rank of a finitely generated free module over a commutative ring (B.9).

**Lemma 2.3.** Let $A$ be a $k$-algebra and let $P, P' \in \mathcal{A}^\text{proj}$.

(a) $\text{Tr}_P(\phi \circ \psi) = \text{Tr}_P(\psi \circ \phi)$ for all $\phi, \psi \in \text{End}_A(P)$.

(b) If $\phi \in \text{End}_A(P)$, $\phi' \in \text{End}_A(P')$, then $\text{Tr}_{P \oplus P'}(\phi \oplus \phi') = \text{Tr}_P(\phi) + \text{Tr}_{P'}(\phi')$.

In particular, $\text{rank}(P \oplus P') = \text{rank } P + \text{rank } P'$.

\(^1\)This map was introduced independently by Hattori [98] and by Stallings [190].
we obtain a well-defined group homomorphism $$\alpha$$ along the exact same lines as the construction of 2.1.3. The Grothendieck Groups

In terms of matrices, this can also be seen as follows. Write

$$\Phi = \text{Tr}(2.7)$$

and the functor $$\text{Func}$$.

Indeed, viewing the isomorphism $$\text{End}_{A}(P) \cong P^\vee \otimes A P$$ as an identification, it suffices to establish the trace property $$\text{Tr}_P(\phi \otimes \psi) = \text{Tr}_P(\psi \circ \phi)$$ for $$\phi = f \otimes p$$ and $$\psi = f' \otimes p'$$ with $$f, f' \in P^\vee$$ and $$p, p' \in P$$. Note that composition in $$\text{End}_{A}(P)$$ takes the form $$\phi \circ \psi = (f \otimes p) \circ (f' \otimes p') = \langle f, p' \rangle f' \otimes p$$. Therefore,

$$\text{Tr}_P(\phi \circ \psi) = \langle f, p' \rangle \langle f', p \rangle + [A, A] = \langle f', p \rangle \langle f, p' \rangle + [A, A] = \text{Tr}_P(\psi \circ \phi).$$

(b) Put $$Q = P \oplus P'$$ and $$\Phi = \phi \oplus 0_{P'}$$, $$\Phi' = 0_P \oplus \phi' \in \text{End}_{A}(Q)$$. Then $$\phi \oplus \phi' = \Phi + \Phi'$$. It is easy to see that $$\text{Tr}_Q(\Phi) = \text{Tr}_P(\phi)$$ and likewise for $$\Phi'$$.

Indeed, viewing $$\text{End}_{A}(P) \cong P^\vee \otimes A P$$ as a direct summand of $$\text{End}_{A}(Q) \cong Q^\vee \otimes A Q$$ in the canonical way, we have $$\Phi = \phi$$. Thus, the desired formula $$\text{Tr}_{P \oplus P'}(\phi \oplus \phi') = \text{Tr}_P(\phi) + \text{Tr}_{P'}(\phi')$$ follows by linearity of $$\text{Tr}_Q$$. Since $$\text{Id}_Q = \text{Id}_P \oplus \text{Id}_{P'}$$, we also obtain the rank formula $$\text{rank } Q = \text{rank } P + \text{rank } P'$$.

\[\Phi = \text{Tr}(2.7)\]

Functionality. We briefly address the issue of changing the algebra. So let $$\alpha : A \to B$$ be a map in $$\text{Alg}_k$$. Then we have the $$k$$-linear map $$\text{Tr} \alpha : \text{Tr } A \to \text{Tr } B$$ (§1.5.2) and the functor $$\alpha_* = \text{Ind}_{A}B : \text{Aproj} \to \text{Bproj}$$ from the previous paragraph. For any $$P \in \text{Aproj}$$ and any $$\phi \in \text{End}_{A}(P)$$, we have the formula

\[\text{(2.7)} \quad \text{Tr}_{\alpha_* P}(\alpha_* \phi) = (\text{Tr } \alpha)(\text{Tr } P(\phi)).\]

Indeed, with $$\phi = f \otimes p$$ as above, we have $$\alpha_* \phi = (\text{Id}_B \otimes f) \otimes (1_B \otimes p)$$ and $$\langle (\text{Id}_B \otimes f, 1_B \otimes p) \rangle = \alpha(\langle f, p \rangle)$$, proving (2.7). With $$\phi = \text{Id}_P$$, (2.7) gives

\[\text{(2.8)} \quad \text{rank}(\alpha_* P) = (\text{Tr } \alpha)(\text{rank } P).\]

In terms of matrices, this can also be seen as follows. Write $$P \cong A_{\text{reg}}^e$$ as in (2.1). Then $$\text{rank } P = \sum_i e_{ii} + [A, A]$$ and $$\alpha_* P \cong B_{\text{reg}}^e \text{Mat}_n(\alpha)(e)$$, and so $$\text{rank}(\alpha_* P) = \sum_i \alpha(e_{ii}) + [B, B] = (\text{Tr } \alpha)(\text{rank } P)$$.

2.1.3. The Grothendieck Groups $$K_0(A)$$ and $$\mathcal{P}(A)$$

Finitely Generated Projectives. Using $$\text{Aproj}$$ in place of $$\text{Rep}_{\text{fin }}A$$, one can construct an abelian group along the exact same lines as the construction of $$\mathcal{P}(A)$$ in §1.5.5. The group in question has generators $$[P]$$ for $$P \in \text{Aproj}$$ and relations $$[Q] = [P] + [R]$$ for each short exact sequence $$0 \to P \to Q \to R \to 0$$ in $$\text{Aproj}$$.

Since all these sequences split by Proposition 2.1, this means that we have a relation $$[Q] = [P] + [R]$$ whenever $$Q \cong P \oplus R$$ in $$\text{Aproj}$$. The resulting abelian group is commonly denoted by $$K_0(A)$$.

Given an abelian group $$(\mathcal{G}, +)$$ and a rule assigning to each $$P \in \text{Aproj}$$ a value $$f(P) \in \mathcal{G}$$ in such a way that $$f(Q) = f(P) + f(R)$$ whenever $$Q \cong P \oplus R$$ in $$\text{Aproj}$$, we obtain a well-defined group homomorphism $$K_0(A) \to \mathcal{G}$$, $$[P] \mapsto f(P)$$.

The construction of $$K_0$$ is functorial: for any map $$\alpha : A \to B$$ in $$\text{Alg}_k$$, the induction functor $$\alpha_* = \text{Ind}_{A}B : \text{Aproj} \to \text{Bproj}$$ commutes with direct sums, and
hence it yields a well-defined homomorphism of abelian groups,

\[ K_0(\alpha) : K_0(A) \longrightarrow K_0(B) \]

\[ [P] \longmapsto [\text{Ind}_A^B P] \]

In this way, we obtain a functor \( K_0 : \text{Alg}_k \rightarrow \text{AbGroups} \).

We remark that, for \( P, Q \in A\text{proj} \), the equality \([P] = [Q]\) in \( K_0(A)\) means that \( P \) and \( Q \) are **stably isomorphic** in the sense that \( P \oplus A_{\text{reg}}^r \cong Q \oplus A_{\text{reg}}^r \) for some \( r \in \mathbb{Z}_+ \) (Exercise 2.1.9). We could of course conceivably perform an analogous construction using the category of \( A\text{Proj} \) of all projectives of \( A \) in place of \( A\text{proj} \). However, the resulting group would be trivial in this case (Exercise 2.1.6).

**Finite-Dimensional Projectives.** For the purposes of representation theory, we shall often be concerned with the full subcategory \( \text{Proj}_{\text{fin}} A \) of \( A\text{proj} \) consisting of all finite-dimensional projectives of \( A \). The corresponding Grothendieck group, constructed from \( \text{Proj}_{\text{fin}} A \) exactly as we did for \( A\text{proj} \), will be denoted by

\[ \mathcal{P}(A) \]

The following proposition sorts out the group theoretical structure of \( \mathcal{P}(A) \) in a manner analogous to Proposition 1.46, where the same was done for \( \mathcal{A}(A) \). While the latter result was a consequence of the Jordan-Hölder Theorem (Theorem 1.18), the operative fact in Proposition 2.4 below is the Krull-Schmidt Theorem (§1.2.6).

**Proposition 2.4.**

(a) \( \mathcal{P}(A) \) is isomorphic to the free abelian group with basis the set of isomorphism classes of finite-dimensional indecomposable projectives of \( A \).

(b) For \( P, Q \in \text{Proj}_{\text{fin}} A \), we have \([P] = [Q]\) in \( \mathcal{P}(A)\) if and only if \( P \cong Q \).

**Proof.** By the Krull-Schmidt Theorem, any \( P \in \text{Proj}_{\text{fin}} A \) can be decomposed into a finite direct sum of indecomposable summands and this decomposition is unique up to the order the summands and their isomorphism type. Thus, letting \( I \) denote a full set of representatives for isomorphism classes of finite-dimensional indecomposable projectives of \( A \), we have

\[ P \cong \bigoplus_{I \in I} I^{n_I(P)} \]

for unique \( n_I(P) \in \mathbb{Z}_+ \), almost all of which are zero. Evidently, \( n_I(Q) = n_I(P) + n_I(R) \) if \( Q \cong P \oplus R \) in \( \text{Proj}_{\text{fin}} A \); so we obtain a well-defined group homomorphism

\[ \mathcal{P}(A) \longrightarrow \mathbb{Z}^{\oplus I} \]

\[ [P] \longmapsto (n_I(P))_I \]
The map sending the standard \( \mathbb{Z} \)-basis element \( e_I = (\delta_{I,J})_J \in \mathbb{Z}^\oplus J \) to \( [I] \in \mathcal{P}(A) \) is inverse to the above homomorphism, and so we have in fact constructed an isomorphism. This proves (a), and (b) is an immediate consequence as well. \( \Box \)

**The Cartan Homomorphism.** Since any \( P \in \text{Proj}_{\text{fin}} A \) is of course also an object of \( \text{A}_{\text{proj}} \) and of \( \text{Rep}_{\text{fin}} A \), the symbol \( [P] \) can be interpreted in \( \mathcal{P}(A) \) as well as in \( K_0(A) \) and in \( \mathcal{R}(A) \). In fact, it is clear from the definition of these groups that there are group homomorphisms

\[
\begin{align*}
  c : \mathcal{P}(A) & \longrightarrow \mathcal{R}(A) \\
  [P] & \longmapsto [P]
\end{align*}
\]

and an analogous homomorphism \( \mathcal{P}(A) \rightarrow K_0(A) \). The map (2.9) is particularly important; it is called the **Cartan homomorphism**. Despite the deceptively simple looking expression above, \( c \) need not be injective, whereas the homomorphism \( \mathcal{P}(A) \rightarrow K_0(A) \) is in fact always mono (Exercises 2.1.10, 2.1.13).

**A Pairing Between** \( K_0(A) \) **and** \( \mathcal{R}(A) \). For any finitely generated \( V \in \text{A}_{\text{Mod}} \) and any \( W \in \text{Rep}_{\text{fin}} A \), the \( \mathbb{K} \)-vector space \( \text{Hom}_A(V, W) \) is finite dimensional, because \( A_{\text{reg}} \otimes V \) for some \( n \) and so \( \text{Hom}_A(V, W) \hookrightarrow \text{Hom}_A(A_{\text{reg}} \otimes W) \cong W^\oplus n \). Thus, we may define

\[
(2.10) \quad \langle V, W \rangle \overset{\text{def}}{=} \dim_{\mathbb{K}} \text{Hom}_A(V, W)
\]

Now let \( P \in \text{A}_{\text{proj}} \). Then the functor \( \text{Hom}_A(P, \cdot) \) is exact by (2.2) and we obtain a group homomorphism \( \langle P, \cdot \rangle : \mathcal{R}(A) \rightarrow \mathbb{Z} \). For any \( V \in \text{Rep}_{\text{fin}} A \), the functor \( \text{Hom}_A(\cdot, V) \) does at least commute with finite direct sums. Thus, we also have a group homomorphism \( \langle \cdot, V \rangle : \mathcal{P}(A) \rightarrow \mathbb{Z} \). The value \( \langle P, V \rangle \) only depends on the classes \( [P] \in \mathcal{P}(A) \) and \( [V] \in \mathcal{R}(A) \), giving a bi-additive pairing

\[
\begin{align*}
  K_0(A) \times \mathcal{R}(A) & \longrightarrow \mathbb{Z} \\
  ([P], [V]) & \longmapsto \langle P, V \rangle
\end{align*}
\]

Under suitable hypotheses, this pairing gives a “duality” between \( K_0(A) \) and \( \mathcal{R}(A) \) that will play an important role later (e.g., §3.4.2).
2.1. Projectives

Hattori-Stallings Ranks and Characters. By Lemma 2.3(b), the Hattori-Stallings rank gives a well-defined group homomorphism

\[
\text{rank}: K_0(A) \longrightarrow \text{Tr } A
\]

(2.11)

\[
[P] \longmapsto \text{rank } P
\]

If \( \alpha: A \to B \) is a homomorphism of \( k \)-algebras, then the following diagram commutes by (2.8):

\[
\begin{array}{ccc}
K_0(A) & \xrightarrow{K_0(\alpha)} & K_0(B) \\
\downarrow \text{rank} & & \downarrow \text{rank} \\
\text{Tr } A & \xrightarrow{\text{Tr } \alpha} & \text{Tr } B
\end{array}
\]

(2.12)

The following proposition, due to Bass [10], further hints at the aforementioned duality between \( K_0(A) \) and \( \mathcal{R}(A) \) by relating the pairing (2.10) to the evaluation pairing between \( \text{Tr } A \) and \( A^{\star}_{\text{trace}} \equiv (\text{Tr } A)^{\star} \). Recall that \( \chi: \mathcal{R}(A) \to A^{\star}_{\text{trace}} \) is the character homomorphism (1.58).

Proposition 2.5. The following diagram commutes:

\[
\begin{array}{ccc}
K_0(A) \times \mathcal{R}(A) & \xrightarrow{(\ldots)} & \mathbb{Z} \\
\downarrow \text{rank} \times \chi & & \downarrow \text{can.} \\
\text{Tr } A \times A^{\star}_{\text{trace}} & \xrightarrow{\text{evaluation}} & \mathbb{K}
\end{array}
\]

Proof. The proposition states that, for \( P \in \mathcal{A}_{\text{proj}} \) and \( V \in \text{Rep}_{\text{fin}} A \),

\[
\langle \chi_V, \text{rank } P \rangle = \dim_k \text{Hom}_A(P, V) 1_k
\]

This is clear if \( P \equiv A^{\text{reg}}_n \), for, then \( \text{Hom}_A(P, V) \equiv V^{\otimes n} \) and \( \langle \chi_V, \text{rank } P \rangle = \langle \chi_V, n + [A, A] \rangle = n \dim_k V \). The general case elaborates on this observation. In detail, fix \( A \)-module maps \( P \xrightarrow{\mu} F = A^{\text{reg}}_n \) for some \( n \) with \( \pi \circ \mu = \text{Id}_P \). The functor \( \text{Hom}_A(\cdot, V) \) then yields \( k \)-linear maps

\[
\begin{array}{ccc}
\text{Hom}_A(P, V) & \xrightarrow{\pi^* = \cdot \circ \pi} & \text{Hom}_A(F, V) \equiv V^{\otimes n} \\
\mu^* = \cdot \circ \mu
\end{array}
\]

with \( \mu^* \circ \pi^* = \text{Id}_{\text{Hom}_A(P, V)} \). Thus, \( h := \pi^* \circ \mu^* \in \text{End}_k(\text{Hom}_A(F, V)) \) is an idempotent with \( \text{Im } h \equiv \text{Hom}_A(P, V) \). Therefore, by standard linear algebra (Exercise 1.5.1(b)),

\[
\dim_k \text{Hom}_A(P, V) 1_k = \text{trace } h.
\]
Let \((e_i, e^i)_{i=1}^{n}\) be dual bases for \(F = A_{\text{reg}}^n\). Then we obtain dual bases \((x_i, x^i)_{i=1}^{n}\) for \(P\) by putting \(x_i = \pi(e_i)\) and \(x^i = e^i \circ \mu\). Chasing the idempotent \(h\) through the isomorphism
\[
\text{End}_k(\text{Hom}_A(F, V)) \xrightarrow{\sim} \text{End}_k(V^\otimes n) \xrightarrow{\text{Lemma 1.4}} \text{Mat}_n(\text{End}_k(V))
\]
coming from \(\text{Hom}_A(F, V) \xrightarrow{\sim} V^\otimes n, f \mapsto (f(e_i))\), one sees that \(h\) corresponds to the matrix \((h_{i,j}) \in \text{Mat}_n(\text{End}_k(V))\) that is given by \(h_{i,j}(v) = \langle x^i, x_j \rangle_V v\) for \(v \in V\). Therefore,
\[
\text{trace } h = \sum_i \text{trace } h_{i,i} = \sum_i \text{trace}(\langle x^i, x_i \rangle_V) = \langle \chi_V, \sum_i \langle x^i, x_i \rangle \rangle = \langle \chi_V, \text{rank } P \rangle
\]
and the proof is complete. \(\square\)

2.1.4. Finite-Dimensional Algebras

We now turn our attention to the case where the algebra \(A\) is finite dimensional. Then the categories \(A\text{proj}\) and \(\text{Proj}_{\text{fin}} A\) coincide and so \(K_0(A) = \mathcal{P}(A)\). Our first goal will be to describe the indecomposable projectives of \(A\). This will result in more explicit descriptions of the group \(\mathcal{P}(A)\) and of the Cartan homomorphism \(c: \mathcal{P}(A) \to \mathcal{R}(A)\) than previously offered in Proposition 2.4 and in (2.9).

Lifting Idempotents

We start with a purely ring theoretic lemma, for which the algebra \(A\) need not be finite dimensional. An ideal \(I\) of \(A\) is called \(\text{nil}\) if all elements of \(I\) are nilpotent. A family \((e_i)_{i \in I}\) of idempotents of \(A\) is called \(\text{orthogonal}\) if \(e_ie_j = \delta_{i,j}e_i\) for \(i, j \in I\).

Lemma 2.6. Let \(I\) be a nil ideal of \(A\) and let \(f_1, \ldots, f_n\) be orthogonal idempotents of \(A/I\). Then there exist orthogonal idempotents \(e_i \in A\) such that \(e_i + I = f_i\).

Proof. First, consider the case \(n = 1\) and write \(f = f_1 \in A/I\). Let \(-: A \to A/I\) denote the canonical map and fix any \(a \in A\) such that \(\bar{a} = f\). Then the element \(b = 1 - a \in A\) satisfies \(ab = ba = a - a^2 \in I\), and hence \((ab)^m = 0\) for some \(m \in \mathbb{N}\). Therefore, by the Binomial Theorem, \(1 = (a + b)^{2m} = e + e'\) with \(e = \sum_{i=0}^{m} \binom{2m}{i} a^{2m-i} b^i\) and \(e' = \sum_{i=m+1}^{2m} \binom{2m}{i} a^{2m-i} b^i\). By our choice of \(m\), we have \(ee' = e'e = 0\) and so \(e = e(e + e') = e^2\) is an idempotent. Finally, \(e \equiv a^{2m} \equiv a \mod I\), whence \(\bar{e} = \bar{a} = f\) as desired. Note also that \(e\) is a polynomial in \(a\) with integer coefficients and zero constant term.

Now let \(n > 1\) and assume that we have already constructed \(e_1, \ldots, e_{n-1} \in A\) as in the lemma. Then \(x = \sum_{i=1}^{n-1} e_i\) is an idempotent of \(A\) such that \(e_i x = xe_i = e_i\) for \(1 \leq i \leq n - 1\). Fix any \(a \in A\) such that \(\bar{a} = f_n\) and put \(a' = (1 - x)a(1 - x) \in A\). Then \(xa' = a'x = 0\). Furthermore, since \(\bar{a} = \sum_{i=1}^{n-1} f_i\) and \(\bar{a} = f_n\) are orthogonal idempotents of \(A/I\), we have \(a' = f_n\). Now construct the idempotent \(e_n \in A\) with
\[ e_n = f_n \text{ from } a' \text{ as in the first paragraph. Since } e_n \text{ is a polynomial in } a' \text{ with integer coefficients and zero constant term, it follows that } xe_n = e_n x = 0. \] Therefore, \( e_i e_n = e_i xe_n = 0 \) and, similarly, \( e_n e_i = 0 \) for \( i \neq n \), completing the proof. \( \square \)

**Projective Covers**

Let us now assume that \( A \in \text{Alg}_k \) is finite dimensional. We have already repeatedly used the fact that, for any \( V \in \text{Rep} A \), there exists an epimorphism \( P \to V \) with \( P \) projective or even free. It turns out that now there is a "minimal" choice for such an epimorphism, which is essentially unique. To describe this choice, consider the completely reducible factor

\[
\text{head } V \overset{\text{def}}{=} V/(\text{rad } A).V
\]

This construction is functorial: \( \text{head } \cdot \overset{A \text{s.p.}}{\to} \text{head } \cdot \), where \( A \text{s.p.} = A/\text{rad } A \) is the semisimplification of \( A \).

**Theorem 2.7.** Let \( A \in \text{Alg}_k \) be a finite dimensional. Then, for any \( V \in \text{Rep} A \), there exists a \( P \in A\text{Proj} \) and an epimorphism \( \phi: P \to V \) satisfying the following equivalent conditions:

(i) \( \text{Ker } \phi \subseteq (\text{rad } A).P \).

(ii) \( \text{head } \phi: \text{head } P \to \text{head } V \).

(iii) Every epimorphism \( \phi': P' \to V \) with \( P' \in A\text{Proj} \) factors as \( \phi' = \phi \circ \pi \) for some epimorphism \( \pi: P' \to P \).

In particular, \( P \) is determined by \( V \) up to isomorphism.

**Proof.** We start by proving the existence of an epimorphism \( \phi \) satisfying (i).

First assume that \( V \) is irreducible. Then \( V \) is a direct summand of the regular representation of \( A^{\text{s.p.}} = A/\text{rad } A \), and hence \( V \cong A^{\text{s.p.}}.f \) for some idempotent \( f \in A^{\text{s.p.}} \). Since \( \text{rad } A \) is nil, even nilpotent, Lemma 2.6 guarantees the existence of an idempotent \( e \in A \) so that \( \bar{e} = f \) under the canonical map \( A \to A^{\text{s.p.}} \). Putting \( P = Ae \) and \( \phi = \bar{\cdot} \), we obtain a projective \( P \in A\text{Proj} \) and an epimorphism \( \phi: P \to V \) satisfying \( \text{Ker } \phi = Ae \cap \text{rad } A = (\text{rad } A)e = (\text{rad } A).P \) as required.

Next assume that \( V \) is completely reducible and write \( V \cong \bigoplus_{S \in \text{Irr } A} S^{\oplus m(S,V)} \) as in (1.44). For each \( S \), let \( \phi_S: P_S \to S \) be the epimorphism constructed in the previous paragraph. Then the following map satisfies the requirements of (i):

\[
\phi = \bigoplus_{S \in \text{Irr } A} \phi_S^{\oplus m(S,V)}: \quad P = \bigoplus_{S \in \text{Irr } A} P_S^{\oplus m(S,V)} \to \bigoplus_{S \in \text{Irr } A} S^{\oplus m(S,V)} \to V.
\]
For general \( V \), consider the epimorphism \( \phi: P \to \text{head} V \) constructed in the previous paragraph. Proposition 2.1 yields a morphism \( \tilde{\phi} \) as in the diagram

\[
\begin{array}{ccc}
P & \xrightarrow{\exists \tilde{\phi}} & \text{head} V \\
\downarrow \phi & & \downarrow \\
V & \xrightarrow{\text{can}} & \text{head} V
\end{array}
\]

Since \( \phi = \tilde{\phi} \circ \text{can} \) is surjective, \( \text{Im} \tilde{\phi} + (\text{rad} A).V = V \). Iterating this equality, we obtain \( \text{Im} \tilde{\phi} + (\text{rad} A)^i.V = V \) for all \( i \). Hence, \( \text{Im} \tilde{\phi} = V \), because \( \text{rad} A \) is nilpotent. Moreover, \( \text{Ker} \phi \subseteq \text{Ker} \tilde{\phi} \subseteq (\text{rad} A).P \). This completes the proof of the existence claim in the theorem.

In order to prove the equivalence of (i)–(iii), note that \( \phi: P \to V \) gives rise to an epimorphism \( \text{head} \phi: \text{head} P \to \text{head} V \) with

\[
\text{Ker}(\text{head} \phi) = \phi^{-1}(\text{rad} A).V)/(\text{rad} A).P = (\text{Ker} \phi + (\text{rad} A).P)/(\text{rad} A).P.
\]

Therefore, \( \text{head} \phi \) is an isomorphism if and only if \( \text{Ker} \phi \subseteq (\text{rad} A).P \), proving the equivalence of (i) and (ii). Now assume that \( \phi \) satisfies (i) and let \( \phi' \) be as in (iii). Then Proposition 2.1 yields the diagram

\[
\begin{array}{ccc}
P' & \xrightarrow{\exists \pi} & P \\
\downarrow \phi' & & \downarrow \phi \\
P & \xrightarrow{\phi} & V
\end{array}
\]

As above, it follows from surjectivity of \( \phi' \) that \( P = \text{Im} \pi + \text{Ker} \phi = \text{Im} \pi + (\text{rad} A).P \) and iteration of this equality gives \( P = \text{Im} \pi \). This shows that (i) implies (iii). For the converse, assume that \( \phi \) satisfies (iii) and pick some epimorphism \( \phi': P' \to V \) with \( P' \in \text{AProj} \) and \( \text{Ker} \phi' \subseteq (\text{rad} A).P' \). By (iii), there exists an epimorphism \( \pi: P' \to P \) with \( \phi' = \phi \circ \pi \). Therefore, \( \text{Ker} \phi = \pi(\text{Ker} \phi') \subseteq (\text{rad} A).\pi(P') = (\text{rad} A).P \) and so \( \phi \) satisfies (i). This establishes the equivalence of (i)–(iii).

Finally, for uniqueness, let \( \phi: P \to V \) and \( \phi': P' \to V \) both satisfy (i)–(iii). Then there are epimorphisms \( P' \xrightarrow{\pi} P' \) such that \( \phi \circ \pi' = \phi' \) and \( \phi' \circ \pi = \phi \). Consequently, \( \phi = \phi \circ \pi' \circ \pi \) and so \( \text{Ker} \pi \subseteq \text{Ker} \phi \subseteq (\text{rad} A).P \). On the other hand, \( P = Q \oplus \text{Ker} \pi \) for some \( Q \), because the epimorphism \( \pi \) splits. Therefore, \( \text{Ker} \pi = (\text{rad} A).\text{Ker} \pi \), which forces \( \text{Ker} \pi = 0 \) by nilpotency of \( \text{rad} A \). Hence \( \pi \) is an isomorphism and the proof of the theorem is complete. \( \square \)

The projective constructed in the theorem above for a given \( V \in \text{Rep} A \) is called the **projective cover** of \( V \); it will be denoted by \( PV \).
Thus, we have an epimorphism $PV \to V$ and $PV$ is minimal in the sense that $PV$ is isomorphic to a direct summand of every $P \in \mathcal{A}\text{Proj}$ such that $P \to V$ by (iii) of Theorem 2.7. Moreover, (ii) states that

$$(2.13) \quad \text{head } PV \cong \text{head } V.$$ 

Exercise 2.1.7 explores some further properties of the operator $P$.

**Principal Indecomposable Representations**

Since $A$ is assumed finite dimensional, the regular representation $A_{\text{reg}}$ decomposes into a finite direct sum of indecomposable representations. A full representative set of the isomorphism types of the summands occurring in this decomposition is called a set of **principal indecomposable representations** of $A$. All principal indecomposables evidently belong to $\mathcal{Proj}_{\text{fin}} A$ and they are unique up to isomorphism by the Krull-Schmidt Theorem (§1.2.6). The following proposition lists some further properties of the principal indecomposable representations. Recall that, for any $V \in \mathcal{Rep}_{\text{fin}} A$ and any $S \in \text{Irr } A$, the multiplicity of $S$ as a composition factor of $V$ is denoted by $\mu(S, V)$; see the Jordan-Hölder Theorem (Theorem 1.18).

**Proposition 2.8.** Let $A \in \text{Alg}_k$ be finite dimensional. Then:

(a) The principal indecomposable representations of $A$ are exactly the projective covers $PS$ with $S \in \text{Irr } A$; they are a full representative set of the isomorphism classes of all indecomposable projectives of $A$.

(b) $A_{\text{reg}} \cong \bigoplus_{S \in \text{Irr } A} (PS)^{\oplus \dim D(S) S}.$

(c) $(PS, V) = \mu(S, V) \dim_k D(S)$ for any $V \in \mathcal{Rep}_{\text{fin}} A$ and $S \in \text{Irr } A$.

**Proof.** Since $\text{head } PS \cong S$ by (2.13), the various $PS$ are pairwise non-isomorphic and they are all indecomposable. Now let $P \in \mathcal{A}\text{Proj}$ be an arbitrary indecomposable projective. Since $P$ is a submodule of $A_{\text{reg}}$ for some set $I$, there exists a nonzero homomorphism $P \to A_{\text{reg}}$, and hence there certainly exists an epimorphism $P \to S$ for some $S \in \text{Irr } A$. But then $PS$ is isomorphic to a direct summand of $P$ by Theorem 2.7(iii), and hence $P \cong PS$. Thus, the collection $PS$ with $S \in \text{Irr } A$ forms a full set of non-isomorphic indecomposable projectives for $A$. To see that this collection also coincides with the principal indecomposables, observe that $P_{\text{reg}}^{\oplus P} \cong A_{\text{reg}}$, because the canonical map $A \to A^{\oplus P}$ has kernel $\text{rad } A$. Since Wedderburn’s Structure Theorem gives the decomposition $A_{\text{reg}}^{\oplus P} \cong \bigoplus_{S \in \text{Irr } A} S^{\oplus \dim D(S) S}$, the isomorphism in (b) now follows by additivity of the operator $P \cdot$ on direct sums (Exercise 2.1.7). This proves (a) as well as (b).

For (c), note that the function $(PS, \cdot) = \dim_k \text{Hom}_A(PS, \cdot)$ is additive on short exact sequences in $\mathcal{Rep}_{\text{fin}} A$ by exactness of the functor $\text{Hom}_A(PS, \cdot)$, and so is the multiplicity $\mu(S, \cdot)$ by (1.32). Therefore, by considering a composition series of $V$, ...
one reduces (c) to the case where \( V \in \text{Irr } A \). But then \( \mu(S, V) = \delta_{S,V} \) and
\[
\text{Hom}_A(PS, V) \cong \text{Hom}_A(\text{head } PS, V) \cong \text{Hom}_A(S, V) = \delta_{S,V} D(S)
\]
by Schur's Lemma. The formula in (c) follows from this. \( \Box \)

As a special case of the multiplicity formula in Proposition 2.8(c), we note the so-called orthogonality relations:
\[
(2.14) \quad (PS, S') = \delta_{S,S'} \dim_k D(S) \quad (S, S' \in \text{Irr } A)
\]
The multiplicity formula in Proposition 2.8(c) and the orthogonality relations (2.14) have a particularly appealing form when the base field \( k \) is a splitting field for \( A \); for, then \( \dim_k D(S) = 1 \) for all \( S \in \text{Irr } A \).

**The Cartan Matrix**

In our current finite-dimensional setting, the Grothendieck groups \( \mathcal{P}(A) = K_0(A) \) and \( \mathcal{R}(A) \) are both free abelian of finite rank equal to the size of \( \text{Irr } A \). Indeed, the classes \([S] \in \mathcal{R}(A)\) with \( S \in \text{Irr } A \) provide a \( \mathbb{Z} \)-basis of \( \mathcal{R}(A) \) (Proposition 1.46) and the classes \([PS] \in \mathcal{P}(A)\) form a \( \mathbb{Z} \)-basis of \( \mathcal{P}(A) \) (Propositions 2.4(a) and 2.8(a)). In terms of these bases, the Cartan homomorphism \( c: \mathcal{P}(A) \to \mathcal{R}(A) \) in (2.9) has the following description:

\[
\mathcal{P}(A) \cong \mathbb{Z}^{\oplus \text{Irr } A} \xrightarrow{\psi} \mathcal{R}(A) \cong \mathbb{Z}^{\oplus \text{Irr } A} \xrightarrow{\psi} \sum_{S' \in \text{Irr } A} \mu(S', PS)[S']
\]

Thus, the Cartan homomorphism can be described by the following integer matrix:
\[
(2.15) \quad C = \left( \mu(S', PS) \right)_{S', S \in \text{Irr } A}.
\]
This matrix is called the **Cartan matrix** of \( A \). Note that all entries of \( C \) belong to \( \mathbb{Z}_+ \) and that the diagonal entries are strictly positive. If \( k \) is a splitting field for \( A \), then the Cartan matrix takes the following form by Proposition 2.8(c):
\[
(2.16) \quad C = \left( (PS', PS) \right)_{S', S \in \text{Irr } A}.
\]

**Characters of Projectives**

In this paragraph, we will show that the Hattori-Stallings rank of \( P \in \text{Proj}_{\text{fin }} A \) determines the character \( \chi_P \). The reader is reminded that the character map \( \chi: \mathcal{R}(A) \to A^*_\text{trace} = (\text{Tr } A)^* \) has image in the subspace \( C(A) \equiv (\text{Tr } A^a)^* \) of \( (\text{Tr } A)^* \); see (1.55). Let \( A \) denote right and left multiplication with \( a, b \in A \), respectively, and define a \( k \)-linear map \( t = t_A: A \to A^* \) by
\[
(2.17) \quad \langle t(a), b \rangle \overset{\text{def}}{=} \text{trace}(b_A \circ Aa)
\]
Note that if \( a \) or \( b \) belong to \([A, A]\), then \( b_A \circ Aa \in [\text{End}_k(A), \text{End}_k(A)]\) and so \( \text{trace}(b_A \circ Aa) = 0 \). Moreover, if \( a \) or \( b \) belong to \( \text{rad} A \), then the operator \( b_A \circ Aa \) is nilpotent, and so \( \text{trace}(b_A \circ Aa) = 0 \) again. Therefore, the map \( t \) can be refined as in the following commutative diagram, with \( t \) denoting all refinements:

\[
\begin{array}{ccc}
A & \xrightarrow{t} & A^* \\
\downarrow \text{can.} & & \downarrow \text{can.} \\
\text{Tr} A & \xrightarrow{t} & \text{A}_{\text{trace}}^* = (\text{Tr} A)^* \\
\downarrow \text{can.} & & \downarrow \text{can.} \\
\text{Tr} A^\text{s.p.} & \xrightarrow{t} & (\text{Tr} A^\text{s.p.})^* \cong C(A)
\end{array}
\] (2.18)

The following proposition is due to Bass [10].

**Proposition 2.9.** Let \( A \in \text{Alg}_k \) be finite dimensional. The following diagram commutes:

\[
\begin{array}{ccc}
\mathcal{P}(A) & \xrightarrow{c} & \mathcal{R}(A) \\
\downarrow \text{rank} & & \downarrow \chi \\
\text{Tr} A & \xrightarrow{t} & (\text{Tr} A)^*
\end{array}
\]

**Proof.** We need to check the equality \( \langle \chi_P, a \rangle = \text{trace}(a_A \circ A \text{rank} P) \) for \( P \in \text{Proj}_\text{fin} A \) and \( a \in A \). Fix dual bases \((b_i, b^j)_i\) for \( A|_{\text{Vec}_k} \) and let \((x_j, x^j)_j\) be dual bases for \( P|_{\text{Vec}_k} \); this follows from the calculation, for \( p \in P \),

\[
p = \sum_{i,j} \langle x^j, p \rangle x_j = \sum_{i,j} \langle b^j, \langle x^j, p \rangle \rangle b_i x_j = \sum_{i,j} \langle b^j \circ x^j, p \rangle b_i x_j.
\]

Thus, \( \text{Id}_P \) corresponds to \( \sum_{i,j} b_i \cdot x_j \otimes b^j \circ x^j \in P \otimes P^* \) under the standard isomorphism \( \text{End}_k(P) \cong P \otimes P^* \) and, for any \( a \in A \), the endomorphism \( a_P \in \text{End}_k(P) \) corresponds to \( \sum_{i,j} ab_i \cdot x_j \otimes b^j \circ x^j \). Therefore,

\[
\langle \chi_P, a \rangle = \text{trace} a_P = \sum_{i,j} \langle b^j \circ x^j, ab_i \cdot x_j \rangle = \sum_{i,j} \langle b^j, ab_i \langle x^j, x_j \rangle \rangle = (2.6) \sum_i \langle b^j, ab_i \text{rank} P \rangle = \text{trace}(a_A \circ A \text{rank} P)
\]

as claimed. \( \square \)

**The Hattori-Stallings Rank Map**

If \( k \) is a splitting field for \( A \), then the character map yields an isomorphism of vector spaces (Proposition 1.49),

\[
\mathcal{R}_k(A) = \mathcal{R}(A) \otimes_k \mathbb{Z}_k \xrightarrow{\sim} C(A) \equiv (\text{Tr} A^\text{s.p.})^*.
\]
Our goal in this paragraph is to prove a version of this result for $P(A)$, with the Hattori-Stallings ranks replacing characters. This will further highlight the duality between $P(A)$ and $R(A)$. Let

$$\text{rank}_k : P_k(A) \overset{\text{def}}{=} P(A) \otimes \mathbb{Z}_k \rightarrow \text{Tr} A$$

denote the $k$-linear extension of the Hattori-Stallings rank map.

**Theorem 2.10.** Let $A \in \text{Alg}_k$ be finite dimensional and let $\tau : \text{Tr} A \rightarrow \text{Tr} A^s.p.$ denote the canonical epimorphism.

(a) If $\text{char}(k) = 0$, then $\tau \circ \text{rank}$ is a group monomorphism $P(A) \hookrightarrow \text{Tr} A^s.p.$

(b) If $k$ is a splitting field for $A$, then we have a $k$-linear isomorphism

$$\tau \circ \text{rank}_k : P_k(A) \xrightarrow{\sim} \text{Tr} A^s.p.$$ 

Thus, the images of the principal indecomposable representations $PS$ with $S \in \text{Irr} A$ form a $k$-basis of $\text{Tr} A^s.p.$.

**Proof.** Put $\rho = \tau \circ \text{rank}_k : P_k(A) \rightarrow \text{Tr} A^s.p.$ Then, for $S, S' \in \text{Irr} A$, we have

$$\langle \chi_{S'}, \rho[PS] \rangle = \langle \chi_{S'}, \text{rank} PS \rangle = \text{Proposition 2.5} \left(PS, S' \right) 1_k$$

$$= \delta_{S,S'} \dim_k D(S) 1_k$$

Thus, the images $\rho[PS]$ with $\dim_k D(S) 1_k \neq 0$ form a $k$-linearly independent subset of $\text{Tr} A^s.p.$ If $\text{char} k = 0$ or if $k$ is a splitting field for $A$, then this holds for all $S \in \text{Irr} A$. Since $P_k(A)$ is generated by the classes $[PS] \otimes 1$ (Propositions 2.4(a) and 2.8(a)), we obtain a $k$-linear embedding $\rho : P_k(A) \hookrightarrow \text{Tr} A^s.p.$ in these cases.

If $\text{char} k = 0$, then the canonical map $P(A) \overset{\sim}{=} \mathbb{Z}^{\oplus \text{Irr} A} \rightarrow P_k(A) \overset{\text{def}}{=} k^{\oplus \text{Irr} A}$ is an embedding, proving (a). For a splitting field $k$, we have $\dim_k \text{Tr} A^s.p. = \dim_k C(A) = \# \text{Irr} A$ (Theorem 1.44) and (b) follows. \qed

**Exercises for Section 2.1**

*Without any mention to the contrary, $A \in \text{Alg}_k$ is arbitrary in these exercises.*

2.1.1 (Injectives). A module $I \in _A\text{Mod}$ is called **injective** if $I$ satisfies the following equivalent conditions:

(i) Given a monomorphism $f : M \hookrightarrow N$ and an arbitrary $g : M \rightarrow I$ in $_A\text{Mod}$, there exists a “lift” $\overline{g} : N \rightarrow I$ in $_A\text{Mod}$ such that $\overline{g} \circ f = g$:

$$\begin{array}{c}
I \\
\downarrow \overline{g} \\
N \\
\downarrow f \\
M
\end{array}$$
(ii) Every monomorphism $f : I \hookrightarrow M$ in $\mathbf{A}_{\text{Mod}}$ splits: there exists $s : M \to I$ such that $s \circ f = \text{Id}_I$.

(iii) The functor $\text{Hom}_A(\cdot, I) : \mathbf{A}_{\text{Mod}} \to \text{Vect}_k$ is exact.

(a) Prove the equivalence of the above conditions.

(b) Let $A \to B$ be an algebra map. Show that $\text{Coid}^B_A : \mathbf{A}_{\text{Mod}} \to \mathbf{B}_{\text{Mod}}$ (§1.2.2) sends injectives of $A$ to injectives of $B$.

(c) Let $(M_i)_i$ be a family of $A$-modules. Show that the direct product $\prod_i M_i$ is injective if and only if all $M_i$ are injective.

2.1.2 (Semisimplicity). Show that the following are equivalent: (i) $A$ is semisimple; (ii) all $V \in \mathbf{A}_{\text{Mod}}$ are projective; (iii) all $V \in \mathbf{A}_{\text{Mod}}$ are injective.

2.1.3 (Morita contexts). A Morita context consists of the following data: algebras $A, B \in \text{Alg}_k$, bimodules $V \in \mathbf{A}_{\text{Mod}}_B$ and $W \in \mathbf{B}_{\text{Mod}}_A$, and bimodule homomorphisms $f : V \otimes_B W \to A$ and $g : W \otimes_A V \to B$ with $A, B$ being the regular bimodules. Writing $f(v \otimes w) = vw$ and $g(w \otimes v) = wv$, the maps $f$ and $g$ are required to satisfy the associativity conditions $(vw)v' = v(wv')$ and $(wv)w' = w(vw')$ for all $v, v' \in V$ and $w, w' \in W$. Under the assumption that $g$ is surjective, prove:

(a) $g$ is an isomorphism.

(b) Every left $B$-module is a homomorphic image of a direct sum of copies of $W$ and every right $B$-module is an image of a direct sum of copies of $V$.

(c) $V$ and $W$ are finitely generated projective as $A$-modules.

2.1.4 (Morita contexts and finiteness conditions). This problem assumes familiarity with Exercise 2.1.3 and uses the same notation. Let $(A, B, V, W, f, g)$ be a Morita context such that $A$ is right noetherian and $g$ is surjective. Prove:

(a) $B$ is right noetherian and $V$ is finitely generated as right $B$-module.

(b) If $A$ is also affine, then $B$ is affine as well.

2.1.5 (Hattori-Stallings rank). Let $e = (e_{ij}) \in \text{Mat}_n(A)$ be an idempotent matrix and let $P = A_{\text{reg}} e$ as in (2.1). Show that $\text{rank } P = \sum_i e_{ii} + [A, A]$.

2.1.6 (“Eilenberg swindle”). Let $P \in \mathbf{A}_{\text{Proj}}$ be arbitrary and let $F$ be a free $A$-module such that $F = P' \oplus Q$ with $P' \cong P$. Show that $F^{\oplus \mathbb{N}} = F \oplus F \oplus F \oplus \cdots$ is a free $A$-module satisfying $P \oplus F^{\oplus \mathbb{N}} \cong F^{\oplus \mathbb{N}}$. Conclude that if $K_0^A(A)$ is constructed exactly as $K_0(A)$ but using arbitrary projectives of $A$, then $K_0^A(A) = \{0\}$.

2.1.7 (Properties of projective covers). Assume that $A$ is finite dimensional. Let $V, W \in \text{Rep } A$ and let $\alpha : PV \to V$, $\beta : PW \to W$ be the projective covers. Prove:

(a) If $\phi : V \to W$ is a homomorphism in $\text{Rep } A$, then there exists a lift $\overline{\phi} : PV \to PW$ with $\phi \circ \alpha = \beta \circ \overline{\phi}$. Furthermore, any such $\overline{\phi}$ is surjective iff $\phi$ is surjective.

(b) $P(\text{head } V) \cong PV$.

(c) $P(V \oplus W) \cong PV \oplus PW$.  

2.1.8 (Local algebras). Assume that $A$ is finite dimensional and local, that is, $A^{s.p.} = A/\text{rad} A$ is a division algebra. Show that every $P \in \text{Proj}_m A$ is free.

2.1.9 (Equality in $K_0(A)$ and stable isomorphism). (a) For $P, Q \in A\text{proj}$, show that $[P] = [Q]$ holds in $K_0(A)$ if and only if $P$ and $Q$ are stably isomorphic, that is, $P \oplus A^r_{\text{reg}} \cong Q \oplus A^r_{\text{reg}}$ for some $r \in \mathbb{Z}_+$.

(b) Assume that $\text{Mat}_n(A)$ is directly finite for all $n$, that is, $xy = 1_{n \times n}$ implies $yx = 1_{n \times n}$ for $x, y \in \text{Mat}_n(A)$—this holds, for example, whenever $A$ is commutative or (right or left) noetherian; see [85, Chapter 5]. Show that $[P] \neq 0$ in $K_0(A)$ for any $0 \neq P \in A\text{proj}$.

2.1.10 ($K_0(A)$ and $\mathcal{P}(A)$). Show that the map $\mathcal{P}(A) \xrightarrow{} K_0(A)$, $[P] \mapsto [P]$, is a monomorphism of groups.

2.1.11 (Grothendieck groups of $A$ and $A^{s.p.}$). Assume that $A$ is finite dimensional. Recall that inflation along the canonical map $\alpha : A \rightarrow A^{s.p.}$ gives an isomorphism $\mathcal{R}(\alpha) : \mathcal{R}(A^{s.p.}) \xrightarrow{} \mathcal{R}(A)$ (Lemma 1.47). Show that induction along $\alpha$ gives an isomorphism $K_0(\alpha) : K_0(A) = \mathcal{P}(A) \xrightarrow{} K_0(A^{s.p.}) = \mathcal{P}(A^{s.p.}) = \mathcal{R}(A^{s.p.})$, $[P] \mapsto [\text{head } P]$.

2.1.12 (Some Cartan matrices). (a) Let $A = k[x]/(x^n)$ with $n \in \mathbb{N}$. Show that the Cartan matrix of $A$ is the $1 \times 1$-matrix $C = (n)$.

(b) Let $A$ be the algebra of upper triangular $n \times n$-matrices over $k$. Show that the Cartan matrix of $A$ is the upper triangular $n \times n$-matrix

$$C = \begin{pmatrix} 1 & \cdots & 1 \\ \vdots & \ddots & \vdots \\ 1 & \cdots & 1 \end{pmatrix}.$$ 

2.1.13 (Cartan matrix of the Sweedler algebra). Let $\text{char} k \neq 2$. The algebra $A = k(x, y)/(x^2, y^2 - 1, xy + yx)$ is called the Sweedler algebra.

(a) Realize $A$ as a homomorphic image of the quantum plane $Q_q(k^2)$ with $q = -1$ (Exercise 1.1.15) and use this to show that $\dim_k A = 4$.

(b) Show that rad $A = (x)$ and $A^{s.p.} \cong k \times k$. There are two irreducible $A$-modules, $k_x$, with $x.1 = 0$ and $y.1 = \pm 1$.

(c) Show that $e_{\pm} = \frac{1}{2}(1 \pm y) \in A$ are idempotents with $A = Ae_+ \oplus Ae_-$ and $xe_{\pm} = e_\pm x$. Conclude that $\mathcal{P}k_{\pm} = Ae_{\pm}$ and that the Cartan matrix of $A$ is $C = (1 \ 1)$.

2.2. Frobenius and Symmetric Algebras

This section features a special class of finite-dimensional algebras, called Frobenius algebras, with particular emphasis on the subclass of symmetric algebras. As we will see, all finite-dimensional semisimple algebras are symmetric and it is in fact often quite useful to think of semisimple algebras in this larger context. We will learn later that Frobenius algebras encompass all group algebras of finite groups and, more generally, all finite-dimensional Hopf algebras.
The material in this section is admittedly rather technical and focused on explicit formulae. However, the tools deployed here will see some heavy use in Chapter 12.

### 2.2.1. Definition of Frobenius and Symmetric Algebras

Recall that every $A \in \text{Alg}_k$ carries the regular $(A, A)$-bimodule structure (Example 1.2): the left and right actions of $a \in A$ on $A$ are respectively given by left multiplication, $A \cdot a$, and by right multiplication, $a \cdot A$. This structure gives rise to a bimodule structure on the linear dual, $A^* = \text{Hom}_k(A, k)$, for which it is customary to use the following notation:

$$a \cdot f \cdot b = f \circ b_A \circ a \quad (a, b \in A, f \in A^*).$$

Using $\langle \cdot, \cdot \rangle : A^* \times A \to k$ to denote the evaluation pairing, the $(A, A)$-bimodule action becomes

$$(2.19) \quad \langle a \cdot f \cdot b, c \rangle = \langle f, bca \rangle \quad (a, b, c \in A, f \in A^*).$$

The algebra $A$ is said to be Frobenius if $A^*$, viewed as a left $A$-module, is isomorphic to the left regular $A$-module $A_{\text{reg}} = A \cdot A$. We will see in Lemma 2.11 below that this condition is equivalent to corresponding right $A$-module condition. If $A^*$ and $A$ are in fact isomorphic as $(A, A)$-bimodules—this is not automatic from the existence of a one-sided module isomorphism—then the algebra $A$ is called symmetric. Note that even a mere isomorphism $A^* \cong A$ in $\text{Vect}_k$ forces $A$ to be finite dimensional (Appendix B); so Frobenius algebras will necessarily have to be finite dimensional.

### 2.2.2. Frobenius Data

For any finite-dimensional algebra $A$, a left $A$-module isomorphism $A^* \cong A$ amounts to the existence of an element $\lambda \in A^*$ such that $A^* = A \cdot \lambda$; likewise for the right side. The next lemma shows in particular that any left $A$-module generator $\lambda \in A^*$ also generates $A^*$ as right $A$-module and conversely. In the lemma, we will dispense with the summation symbol, and we shall continue to do so below:

*Summation over indices occurring twice is implied throughout Section 2.2.*

**Lemma 2.11.** Let $A \in \text{Alg}_k$ be a finite dimensional. Then the following are equivalent for any $\lambda \in A^*$:

1. $A^* = A \cdot \lambda$.
2. There exist elements $x_i, y_i \in A$ ($i = 1, \ldots, \dim_k A$) satisfying the following equivalent conditions:

   $$a = x_i \langle \lambda, ay_i \rangle \quad \text{for all } a \in A;$$
   $$\langle \lambda, x_i y_j \rangle = \delta_{i,j};$$
   $$a = y_i \langle \lambda, x_i a \rangle \quad \text{for all } a \in A.$$
(iii) $A^* = \lambda \leftarrow A$.

**Proof.** Conditions (2.20) and (2.21) both state that $(x_i)_i$ is a $k$-basis of $A$ and $(y_i \rightarrow \lambda)_i$ is the corresponding dual basis of $A^*$. Thus, (2.20) and (2.21) are equivalent, and they certainly imply that $\lambda$ generates $A^*$ as left $A$-module; so (i) follows. Conversely, if (i) is satisfied, then for any $k$-basis $(x_i)_i$ of $A$, the dual basis of $A^*$ has the form $(y_i \rightarrow \lambda)_i$ for suitable $y_i \in A$, giving (2.20), (2.21). Similarly, (2.22) and (2.21) both state that $(y_i)_i$ and $(\lambda \leftarrow x_i)_i$ are dual bases of $A$ and $A^*$, and the existence of such bases is equivalent to (iii). □

**Frobenius Form.** To summarize the foregoing, a finite-dimensional algebra $A$ is a Frobenius algebra if and only if there is a linear form $\lambda \in A^*$ satisfying the equivalent conditions of Lemma 2.11; any such $\lambda$ is called a Frobenius form. Note that the equality $A^* = \lambda \leftarrow A$ is equivalent to the condition that $0 \neq a \in A$ implies $\langle \lambda, Aa \rangle \neq 0$, which in turn is equivalent to the corresponding condition for $aA$. Thus, a Frobenius form is a linear form $\lambda \in A^*$ such that $\ker \lambda$ contains no nonzero left ideal or, equivalently, no nonzero right ideal of $A$. We will think of a Frobenius algebra as the pair

$$(A, \lambda).$$

A homomorphism of Frobenius algebras $f \colon (A, \lambda) \rightarrow (B, \mu)$ is a map in $\text{Alg}_k$ such that $\mu \circ f = \lambda$.

**Dual Bases.** The elements $(x_i, y_i)_i$ in Lemma 2.11 are called dual bases for $(A, \lambda)$. The identities (2.20) and (2.22) can be expressed by the diagram

$$\begin{array}{ccc}
\text{End}_k(A) & \xrightarrow{\sim} & A \otimes A^* \\
\psi & & \psi \\
\text{Id}_A & \xleftarrow{\text{can}} & x_i \otimes (y_i \rightarrow \lambda) = y_i \otimes (\lambda \leftarrow x_i)
\end{array}$$

(2.23)

**Nakayama Automorphism.** For a given Frobenius form $\lambda \in A^*$, Lemma 2.11 implies that

$$\lambda \rightarrow a = v_A(a) \rightarrow \lambda \quad (a \in A)$$

(2.24)

for a unique $v_A(a) \in A$. Thus, $\langle \lambda, ab \rangle = \langle \lambda, bv_A(a) \rangle$ for $a, b \in A$. This determines an automorphism $v_A \in \text{Aut}_{\text{Alg}}(A)$, which is called the Nakayama automorphism of $(A, \lambda)$. In terms of dual bases $(x_i, y_i)$, the Nakayama automorphism is given by

$$v_A(a) \overset{(2.22)}{=} y_i \langle \lambda, x_i v_A(a) \rangle = y_i \langle \lambda, ax_i \rangle.$$
Changing the Frobenius Form. The data associated to $A$ that we have assembled above, starting with a choice of Frobenius form $\lambda \in A^*$, are unique up to units. Indeed, for each unit $u \in A^\times$, the form $u \cdot \lambda \in A^*$ is also a Frobenius form and all Frobenius forms of $A$ arise in this way, because they are just the possible generators of the left $A$-module $A A^* \cong A^* A$. The reader will easily check that if $(x_i, y_i)$ are dual bases for $(A, \lambda)$, then $(x_i, y_i u^{-1})$ are dual bases for $(A, u \cdot \lambda)$ and the Nakayama automorphisms are related by

$$v_{u \cdot \lambda}(a) = uv_{\lambda}(a)u^{-1} \quad (a \in A).$$

2.2.3. Casimir Elements

Let $(A, \lambda)$ be a Frobenius algebra. The elements of $A \otimes A$ that correspond to $\Id_A$ under the two isomorphism $\End_k(A) \cong A \otimes A^* \cong A \otimes A$ that are obtained by identifying $A^*$ and $A$ via $\lambda \cdot - \lambda$ and $\lambda -\lambda$, will be referred to as the Casimir elements of $(A, \lambda)$; they will be denoted by $c_{-\lambda}$ and $c_{\lambda-\lambda}$, respectively. By (2.23), the Casimir elements are given by

$$c_{-\lambda} = x_i \otimes y_i \quad \text{and} \quad c_{\lambda-\lambda} = y_i \otimes x_i$$

Thus $c_{\lambda-\lambda} = \tau(c_{-\lambda})$, where $\tau \in \Aut_{\text{Alg}}(A \otimes A)$ is the switch map, $\tau(a \otimes b) = b \otimes a$. The Casimir elements do depend on $\lambda$ but not on the choice of dual bases $(x_i, y_i)$.

Lemma 2.12. Let $(A, \lambda)$ be a Frobenius algebra with Nakayama automorphism $v_{\lambda} \in \Aut_{\text{Alg}}(A)$. The following identities hold in the algebra $A \otimes A$, with $a, b \in A$:

(a) $c_{-\lambda} = (\Id \otimes v_{\lambda})(c_{\lambda-\lambda}) = (v_{\lambda} \otimes \Id)(c_{\lambda-\lambda})$;
(b) $c_{\lambda-\lambda} = v_{\lambda} \circ (\Id)(c_{\lambda-\lambda}) = (v_{\lambda} \otimes v_{\lambda})(c_{\lambda-\lambda})$;
(c) $(a \otimes b)c_{-\lambda} = c_{-\lambda}(b \otimes v_{\lambda}(a))$;
(d) $(a \otimes b)c_{\lambda-\lambda} = c_{\lambda-\lambda}(b \otimes v_{\lambda}(a))$.

Proof. The identities in (b) and (d) follow from those in (a) and (c) by applying $\tau$; we will focus on the latter. For (a), we calculate

$$x_i \otimes (\lambda - \lambda) = y_i \otimes (\lambda - x_i) \quad \text{by (2.23)}, \quad y_i \otimes (\lambda - x_i) \otimes (\lambda - x_i) \quad \text{by (2.24)}, \quad y_i \otimes (\lambda - x_i) \otimes (\lambda - x_i).$$

This gives $x_i \otimes y_i = y_i \otimes v_{\lambda}(x_i)$ or $c_{-\lambda} = (\Id \otimes v_{\lambda})(c_{\lambda-\lambda})$ by (2.27). Applying $v_{\lambda} \otimes \Id$ to this identity, we obtain $(v_{\lambda} \otimes \Id)(c_{-\lambda}) = (v_{\lambda} \otimes v_{\lambda})(c_{\lambda-\lambda})$ and then $\tau$ yields $(\Id \otimes v_{\lambda})(c_{\lambda-\lambda}) = (v_{\lambda} \otimes v_{\lambda})(c_{\lambda-\lambda})$. This proves (a). Part (c) follows from the computations

$$ax_i \otimes y_i \equiv x_j \otimes y_j \otimes (\lambda - x_i) \otimes (\lambda - x_i) \equiv x_j \otimes (\lambda - x_i) \otimes (\lambda - x_i).$$

In later chapters, we will consider similar elements, also called Casimir elements, for semisimple Lie algebras; see (5.54) and §6.2.1.
and
\[ x_i \otimes b y_i \equiv x_i \otimes y_j \langle \lambda, x_j b y_i \rangle = x_i \langle \lambda, x_j b y_i \rangle \otimes y_j \equiv x_j b \otimes y_j. \]

**Casimir Operator and Higman Trace.** We now discuss two closely related operators that were originally introduced by D. G. Higman [99]. Continuing with the notation of (2.23), (2.27), they are defined by

\[ \gamma \rightarrow \lambda : A \rightarrow \text{Tr} A \rightarrow A \]
\[ \gamma \leftarrow \lambda : A \rightarrow \mathcal{Z} A \leftarrow A \]

(2.28)

\[ a \rightarrow x_i a y_j \quad a \rightarrow y_i a x_j \]

The operator \( \gamma \rightarrow \lambda \) will be called the **Higman trace** and \( \gamma \leftarrow \lambda \) will be referred to as the **Casimir operator**. The following lemma justifies the claims, implicit in (2.28), that the Higman trace does indeed factor through the universal trace \( \text{Tr} : A \rightarrow \text{Tr} A = A/[A, A] \) and the Casimir operator has values in the center \( \mathcal{Z} A \).

**Lemma 2.13.** Let \((A, \lambda)\) be a Frobenius algebra with Nakayama automorphism \( \nu \lambda \in \text{Aut}_{\text{Alg}}(A) \). Then, for all \( a, b, c \in A \),

\[ a \gamma \rightarrow \lambda (bc) = \gamma \rightarrow \lambda (cb) \nu \lambda (a) \quad \text{and} \quad a \gamma \leftarrow \lambda (bc) = (\nu \lambda (c) b) a. \]

**Proof.** The identity in Lemma 2.12(c) states that \( ax_i \otimes by_i = x_i b \otimes y_i \nu \lambda (a) \). Multiplying this on the right with \( c \otimes 1 \) and then applying the multiplication map of \( A \) gives \( ax_i cb y_i = x_i b c y_i \nu \lambda (a) \) or, equivalently, \( a \gamma \rightarrow \lambda (cb) = \gamma \rightarrow \lambda (bc) \nu \lambda (a) \). The formula for \( \gamma \leftarrow \lambda \) follows in the same way from Lemma 2.12(d).

\[ \square \]

### 2.2.4. Traces

In this subsection, we use the Frobenius structure to derive some trace formulae that will be useful later on. To start with, the left and right regular representation of any Frobenius algebra \( A \) have the same character. Indeed, for any \( a \in A \), we compute

\[ \text{trace}(a_A) = \text{trace}(a_{A^*}) = \text{trace}((A a)^*) = \text{trace}(A a), \]

where the first equality uses the fact that \( A^* \cong A \) in \( \text{Mod}_A \) and the second is due to the switch in sides when passing from \( A \) to \( A^* \) in (2.19).

**Lemma 2.14.** Let \((A, \lambda)\) be a Frobenius algebra with dual bases \((x_i, y_i)\). Then, for any \( f \in \text{End}_k(A) \),

\[ \text{trace}(f) = \langle \lambda, f(x_i)y_i \rangle = \langle \lambda, x_i f(y_i) \rangle. \]
Proof. By (2.23), we have

\[
\begin{array}{ccc}
\text{End}_k(A) & \xrightarrow{\text{can.}} & A \otimes A^* \\
\downarrow f & & \downarrow \text{can.} \\
\text{End}_k(A) & \xrightarrow{\text{trace}} & A \otimes A^*
\end{array}
\]

Since the trace function on \(\text{End}_k(A)\) becomes evaluation on \(A \otimes A^*\), we obtain the formula in the lemma.

With \(f = b_A \circ A a\) for \(a, b \in A\), Lemma 2.14 yields the following expression for the map \(t: A \to A^*\) from (2.17) in terms of the Higman trace:

\[
\text{(2.29)} \quad \text{trace}(b_A \circ A a) = \langle \lambda, b_A \gamma_A(a) \rangle = \langle \lambda, \gamma_A(b) a \rangle.
\]

Equation (2.29) shows in particular that the left and right regular representation of \(A\) have the same character, as was already shown above:

\[
\text{(2.30)} \quad \chi_{\text{reg}}(a) = \text{trace}(A a) = \text{trace}(A a) = \langle \lambda, \gamma_A(1) a \rangle = \langle \lambda, a \gamma_A(1) \rangle.
\]

2.2.5. Symmetric Algebras

Recall that the algebra \(A\) is symmetric if there is an isomorphism \(A \xrightarrow{\sim} A^*\) in \(A\text{-Mod}_A\). In this case, the image of \(1 \in A\) will be a Frobenius form \(\lambda \in A^*\) such that \(a \mapsto \lambda = \lambda \mapsto a\) holds for all \(a \in A\). Thus:

\[
\text{(2.31)} \quad \nu_A = \text{Id}_A \quad \text{or, equivalently,} \quad \lambda \in A^*_{\text{trace}}.
\]

Recall also that Frobenius forms \(\lambda \in A^*\) are characterized by the condition that \(\text{Ker} \lambda\) contains no nonzero left or right ideal of \(A\). For \(\lambda \in A^*_{\text{trace}}\), this is equivalent to saying that \(\text{Ker} \lambda\) contains no nonzero two-sided ideal of \(A\), because \(\langle \lambda, A a \rangle = \langle \lambda, A a \rangle = \langle \lambda, a A \rangle = \langle \lambda, a A \rangle\) for \(a \in A\). Thus, a finite-dimensional algebra \(A\) is symmetric if and only if there is a trace form \(\lambda \in A^*_{\text{trace}}\) such that \(\text{Ker} \lambda\) contains no nonzero ideal of \(A\). In light of (2.26), a symmetric algebra is also the same as a Frobenius algebra \(A\) possessing a Frobenius form \(\lambda \in A^*\) such that the Nakayama automorphism \(\nu_A\) is an inner automorphism of \(A\), in which case the same holds for any Frobenius form of \(A\). When dealing with a symmetric algebra \(A\), it will be convenient to always fix a Frobenius form as in (2.31); this then determines \(\lambda\) up to a central unit of \(A\).

Casimir Element and Trace. Let us note some consequences of (2.31). First, \(c_{\sim \lambda} = c_{\lambda_{\sim}}\) by Lemma 2.12(a). Therefore, the Casimir operator is the same as the Higman trace, \(\gamma_{\sim \lambda} = \gamma_{\lambda_{\sim}}\). We will simply write \(c_A\) and \(\gamma_A\), respectively, and refer to \(\gamma_A\) as the Casimir trace of \((A, \lambda)\). Fixing dual bases \((x_i, y_i)\) for \((A, \lambda)\) as in (2.27) and (2.28), the Casimir element and trace are given by

\[
\text{(2.32)} \quad c_A = x_i \otimes y_i = y_i \otimes x_i
\]
and

\[ \gamma_A: A \to \text{Tr} A \to Z'A \]

\[ a \mapsto x_i y_i = y_i a x_i \]

The square of the Casimir element, \( c^2 \), belongs to \( Z'(A \otimes A) = Z'A \otimes Z'A \), because \((a \otimes b)c_A = c_A(b \otimes a)\) for all \( a, b \in A \) by Lemma 2.12(c).

**Example 2.15** (Matrix algebras). Let \( A = \text{Mat}_n(k) \) be the \( n \times n \) matrix algebra. Then we can take the ordinary matrix trace as Frobenius form:

\[ \lambda = \text{trace} \]

Dual bases for this form are provided by the standard matrices \( e_{j,k} \), with 1 in the \((j, k)\)-position and 0s elsewhere: \( \text{trace}(e_{j,k} e_{k',j'}) = \delta_{j,j'} \delta_{k,k'} \). Thus, with implied summation over both \( j \) and \( k \), the Casimir element is

\[ e_{\text{trace}} = e_{j,k} \otimes e_{k,j} \]

and its square is \( c^2_{\text{trace}} = 1_{n \times n} \otimes 1_{n \times n} \). By (2.33), the Casimir trace of a matrix \( a = (a_{i,m}) \in A \) is \( e_{j,k} a e_{k,j} = a_{k,k} e_{j,j} \); so

\[ \gamma_{\text{trace}}(a) = \text{trace}(a) 1_{n \times n} \quad (a \in A) \]

Now (2.30) gives the formula \( \chi_{\text{reg}}(a) = n \text{ trace}(a) \) for the regular character. (This was already observed in Exercise 1.5.2.)

### 2.2.6. Semisimple Algebras as Symmetric Algebras

**Proposition 2.16.** Every finite-dimensional semisimple \( k \)-algebra is symmetric.

**Proof.** Let \( A \in \text{Alg}_k \) be finite dimensional and semisimple. Wedderburn’s Structure Theorem allows us to assume that \( A \) is in fact simple, because a finite direct product of algebras is symmetric if all its components are (Exercise 2.2.1). By Theorem 1.44(b), we also know that \( A_{\text{trace}} \neq 0 \). If \( \lambda \) is any nonzero trace form, then \( \ker \lambda \) contains no nonzero ideal of \( A \), by simplicity. Thus, \( \lambda \) serves as a Frobenius form for \( A \).

Now let \( A \) be split semisimple. The algebra structure of \( A \) is completely determined by the degrees of the irreducible representations of \( A \), in view of the Wedderburn isomorphism (1.46):

\[ A \xrightarrow{\sim} \prod_{S \in \text{Irr} A} \text{End}_k(S) \cong \prod_{S \in \text{Irr} A} \text{Mat}_{\text{dim} S}(k) \]

\[ a \mapsto (a_S) \]
The determination of all \( \dim_k S \) for a given \( A \) can be a formidable task, however, while it is comparatively easy to come up with a Frobenius form \( \lambda \in A_{\text{trace}}^* \) and assemble the Frobenius data of \( A \). Therefore, it is of interest to figure out what information concerning the degrees \( \dim_k S \) can be gleaned from these data. We shall in particular exploit the Casimir square \( c^2_\lambda \in \mathcal{Z}A \otimes \mathcal{Z}A \) and the value \( \gamma_\lambda(1) \in \mathcal{Z}A \) of the Casimir trace \( \gamma_\lambda \). Note that the operator \( c_S \in \text{End}_k(S) \) is a scalar for \( c \in \mathcal{Z}A \); so we may view \( \gamma_\lambda(1)_S \in k \) for all \( S \in \text{Irr} A \).

Recall that the primitive central idempotent \( e(S) \in \mathcal{Z}A \) is the element corresponding to \( (0, \ldots, 0, \text{Id}_S, 0, \ldots, 0) \in \prod_{S \in \text{Irr} A} \text{End}_k(S) \) under the above isomorphism; so \( \mathcal{Z}A = \bigoplus_{S \in \text{Irr} A} k e(S) \) and

\[
e(S)_T = \delta_{S,T} \text{Id}_S \quad (S, T \in \text{Irr} A).
\]

Our first goal is to give a formula for \( e(S) \) in terms of Frobenius data of \( A \). We will also describe the image of the Casimir square \( c^2_\lambda \) under the following isomorphism coming from the Wedderburn isomorphism:

\[
A \otimes A \xrightarrow{\sim} \prod_{S, T \in \text{Irr} A} \text{End}_k(S) \otimes \text{End}_k(T)
\]

(2.34)

\[
a \otimes b \longmapsto (a \otimes b)_{S, T} := (a_S \otimes b_T)
\]

**Theorem 2.17.** Let \( A \in \text{Alg}_k \) be split semisimple, with Frobenius form \( \lambda \in A_{\text{trace}}^* \). Then, for all \( S, T \in \text{Irr} A \),

(a) \( e(S) \gamma_\lambda(1)_S = (\dim_k S) (S \otimes \text{Id}_A)(c_\lambda) = (\dim_k S) (\text{Id}_A \otimes S)(c_\lambda) \) and

\[ \gamma_\lambda(1)_S = 0 \text{ if and only if } (\dim_k S) 1_k = 0. \]

(b) \( (c_\lambda)_{S,T} = 0 \text{ if } S \neq T \text{ and } (\dim_k S)^2 (c^2_\lambda)_{S,S} = \gamma_\lambda(1)^2_S. \)

**Proof.** (a) The equality \( (S \otimes \text{Id}_A)(c_\lambda) = (\text{Id}_A \otimes S)(c_\lambda) \) follows from (2.32). In order to prove the equality \( e(S) \gamma_\lambda(1)_S = (\dim_k S) (\text{Id}_A \otimes S)(c_\lambda) \), use (2.20) to write \( e(S) = x_i(\lambda, e(S), y_i) \). We need to show that \( \langle \lambda, e(S)y_i \rangle \gamma_\lambda(1)_S = (\dim_k S) S \gamma_\lambda(1)_S \) for all \( i \) or, equivalently,

\[
\langle \lambda, ae(S) \rangle \gamma_\lambda(1)_S = (\dim_k S) S \gamma_\lambda(a) \quad (a \in A).
\]

(2.35)

For this, we use the regular character:

\[
\chi_{\text{reg}}(ae(S)) = \langle \lambda, ae(S) \gamma_\lambda(1) \rangle = \langle \lambda, ae(S) \gamma_\lambda(1)_S \rangle = \langle \lambda, ae(S) \gamma_\lambda(1)_S \rangle.
\]

(2.30)

On the other hand, the isomorphism \( A_{\text{reg}} \cong \bigoplus_{T \in \text{Irr} A} T \otimes \dim_k T \) from Wedderburn’s Structure Theorem gives \( \chi_{\text{reg}} = \sum_{T \in \text{Irr} A} (\dim_k T) \chi_T \). Since \( e(S) \gamma_T = \delta_{S,T} \chi_S \), we obtain

\[
e(S) \longrightarrow \chi_{\text{reg}} = (\dim_k S) \chi_S.
\]

(2.36)
Thus, $\chi_{\text{reg}}(ae(S)) = (\dim_k S)\chi_S(a)$, proving (2.35). Since $\chi_S$ and $\langle \lambda, , e(S) \rangle$ are both nonzero, (2.35) also shows that $\gamma_A(1)_S = 0$ if and only if $(\dim_k S)1_k = 0$.

(b) For $S \neq T$, the identity $(a \otimes b)c_A = c_A(b \otimes a)$ from Lemma 2.12 gives

$$
(c_A)_{S,T} = ((e(S) \otimes e(T))c_A)_{S,T} = (c_A(e(T) \otimes e(S)))_{S,T}.
$$

It remains to consider the case $S = T$. Here, the identity in Lemma 2.12 gives

$$
(c_A)_{S,T} = (c_A)_{S,S} = 0.
$$

It follows from the above that $\chi_S(c) = (\dim_k S)c_S$. Therefore, writing $a_S = \rho_S(a)$ for $a \in A$, we calculate

$$
(\dim_k S)(\rho_S \circ \gamma_A)(a) = (\chi_S \circ \gamma_A)(a) = \chi_S(a)\gamma_A(x_i,y_i) = \chi_S(a)\gamma_A(1)_S
$$

and further

$$
(\dim_k S)^2 (c_A^2)_{S,S} = (\dim_k S)^2 (\rho_S \circ \rho_S)(\gamma_A \otimes \text{Id})(c_A)
$$

$$
= (\dim_k S)^2 ((\rho_S \circ \gamma_A) \otimes \rho_S)(c_A)
$$

$$
= (\dim_k S) (\chi_S \otimes \rho_S)(c_A)\gamma_A(1)_S
$$

$$
= \rho_S ((\dim_k S)(\chi_S \otimes \text{Id})(c_A))\gamma_A(1)_S
$$

$$
= \rho_S (e(S)\gamma_A(1)_S)\gamma_A(1)_S = \gamma_A(1)_S^2.
$$

This completes the proof of the theorem. $\square$

### 2.2.7. Integrality and Divisibility

Theorem 2.17 is a useful tool in proving certain divisibility results for the degrees of irreducible representations. For this, we recall some standard facts about integrality; proofs can be found in most textbooks on commutative algebra or algebraic number theory. Let $R$ be a ring and let $S$ be a subring of the center $\mathcal{Z}R$. An element $r \in R$ is said to be **integral** over $S$ if $f(r) = 0$ for some monic polynomial $f \in S[x]$. The following facts will be referred to repeatedly, in later sections as well:

- An element $r \in R$ is integral over $S$ if and only if $r \in R'$ for some subring $R' \subseteq R$ such that $R'$ contains $S$ and is finitely generated as an $S$-module.
- If $R$ is commutative, then the elements of $R$ that are integral over $S$ form a subring of $R$ containing $S$; it is called the **integral closure** of $S$ in $R$.
- An element of $\mathbb{Q}$ that is integral over $\mathbb{Z}$ must belong to $\mathbb{Z}$.

The last fact above reduces the problem of showing that a given nonzero $s \in \mathbb{Z}$ divides another $t \in \mathbb{Z}$ to proving that the fraction $t/s$ is merely integral over $\mathbb{Z}$.
2.2. Frobenius and Symmetric Algebras

Corollary 2.18. Let $A$ be a split semisimple algebra over a field $\mathbb{k}$ of characteristic 0 and let $\lambda \in A^*_{\text{trace}}$ be a Frobenius form such that $\gamma_A(1) \in \mathbb{Z}$. Then the following are equivalent:

(i) The degree of every irreducible representation of $A$ divides $\gamma_A(1)$;

(ii) the Casimir element $c_A$ is integral over $\mathbb{Z}$.

Proof. Theorem 2.17 gives the formula

$$ (c_A^2)_{S,S} = \left( \frac{\gamma_A(1)}{\dim_S} \right)^2. $$

If (i) holds, then the isomorphism (2.34) sends $\mathbb{Z}[c_A^2]$ to $\prod_{S \in \text{Irr } A} \mathbb{Z}$, because $(c_A)_{S,T} = 0$ for $S \neq T$ (Theorem 2.17). Thus, $\mathbb{Z}[c_A]$ is a finitely generated $\mathbb{Z}$-module and (ii) follows. Conversely, (ii) implies that $c_A^2$ also satisfies a monic polynomial over $\mathbb{Z}$ and all $(c_A^2)_{S,S}$ satisfy the same polynomial. Therefore, the fractions $\frac{\gamma_A(1)}{\dim_S}$ must be integers, proving (i). □

Corollary 2.19. Let $A$ be a split semisimple $\mathbb{k}$-algebra, with $\text{char } \mathbb{k} = 0$, and let $\lambda \in A^*_{\text{trace}}$ be a Frobenius form for $A$. Furthermore, let $\phi: (A, \lambda) \to (B, \mu)$ be a homomorphism of Frobenius $\mathbb{k}$-algebras and assume that $\gamma_{-\mu}(1) \in \mathbb{k}$. Then, for all $S \in \text{Irr } A$,

$$ \frac{\gamma_{-\mu}(1)}{\dim_S \text{Ind}_A^B S} = \frac{\gamma_A(1)}{\dim_S} \cdot \frac{\gamma_{-\mu}(1)}{\dim_S}. $$

If the Casimir element $c_A$ is integral over $\mathbb{Z}$, then so is the scalar $\frac{\gamma_{-\mu}(1)}{\dim_S \text{Ind}_A^B S} \in \mathbb{k}$.

Proof. Putting $e := e(S)$, we have $S^\oplus \dim_S = Ae$ and so $\text{Ind}_A^B S^\oplus \dim_S = B \phi(e)$. Since $\phi(e) \in B$ is an idempotent, we have $\dim_B B \phi(e) = \text{trace}(B \phi(e))$ (Exercise 1.5.1). Therefore,

$$ \dim \text{Ind}_A^B S^\oplus \dim_S = \text{trace}(B \phi(e)) \equiv \langle \mu, \phi(e) \gamma_{-\mu}(1) \rangle = \langle \mu, \phi(e) \rangle \gamma_{-\mu}(1) $$

$$ \equiv \langle \lambda, e \rangle \gamma_{-\mu}(1) = \frac{(\dim_S)^2}{\gamma_A(1)} \gamma_{-\mu}(1). $$

The claimed equality $\frac{\gamma_{-\mu}(1)}{\dim_S \text{Ind}_A^B S} = \frac{\gamma_A(1)}{\dim_S}$ is immediate from this. Finally, Theorem 2.17 gives $\left( \frac{\gamma_{-\mu}(1)}{\dim_S \text{Ind}_A^B S} \right)^2 = (c_A^2)_{S,S}$, which is integral over $\mathbb{Z}$ if $c_A$ is. □

2.2.8. Separability

A finite-dimensional $\mathbb{k}$-algebra $A$ is called separable if $K \otimes A$ is semisimple for every field extension $K/\mathbb{k}$. The interested reader is referred to Exercises 1.4.10 and 1.5.6 for more on separable algebras. Here, we give a characterization of separability in terms of the Casimir operator, which is due to D. G. Higman [99].
Proposition 2.20. Let $(A, \lambda)$ be a Frobenius algebra. Then $\gamma_{\rightarrow A}$ and $\gamma_{A \leftarrow}$ both vanish on $\operatorname{rad} A$ and their images are contained in the socle of $A_{\text{reg}}$.

Proof. As we have already observed in (2.18), the operator $b_A \circ A a \in \operatorname{End}_k(A)$ is nilpotent if $a$ or $b$ belong to $\operatorname{rad} A$, and hence $\text{trace}(b_A \circ A a) = 0$. Consequently, (2.29) gives $\langle \lambda, A \gamma_{\leftarrow A} \rangle = 0$ and $\langle \lambda, \operatorname{rad} A \cdot A \gamma_{\leftarrow A} \rangle = 0$. Since the Frobenius form $\lambda$ does not vanish on nonzero left ideals, we must have $\gamma_{\leftarrow A} \operatorname{rad} A = 0$ and $\operatorname{rad} A \cdot A \gamma_{\leftarrow A} = 0$. This shows that $\operatorname{rad} A \subseteq \operatorname{Ker} \gamma_{\leftarrow A}$ and $\operatorname{Im} \gamma_{\leftarrow A} \subseteq \operatorname{soc} A_{\text{reg}}$.

For $\gamma_{A \leftarrow}$, we first compute

$$\langle \lambda, b \gamma_{A \leftarrow}(a) \rangle = \langle \lambda, b \gamma_{A \leftarrow}(a) \rangle = \text{trace}(v_A \circ b_A \circ A a).$$

The operator $v_A \circ b_A \circ A a \in \operatorname{End}_k(A)$ is again nilpotent if $a$ or $b$ belong to $\operatorname{rad} A$, because its $n$th power has image in $(\operatorname{rad} A)^n$ in this case. We can now repeat the above reasoning verbatim to obtain the same conclusions for $\gamma_{A \leftarrow}$. \hfill \square

For any Frobenius algebra $(A, \lambda)$, the Casimir operator $\gamma_{A \leftarrow}: A \rightarrow \mathcal{Z} A$ is $\mathcal{Z} A$-linear. Hence, the image $\gamma_{A \leftarrow}(A)$ is an ideal of $\mathcal{Z} A$. This ideal does not depend on the choice of Frobenius form $\lambda$; indeed, if $\lambda' \in A^*$ is another Frobenius form, then $\gamma_{A \leftarrow}(a) = \gamma_{A \leftarrow}(ua)$ for some unit $u \in A^*$ (§2.2.2). Thus, we may define

$$\Gamma A \overset{\text{def}}{=} \gamma_{A \leftarrow}(A).$$

Theorem 2.21. The following are equivalent for a finite-dimensional $A \in \operatorname{Alg}_k$:

(i) $A$ is separable;

(ii) $A$ is symmetric and $\Gamma A = \mathcal{Z} A$;

(iii) $A$ is Frobenius and $\Gamma A = \mathcal{Z} A$.

Proof. The proof of (i) $\Rightarrow$ (ii) elaborates on the proof of Proposition 2.16; we need to make sure that the current stronger separability hypothesis on $A$ also gives $\Gamma A = \mathcal{Z} A$. As in the earlier proof, Exercise 2.2.1 allows us to assume that $A$ is simple. Thus, $F := \mathcal{Z} A$ is a field and $F/k$ is a finite separable field extension (Exercise 1.5.6). It suffices to show that $\Gamma(A) \neq 0$. For this, let $\overline{F}$ denote an algebraic closure of $F$. Then $A \otimes_F \overline{F} \cong \operatorname{Mat}_n(\overline{F})$ for some $n$ (Exercise 1.1.13). The ordinary trace map trace: $\operatorname{Mat}_n(\overline{F}) \rightarrow \overline{F}$ is nonzero on $A$, since $A$ generates $\operatorname{Mat}_m(\overline{F})$ as $\overline{F}$-vector space. It is less clear, that the restriction of the trace map to $A$ has values in $F$, but this is in fact the case (e.g., Reiner [173, (9.3)]), giving a trace form\footnote{This map is called the \textit{reduced trace} of the central simple $F$-algebra $A$.}: $A \rightarrow F$. Since $F/k$ is finite separable field extension, we also have the field trace $T_{F/k}: F \rightarrow k$; this is the same as the regular character $\chi_{\text{reg}}$ of the $k$-algebra $F$ (Exercise 1.5.5). The composite $A := T_{F/k} \circ \text{tr}: A \rightarrow k$ gives a nonzero trace form for the $k$-algebra $A$, which we may take as our Frobenius form.
If \((a_i, b_i)\) are dual \(F\)-bases for \((A, \text{tr})\) and \((e_j, f_j)\) are dual \(k\)-bases for \((F, T_F/k)\), then \((e_j a_i, b_i f_j)\) are dual \(k\)-bases for \((A, A)\): \(\langle \lambda, e_j a_i b_i f_j \rangle = \delta_{i,j} \delta_{i,j}^\ast\). Moreover, 
\[\gamma_{\text{tr}} = \text{tr} \text{ by Example 2.15, and so} \]
\[
\Gamma A = e_j a_i \ b_i f_j = (\gamma_{T_F/k} \circ \gamma_{\text{tr}})(A) = (\gamma_{T_F/k} \circ \text{tr})(A) = \gamma_{T_F/k}(F).
\]
This is nonzero, because \(T_F/k \circ \gamma_{T_F/k} = \gamma_{T_F/k} = \text{reg}\).

The implication (ii)\(\Rightarrow\)(iii) being trivial, let us turn to the proof of (iii)\(\Rightarrow\)(i). Here, we can be completely self-contained. Note that the properties in (iii) are preserved under any field extension \(K/k\): If \(\lambda \in A^\ast\) is a Frobenius form for \(A\) such that \(\gamma_{\lambda(-)}(a) = 1\) for some \(a \in A\), then \(\lambda_K = \text{Id}_K \otimes \lambda\) is a Frobenius form for \(A_K = K \otimes A\)—any pair of dual bases \((x_i, y_i)\) for \((A, A)\) also works for \((A_K, \lambda_K)\)—and \(\gamma_{\lambda_K(-)} = \gamma_{\lambda(-)} = 1\). Thus, it suffices to show that (iii) implies that \(A\) is semisimple. But \(1 = \gamma_{\lambda(-)} \in \text{soc} A\) by Proposition 2.20. Hence \(\text{soc} A = A\), proving that \(A\) is semisimple. \(\square\)

### 2.2.9. Projectives and Injectives

In this subsection, we assume that the reader is familiar with the material in Section 2.1, including Exercise 2.1.1.

We start with some remarks on duality, for which \(A \in \text{Alg}_k\) can be arbitrary. For any \(M \in \text{Mod}_A\), the action \(-\) in (2.19) makes the linear dual \(M^\ast = \text{Hom}_k(M, k)\) a right \(A\)-module: \(\langle f - a, m \rangle = \langle f, am \rangle\). Likewise, the dual \(N^\ast\) of any \(N \in \text{Mod}_A\) becomes a left \(A\)-module via \(-\). Moreover, for any map \(\phi: M \to M'\) in \(\text{Mod}_A\), the dual map \(\phi^\ast: (M')^\ast \to M^\ast\) is a map in \(\text{Mod}_A\) and similarly for maps in \(\text{Mod}_A\). In this way, the familiar contravariant and exact functor \(\cdot^\ast: \text{Vec}_k \to \text{Vec}_k\) restricts to functors, also contravariant and exact,
\[
\cdot^\ast: \text{Mod}_A \to \text{Mod}_A \quad \text{and} \quad \cdot^\ast: \text{Mod}_A \to \text{Mod}_A.
\]
We will focus on finite-dimensional modules \(M \in \text{Mod}_A\) in the proposition below; see Exercise 2.2.7 for the general statement. In this case, the canonical isomorphism (B.22) in \(\text{Vec}_k\) is in fact an isomorphism \(M \cong M^{\ast\ast}\) in \(\text{Mod}_A\).

**Proposition 2.22.** Let \(A \in \text{Alg}_k\) be Frobenius. Then finite-dimensional \(A\)-modules are projective if and only if they are injective.

**Proof.** First, let \(A\) be an arbitrary finite-dimensional \(k\)-algebra and let \(M \in \text{Mod}_A\) be finite-dimensional. We claim that
\[
(2.40) \quad \text{\(M\) is projective } \iff \text{\(M^\ast\) is injective.}
\]
To prove this, assume that \(M\) is projective and let \(f: M^\ast \to N\) be a monomorphism in \(\text{Mod}_A\). We must produce a map \(g: N \to M^\ast\) such that \(g \circ f = \text{Id}_{M^\ast}\) (Exercise 2.1.1). But the dual map \(f^\ast: N^\ast \to M^{\ast\ast}\) splits, since \(M\) is projective; so there is a map \(s: M \to N^\ast\) in \(\text{Mod}_A\) with \(f^\ast \circ s = \text{Id}_M\) (Proposition 2.1). Then \(s^\ast \circ f^\ast = \text{Id}_{M^\ast}\) and \(f = f^{\ast\ast}: M^\ast \cong M^{\ast\ast} \to N \cong N^{\ast\ast}\); so we may take \(g = s^\ast\).
Theorem 2.24. Therefore, is a splitting field for \( P \). Moreover, \( \text{functor} \cdot \) identifies \( T \) with an irreducible left ideal of \( A \). Therefore, we may unambiguously speak of the Nakayama twist \( {}^\nu M \) of \( M \). If \( A \) is symmetric, then \( {}^\nu M \cong M \).

**Proposition 2.23.** Let \( A \in \text{Alg}_k \) be Frobenius and let \( S \in \text{Irr} \ A \). Then \( S \cong {}^\nu (\text{soc} \ PS) \).

**Proof.** Put \( P = PS \) and \( T = \text{soc} \ P \). The inclusion \( T \hookrightarrow P \) gives rise to an epimorphism \( P^* \to T^* \) in \( \text{Mod}_A \). Here, \( T^* \) is completely reducible, because the functor \( \cdot^* \) commutes with finite direct sums and preserves irreducibility of \( A \)-modules. Moreover, \( P^* \) is projective by Proposition 2.22 and (2.40) and \( P^* \) is indecomposable, because \( P \) is. Therefore, \( P^* \) has an irreducible head, whence \( T^* \) must be irreducible. Thus, \( T \) is irreducible as well. In order to describe \( T \) more precisely, write \( P \cong Ae \) for some idempotent \( e = e^2 \in A \) (Proposition 2.8) and identify \( T \) with an irreducible left ideal of \( A \) satisfying \( T = Te \). Fixing a Frobenius form \( \lambda \in A^* \) for \( A \), the image \( v_\lambda(T) \) is an irreducible left ideal of \( A \) such that

\[
0 \neq \langle \lambda, Te \rangle \overset{(2.24)}{=} \langle \lambda, e v_\lambda(T) \rangle.
\]

Therefore, \( ex \neq 0 \) for some \( x \in v_\lambda(T) \) and \( ae \mapsto aex \) gives an epimorphism \( P \cong Ae \to v_\lambda(T) \). Since head \( P \cong S \), it follows that \( v_\lambda(T) \cong S \), which proves the proposition. \( \square \)

The main result of this subsection concerns the Cartan matrix of \( A \). If \( k \) is a splitting field for \( A \), then the Cartan matrix has entries \( (PS, PS') \), where \( (V, W) = \dim_k \text{Hom}_A(V, W) \) for \( V, W \in \text{Rep}_{\text{fin}} A \); see (2.16).

**Theorem 2.24.** Let \( A \in \text{Alg}_k \) be Frobenius and let \( S, S' \in \text{Irr} \ A \). Then

\[
(PS, PS') = (PS', PS).
\]
In particular, if $A$ is symmetric and $\mathbb{k}$ is a splitting field for $A$, then the Cartan matrix of $A$ is symmetric.

**Proof.** Note that $P^\ast = \nu P \nu$. If $A$ is symmetric, then $\nu P \equiv \nu A \nu$; so it suffices to prove the first assertion. Putting $P = PS$ and $P' = PS'$, we need to show that $(P', P') = (P', P)$. Fix a composition series $0 = V_0 \subset V_1 \subset \cdots \subset V_i = P'$ for $P'$ and put $V_i = V_i/V_{i-1} \in \text{Irr} A$. Since $P$ is projective, $(P', \nu P)$ is additive on short exact sequences in $\text{Rep}_{\text{fin}} A$. Therefore,

$$(P', P') = \sum_i (P', V_i) = \sum_{i; V_i \neq S} \dim_k D(S) = \mu(S, P') \dim_k D(S).$$

Next, $\nu P$ is injective (Proposition 2.22) and $\text{soc} \nu P \equiv \nu (\text{soc} P) \equiv S$ (Proposition 2.23). Therefore, $(\nu P, \nu P)$ is additive on short exact sequences in $\text{Rep}_{\text{fin}} A$ and $(S', \nu P) = \delta_{S', S} \dim_k D(S)$ for $S' \in \text{Irr} A$. The following calculation now gives the desired equality:

$$(P', P') = \sum_i (V_i, \nu P) = \sum_{i; V_i \neq S} \dim_k D(S) = \mu(S, P') \dim_k D(S). \quad \Box$$

**Exercises for Section 2.2**

2.2.1 (Direct and tensor products, matrix rings, corners). Prove:

(a) The direct product $A \times B$ is Frobenius if and only if both $A$ and $B$ are Frobenius; likewise for symmetric. In this case, $\Gamma(A \times B) = \Gamma A \times \Gamma B$.

(b) If $(A, \lambda)$ and $(B, \mu)$ are Frobenius, then so is $(A \otimes B, \lambda \otimes \mu)$; similarly for symmetric. Furthermore, $\Gamma(A \otimes B) = \Gamma A \otimes \Gamma B$.

(c) If $(A, \lambda)$ is Frobenius (or symmetric), then so is $(\text{Mat}_n(A), \lambda_n)$, with $\langle \lambda_n, (a_{i,j}) \rangle := \sum_{i,j} \langle \lambda, a_{i,j} \rangle$, and $\Gamma \text{Mat}_n(A) = (\Gamma A) 1_{n \times n}$.

(d) Let $(A, \lambda)$ be symmetric, with $\lambda \in A^{\ast \text{trace}}$, and let $0 \neq e = e^2 \in A$. Then $(e A e, \lambda_{e A e})$ is also symmetric.

2.2.2 (Center and twisted trace forms). Let $(A, \lambda)$ be a Frobenius algebra. Show that the isomorphism $\lambda - \cdot : A \rightarrow A^{\ast}$ in $\text{Mod}_A$ restricts to an isomorphism

$${\mathcal Z} A \rightarrow \{ f \in A^{\ast} \mid f - a = \nu_A(a) - f \text{ for all } a \in A \}$$

in $\text{Vect}_k$. In particular, if $A$ is symmetric, then $\mathcal{Z} A \cong A^{\ast \text{trace}}$.

2.2.3 (Sweedler algebra). Assume that $\text{char} \mathbb{k} \neq 2$ and let $A$ be the Sweedler $\mathbb{k}$-algebra (Exercise 2.1.13) Thus, $A$ has $\mathbb{k}$-basis $1, x, y, xy$ and multiplication is given by $x^2 = 0, y^2 = 1$ and $xy = -yx$. Define $\lambda \in A^{\ast}$ by $\langle \lambda, x \rangle = 1$ and $\langle \lambda, 1 \rangle = \langle \lambda, y \rangle = \langle \lambda, xy \rangle = 0$. Show:

(a) $(A, \lambda)$ is Frobenius, but $A$ is not symmetric.

(b) The Nakayama automorphism $\nu_\lambda$ is given by $\nu_\lambda(x) = x, \nu_\lambda(y) = -y$. 


(c) The Casimir operator $\gamma_{\lambda\lambda}$ vanishes and the Higman trace $\gamma_{\lambda\lambda}$ is given by $\gamma_{\lambda\lambda}(1) = 4x$ and $\gamma_{\lambda\lambda}(x) = \gamma_{\lambda\lambda}(y) = \gamma_{\lambda\lambda}(xy) = 0$.

**2.2.4 (Socle series of a Frobenius algebra)**. Let $(A, \lambda)$ be a Frobenius algebra. For any ideal $I$ of $A$, put $l.\text{ann}_A I = \{ a \in A \mid aI = 0 \}$ and $r.\text{ann}_A I = \{ a \in A \mid Ia = 0 \}$. The left and right socle series of any finite-dimensional $A \in \text{Alg}_k$ are defined by $l.\text{soc}_n A := l.\text{ann}_A (\text{rad} A)^n$ and $r.\text{soc}_n A := r.\text{ann}_A (\text{rad} A)^n$. Show:

(a) $r.\text{ann}_A I = l.\text{ann}_{\lambda(I)} A$.
(b) $l.\text{soc}_n A = r.\text{soc}_n A$ for all $n$.

**2.2.5 (Casimir operator and Higman trace)**. Let $(A, \lambda)$ be a Frobenius algebra. Show:

(a) $\nu_{\lambda} \circ \gamma_{\lambda\lambda} = \gamma_{\lambda\lambda} \circ \nu_{\lambda}$ and $\nu_{\lambda} \circ \gamma_{\lambda\lambda} = \gamma_{\lambda\lambda} \circ \nu_{\lambda}$.
(b) $\gamma_{\lambda\lambda}(a) = \gamma_{\lambda\lambda}(a)\gamma_{\lambda\lambda}(1)$ and $\gamma_{\lambda\lambda}(a) = \gamma_{\lambda\lambda}(a)\gamma_{\lambda\lambda}(1)$.
(c) The identity $\text{trace}(a_{A} \circ_{A} b) = \text{trace}(b_{A} \circ_{A} a)$ holds for all $a, b \in A$ if and only if $\nu_{\lambda} \circ \gamma_{\lambda\lambda} = \gamma_{\lambda\lambda}$. The identity holds if $A$ is symmetric.

**2.2.6 (Frobenius extensions)**. Let $\phi: B \rightarrow A$ be a $\mathbb{k}$-algebra map and view $A$ as $(B, B)$-bimodule via $\phi$ ($\S 1.2.2$). The extension $A/B$ is called a Frobenius extension if there exists an $(B, B)$-bimodule map $E: A \rightarrow B$ and elements $(x_i, y_i)$ of $A$ such that $a = y_i.E(x_i.a) = E(a.y_i).x_i$ for all $a \in A$. Thus, Frobenius algebras are the same as Frobenius extensions of $\mathbb{k}$, with any Frobenius form playing the role of $E$.

Prove:

(a) For any Frobenius extension $A/B$, there is an isomorphism of functors $\text{Coind}_B \cong \text{Ind}_B^A$; it is given by $\text{Coind}_B W \rightarrow \text{Ind}_B^A W, f \mapsto y_i \otimes f(x_i)$ for $W \in \text{Rep} B$, with inverse given by $a \otimes w \mapsto (a' \mapsto E(a'.a).w)$.
(b) Conversely, if $\text{Coind}_B \cong \text{Ind}_B^A$, then $A/B$ is a Frobenius extension.

**2.2.7 (Projectives and injectives)**. Let $A$ be a Frobenius algebra. Use the isomorphism of functors $\text{Coind}^A_k \cong \text{Ind}^A_k$ (Exercise 2.2.6) and Exercise 2.1.1(b),(c) to show that arbitrary $A$-modules are projective if and only if they are injective.
Part II

Groups
Chapter 3

Groups and Group Algebras

The theory of group representations is the archetypal representation theory, along-side the corresponding theory for Lie algebras (Part III). A representation of a group $G$, by definition, is a group homomorphism

$$G \rightarrow \text{GL}(V),$$

where $V$ is a vector space over a field $\mathbb{k}$ and $\text{GL}(V) = \text{Aut}_\mathbb{k}(V)$ denotes the group of all $\mathbb{k}$-linear automorphisms of $V$. More precisely, such a representation is called a linear representation of $G$ over $\mathbb{k}$. Not being part of the defining data of $G$, the base field $\mathbb{k}$ can be chosen depending on the purpose at hand. The traditional choice, especially for representations of finite groups, is the field $\mathbb{C}$ of complex numbers; such representations are called complex representations of $G$. One encounters a very different flavor of representation theory when the characteristic of $\mathbb{k}$ divides the order of $G$. Representations of this kind are referred to as modular representations of $G$. Our main focus in this chapter will be on non-modular representations of finite groups.

*Throughout this chapter, $\mathbb{k}$ denotes a field and $G$ is a group, not necessarily finite and generally in multiplicative notation. All further hypotheses will be spelled out when they are needed.*

3.1. Generalities

This section lays the foundations of the theory of group representations by placing it in the framework of representations of algebras; this is achieved by means of the group algebra of $G$ over $\mathbb{k}$.
3.1.1. Group Algebras

As a \(k\)-vector space, the group algebra of the group \(G\) over \(k\) is the \(k\)-vector space \(kG\) of all formal \(k\)-linear combinations of the elements of \(G\) (Example A.5). Thus, elements of \(kG\) have the form \(\sum_{x \in G} \alpha_x x\) with unique \(\alpha_x \in k\) that are almost all 0. The multiplication of \(k\) and the (multiplicative) group operation of \(G\) give rise to a multiplication for \(kG\):

\[
\left( \sum_{x \in G} \alpha_x x \right) \left( \sum_{y \in G} \beta_y y \right) \overset{\text{def}}{=} \sum_{x,y \in G} \alpha_x \beta_y xy = \sum_{z \in G} \left( \sum_{x,y \in G \atop xy = z} \alpha_x \beta_y \right) z.
\]

It is a routine matter to check that this yields an associative \(k\)-algebra structure on \(kG\) with unit map \(k \to kG, \lambda \mapsto \lambda 1_G\), where \(1_G\) is the identity element of the group \(G\). The group algebra \(kG\) will also be denoted by \(k[G]\) in cases where the group in question requires more complex notation.

Universal Property. Despite being simple and natural, it may not be immediately apparent why the above definition should be worthy of our attention. The principal reason is provided by the following “universal property” of group algebras. For any \(k\)-algebra \(A\), let \(A^\times\) denote its group of units, that is, the group of invertible elements of \(A\). Then there is a natural bijection

\[
\text{Hom}_{\text{Alg}_k}(kG, A) \cong \text{Hom}_{\text{Groups}}(G, A^\times).
\]

The bijection is given by sending an algebra map \(f : kG \to A\) to its restriction \(f|_G\) to the basis \(G \subseteq kG\) as in (A.4). Observe that each element of \(G\) is a unit of \(kG\) and that \(f|_G\) is indeed a group homomorphism \(G \to A^\times\). Conversely, if \(G \to A^\times\) is any group homomorphism, then its unique \(k\)-linear extension from \(G\) to \(kG\) is in fact a \(k\)-algebra map.

Functoriality. Associating to a given \(k\)-algebra \(A\) its group of units, \(A^\times\), is a “functorial” process: any algebra map \(A \to B\) restricts to a group homomorphism \(A^\times \to B^\times\). The usual requirements on functors with respect to identities and composites are clearly satisfied; so we obtain a functor

\[
\cdot^\times : \text{Alg}_k \to \text{Groups}.
\]

Similar things can be said for the rule that associates to a given group \(G\) its group algebra \(kG\). Indeed, we have already observed above that \(G\) is a subgroup of the group of units \((kG)^\times\). Thus, if \(f : G \to H\) is a group homomorphism, then the composite of \(f\) with the inclusion \(H \hookrightarrow (kH)^\times\) is a group homomorphism \(G \to (kH)^\times\). By (3.2) there is a unique algebra homomorphism \(k f : kG \to kH\)
such that the following diagram commutes:

\[
\begin{array}{ccc}
G & \xrightarrow{f} & H \\
\downarrow & & \downarrow \\
\k G & \xrightarrow{\exists ! \ k f} & \k H
\end{array}
\]

It is straightforward to ascertain that \( k \cdot \) respects identity maps and composites as is required for a functor, and hence we again have a functor

\[ k : \text{Groups} \to \text{Alg}_k. \]

Finally, it is routine to verify that the bijection (3.2) is functorial in both \( G \) and \( A \): the group algebra functor \( k \cdot \) is left adjoint to the unit group functor \( \cdot \times \) in the sense of Section A.4.

### 3.1.2. First Examples and Some Variants

Having addressed the basic formal aspects of group algebras in general, let us now look at two explicit examples of group algebras and describe their algebra structure.

**Example 3.1** (The group algebra of a lattice). Abelian groups isomorphic to some \( \mathbb{Z}^n \) are often referred to as lattices. While it is desirable to keep the natural additive notation of \( \mathbb{Z}^n \), the group law of \( \mathbb{Z}^n \) becomes multiplication in the group algebra \( k[\mathbb{Z}^n] \). In order to resolve this conflict, we will denote an element \( m \in \mathbb{Z}^n \) by \( x^m \) when thinking of it as an element of the group algebra \( k[\mathbb{Z}^n] \). This results in the following rule, which governs the multiplication of \( k[\mathbb{Z}^n] \):

\[
x^{m+m'} = x^m x^{m'} \quad (m, m' \in \mathbb{Z}^n).
\]

Fixing a \( \mathbb{Z} \)-basis \( (e_i)_{i=1}^n \) of \( \mathbb{Z}^n \) and putting \( x_i = x^{e_i} \), each \( x^m \) takes the form

\[
x^m = x_1^{m_1} x_2^{m_2} \ldots x_n^{m_n}
\]

with unique \( m_i \in \mathbb{Z} \). Thus, \( x^m \) can be thought of as a monomial in the \( x_i \) and the group algebra \( k[\mathbb{Z}^n] \) is isomorphic to a Laurent polynomial algebra over \( k \),

\[
k[\mathbb{Z}^n] \cong k[x_1^{\pm 1}, x_2^{\pm 1}, \ldots, x_n^{\pm 1}].
\]

The monomials \( x^m \) with \( m \in \mathbb{Z}_+^n \) span a subalgebra of \( k[\mathbb{Z}^n] \) that is isomorphic to the ordinary polynomial algebra \( k[x_1, x_2, \ldots, x_n] \).

**Example 3.2** (Group algebras of finite abelian groups). Now let \( G \) be finite abelian. Then, for suitable positive integers \( n_i \),

\[
G \cong C_{n_1} \times C_{n_2} \times \cdots \times C_{n_t},
\]

where \( C_n \) denotes the cyclic group of order \( n \). Sending a fixed generator of \( C_n \) to the variable \( x \) gives an isomorphism of algebras \( kC_n \cong k[x]/(x^n - 1) \). Moreover,
the isomorphism (3.3) yields an isomorphism \( kG \cong \bigotimes_{i=1}^{t} k[x]/(x^{n_i} - 1) \).

By field theory, this algebra is a direct product of fields if and only if \( \text{char } k \) does not divide any of the integers \( n_i \) or, equivalently, \( \text{char } k \nmid |G| \). This is a very special case of Maschke’s Theorem (§3.4.1).

Monoid Rings. The definition of the product in (3.1) makes sense with any ring \( R \) in place of \( k \), resulting in the group ring \( RG \) of \( G \) over \( R \). The case \( R = \mathbb{Z} \) will play a role in Sections 8.5 and 10.3. We can also start with an arbitrary monoid \( \Gamma \) rather than a group and obtain the monoid algebra \( k\Gamma \) or the monoid ring \( R\Gamma \) in this way. Finally, (3.1) is also meaningful with possibly infinitely many \( \alpha_x \) or \( \beta_y \) being nonzero, provided the monoid \( \Gamma \) satisfies the following condition:

\[
\{(x, y) \in \Gamma \times \Gamma \mid xy = z\} \text{ is finite for each } z \in \Gamma.
\]

In this case, (3.1) defines a multiplication on the \( R \)-module \( R\Gamma \) of all functions \( \Gamma \to R \), not just on the submodule \( R\Gamma = R^{(\Gamma)} \) of all finitely supported functions. The resulting ring is called the total monoid ring of \( \Gamma \) over \( R \).

Example 3.3 (Power series). Let \( \Gamma \) be the (additive) monoid \( \mathbb{Z}(I) \) of all finitely supported functions \( n : I \to \mathbb{Z} \) for some set \( I \), with pointwise addition: \((n + m)(i) = n(i) + m(i)\). Then, as in Example 3.1, the resulting monoid ring \( R\Gamma \) is isomorphic to the polynomial ring \( R[x_i \mid i \in I] \). Condition (3.4) is easily seen to be satisfied for \( \Gamma \). The total monoid ring is \( R[x_i \mid i \in I] \), the ring of formal power series in the commuting variables \( x_i \) over \( R \).

3.1.3. Representations of Groups and Group Algebras

Recall that a representation of the group \( G \) over \( k \), by definition, is a group homomorphism \( G \to GL(V) = \text{End}_k(V)^\times \) for some \( V \in \text{Vect}_k \). The adjoint functor relation (3.2) gives a natural bijection, \( \text{Hom}_{\text{Alg}_k}(kG, \text{End}_k(V)) \cong \text{Hom}_{\text{Groups}}(G, GL(V)) \). Thus, representations of \( G \) over \( k \) are in natural 1-1 correspondence with representations of the group algebra \( kG \):

\[
\text{representations of } kG \cong \text{representations of } G \text{ over } k.
\]

This observation makes the material of Chapter 1 available for the treatment of group representations. In particular, we may view the representations of \( G \) over \( k \) as a category that is equivalent to \( \text{Rep } kG \) (or \( kG\text{-Mod} \)) and we may speak of homomorphisms and equivalence of group representations as well as of irreducibility, composition series etc. by employing the corresponding notions from Chapter 1.
We will also continue to use the notation of Chapter 1, writing the group homomorphism $G \to \text{GL}(V)$ as $g \mapsto g_v$ and putting $g.v = g_v(v)$ for $v \in V$.

Many interesting irreducible representations for specific groups $G$ will be discussed in detail later on. For now, we just mention one that always exists, for every group, although it does not appear to be very exciting at first sight.¹ This is the so-called **trivial representation**, $\mathbb{1}$, which arises from the trivial group homomorphism $G \to \text{GL}(k) = k^\times$:

$$\mathbb{1} : G \longrightarrow k^\times$$

(3.5)

$$g \longmapsto 1_k$$

**A Word on the Base Field.** In the context of group representations, the base field $k$ is often understood and is notationally suppressed. Thus, for example, $\text{Hom}_{kG}(V, W)$ is frequently written as $\text{Hom}_G(V, W)$ in the literature. Generally (except in Chapter 4), we will write $kG$, acknowledging the base field. We will however say that $k$ is a **splitting field** for $G$ rather than for $kG$. Recall that this means that $D(S) = k$ for all $S \in \text{Irr}_{\text{fin}}kG$ (§1.2.5). Thus, an algebraically closed field $k$ is a splitting field for every group $G$ by Schur’s Lemma. Much less than algebraic closure suffices for finite groups; see Exercise 3.1.5 and also Corollary 4.16.

### 3.1.4. Changing the Group

We have seen that each group homomorphism $H \to G$ lifts uniquely to an algebra map $kH \to kG$. Therefore, we have the restriction (or pullback) functor from $kG$-representations to $kH$-representations and, in the other direction, the induction and coinduction functors (§1.2.2). In the context of finite groups, which will be our main focus, we may concentrate on induction, because the induction and coinduction functors are isomorphic by Proposition 3.4 below. When $H$ is a subgroup of $G$, all of the following notations are commonly used:

$$\downarrow_H = \downarrow^G_H = \text{Res}^G_H = \text{Res}^{kG}_{kH} : \text{Rep}kG \longrightarrow \text{Rep}kH$$

and

$$\uparrow^G = \uparrow_H^G = \text{Ind}^G_H = \text{Ind}^{kG}_{kH} : \text{Rep}kH \longrightarrow \text{Rep}kG.$$  

In this chapter and the next, we will predominantly use the up and down arrows, as they are especially economical and suggestive.²

In the language of Exercise 2.2.6, part (b) of the following proposition states that if $H$ is a finite-index subgroup of $G$, then $kH \hookrightarrow kG$ is a Frobenius extension.

---

¹Nonetheless, $\mathbb{1}$ turns out to have great significance. For example, if $\mathbb{1}$ is projective as $kG$-module, then all $kG$-modules are projective; see the proof of Maschke’s Theorem (§3.4.1).

²Coinduction is often denoted by $\downarrow_H = \downarrow_H^G : \text{Rep}kH \to \text{Rep}kG$ in the literature.
We let \( G/H \) denote the collection of all left cosets \( gH \) (\( g \in G \)) and also a transversal for these cosets.

**Proposition 3.4.** Let \( H \) be a subgroup of \( G \) and let \( W \in \text{Rep} \mathbb{k}H \). Then:

(a) \( W \uparrow^G = \bigoplus_{g \in G/H} gW' \) for some subrepresentation \( W' \subseteq W \uparrow^G \downarrow_H \) with \( W' \equiv W \). In particular, \( \dim_k W \uparrow^G = [G : H] \dim_k W \).

(b) If \( [G : H] < \infty \), then there is a natural isomorphism \( \text{Coind}_{\mathbb{k}H}^{k} W \cong \text{Ind}_{\mathbb{k}H}^{G} W \) in \( \text{Rep} \mathbb{k}G \).

**Proof.** (a) The crucial observation is that \( \mathbb{k}G \) is free as a right \( \mathbb{k}H \)-module: the partition \( G = \bigsqcup_{g \in G/H} gH \) yields the decomposition \( \mathbb{k}G = \bigoplus_{g \in G/H} g \mathbb{k}H \). By \( (B.5) \), it follows that the elements of \( W \uparrow^G = \mathbb{k}G \otimes_{\mathbb{k}H} W \) have the form

\[
\sum_{g \in G/H} g \otimes w_g
\]

with unique \( w_g \in W \). The map \( \mu: W \to W \uparrow^G \downarrow_H \) \( w \mapsto 1 \otimes w \), is a morphism in \( \text{Rep} \mathbb{k}H \) and \( \mu \) is injective, since we may choose the transversal so that \( 1 \in G/H \).

Putting \( W' = \text{Im} \mu \), we have \( W' \equiv W \) and \( gW' = \{ g \otimes w \mid w \in W \} \). Thus, the above form of elements of \( W \uparrow^G \) also implies the remaining assertions of (a).

(b) Consider the following projection of \( \mathbb{k}G \) onto \( \mathbb{k}H \):

\[
\pi_H: \begin{array}{ccc}
\mathbb{k}G & \overset{\psi}{\rightarrow} & \mathbb{k}H \\
\downarrow & & \downarrow \\
\sum_{x \in G} x \otimes x & \rightarrow & \sum_{x \in H} x \otimes x
\end{array}
\]

Thus, \( \pi_H \) is the identity on \( \mathbb{k}H \) and it is easy to see that \( \pi_H \) is a \( (\mathbb{k}H, \mathbb{k}H) \)-bimodule map. Moreover, the following identities holds for every \( a \in \mathbb{k}G \):

\[
a = \sum_{g \in G/H} \pi_H(ag)g^{-1} = \sum_{g \in G/H} g\pi_H(g^{-1}a).
\]

By linearity, it suffices to check the equalities for \( a \in G \), in which case they are immediate. The map \( \pi_H \) leads to the following map in \( \text{Rep} \mathbb{k}H \):

\[
\phi: \begin{array}{ccc}
W & \overset{\sim}{\rightarrow} & \text{Coind}_{\mathbb{k}H}^{G} W \overset{\pi_H}{\rightarrow} \text{Ind}_{\mathbb{k}H}^{G} W \downarrow_H \\
\downarrow & & \downarrow & & \downarrow \\
w & \rightarrow & (b \mapsto b.w) & \rightarrow & (a \mapsto \pi_H(a).w)
\end{array}
\]

By Proposition 1.9(a), \( \phi \) gives rise to the map \( \Phi: \text{Ind}_{\mathbb{k}H}^{G} W \to \text{Coind}_{\mathbb{k}H}^{G} W \) in \( \text{Rep} \mathbb{k}G \), with \( \Phi(a \otimes w) = a.\phi(w) \) for \( a \in \mathbb{k}G \) and \( w \in W \). In the opposite direction, define \( \Psi: \text{Coind}_{\mathbb{k}H}^{G} W \to \text{Ind}_{\mathbb{k}H}^{G} W \) by \( \Psi(f) = \sum_{g \in G/H} g \otimes f(g^{-1}) \). Using (3.7), one verifies without difficulty that \( \Phi \) and \( \Psi \) are inverse to each other. \( \square \)
**Frobenius Reciprocity.** Let $H$ be a subgroup of $G$. Then Proposition 1.9 states that, for any $W \in \text{Rep}_k H$ and $V \in \text{Rep}_k G$, we have a natural isomorphism

\[(3.8) \quad \text{Hom}_k G (W^G, V) \cong \text{Hom}_H (W, V_H)\]

in $\text{Vect}_k$. When $[G : H] < \infty$, then Propositions 1.9 and 3.4 together give

\[(3.9) \quad \text{Hom}_k G (V, W^G) \cong \text{Hom}_H (V, W^H)\]

Explicitly, $\text{Hom}_H (V, W) \ni f \leftrightarrow \sum_{g \in G/H} g \otimes f(g^{-1} \cdot ) \in \text{Hom}_k G (V, W^G)$. Both (3.8) and (3.9) are referred to as Frobenius reciprocity isomorphisms.\(^3\)

**Degree Bounds.** We conclude our first foray into the categorical aspects of group representations by giving some down-to-earth applications to the degrees of irreducible representations. The argument in the proof of part (b) below can be used to similar effect in the more general context of cofinite subalgebras (Exercise 1.2.7).

**Corollary 3.5.** Let $H$ be a subgroup of $G$. Then:

(a) Every $W \in \text{Irr}_k H$ is a subrepresentation of $V_H$ for some $V \in \text{Irr}_k G$. Thus, any upper bound for the degrees of irreducible representations of $kG$ is also an upper bound for the irreducible representations of $kH$.

(b) Assume that $[G : H] < \infty$. Then every $V \in \text{Irr}_k G$ is a subrepresentation of $W^G$ for some $W \in \text{Irr}_k H$. Consequently, if the degrees of irreducible representations of $kH$ are bounded above by $d$, then the irreducible representations of $kG$ have degrees at most $[G : H]d$.

**Proof.** (a) We know by Proposition 3.4(a) that $W^G$ is a cyclic $kG$-module, because $W$ is cyclic. Hence there is an epimorphism $W^G \twoheadrightarrow V$ for some $V \in \text{Irr}_k G$ (Exercise 1.1.3). By (3.8), this epimorphism corresponds to a nonzero map of $kH$-representations $W \twoheadrightarrow V_H$, which must be injective by irreducibility of $W$. This proves the first assertion of (a). The statement about degree bounds is clear.

(b) The restriction $V_H$ is finitely generated as $kH$-module, because $V$ is cyclic and $kG$ is finitely generated as left $kH$-module by virtue of our hypothesis on $[G : H]$. Therefore, there is an epimorphism $V_H \twoheadrightarrow W$ for some $W \in \text{Irr}_k H$, and this corresponds to a nonzero map $V \twoheadrightarrow W^G$ by (3.9). By irreducibility of $V$, the latter map must be injective, proving the first assertion of (b). The degree bound is now a consequence of the dimension formula in Proposition 3.4(a).

\(^3\) (3.8) and (3.9) are also known as the Nakayama relations.
3.1.5. Characters of Finite-Dimensional Group Representations

For any $V \in \text{Rep}_\text{fin} \mathbb{k}G$, we have the associated character,

$$\chi_V \in (\mathbb{k}G)^*_\text{trace} \subseteq (\mathbb{k}G)^*.$$  

Here, $(\mathbb{k}G)^*$ is the space of linear forms on $\mathbb{k}G$ and $(\mathbb{k}G)^*_\text{trace} \equiv (\mathbb{k}G/[\mathbb{k}G, \mathbb{k}G])^*$ is the subspace of all trace forms as in (1.54). Linear forms on $\mathbb{k}G$ can be identified with functions $G \to \mathbb{k}$ by (A.4). In particular, each character $\chi_V$ and, more generally, each trace form on $\mathbb{k}G$ can be thought of as a $\mathbb{k}$-valued function on $G$. Below, we shall describe the kind of functions arising in this manner.

Relations with Conjugacy Classes. The group $G$ acts on itself by conjugation,

$$G \times G \to G,$$

(3.10)

$$(x, y) \mapsto \chi_y := xyx^{-1}$$

The orbits in $G$ under this action are called the conjugacy classes of $G$; the conjugacy class of $x \in G$ will be denoted by $G_x$. Using bilinearity of the Lie commutator $[\cdot, \cdot]$, one computes

$$[\mathbb{k}G, \mathbb{k}G] = \sum_{x, y \in G} \mathbb{k}(xy - yx) = \sum_{x, y \in G} \mathbb{k}(xy - y^x(xy)) = \sum_{x, y \in G} \mathbb{k}(x - yx).$$

Thus, a linear form $\phi \in (\mathbb{k}G)^*$ is a trace form if and only if $\phi(x) = \phi(yx)$ for all $x, y \in G$, that is, $\phi$ is constant on all conjugacy classes of $G$. Functions $G \to \mathbb{k}$ that are constant on conjugacy classes are called ($\mathbb{k}$-valued) class functions; so we will think of characters $\chi_V$ as class functions on $G$. To summarize, we have the following commutative diagram in $\text{Vect}_\mathbb{k}$:

$$\begin{array}{cccc}
(\mathbb{k}G)^* & \sim & \{ \text{functions } G \to \mathbb{k} \} \\
\cup & & \cup \\
(\mathbb{k}G)^*_{\text{trace}} & \sim & \text{cf}_\mathbb{k}(G) \overset{\text{def}}{=} \{ \text{class functions } G \to \mathbb{k} \}
\end{array}$$  

(3.11)

Proposition 3.6. $\# \text{Irr}_{\text{fin}} \mathbb{k}G \leq \# \{ \text{conjugacy classes of } G \}$.

Proof. We have an obvious isomorphism $\text{cf}_\mathbb{k}(G) \equiv \mathbb{k}\mathcal{C}$ in $\text{Vect}_\mathbb{k}$, where $\mathcal{C} = \mathcal{C}(G)$ denotes the set of all conjugacy classes of $G$ and $\mathbb{k}\mathcal{C}$ is the vector space of all functions $\mathcal{C} \to \mathbb{k}$. For the proposition, we may assume that $\mathcal{C}$ is finite; so $\dim_\mathbb{k} \text{cf}_\mathbb{k}(G) = \# \mathcal{C}$. Since $\# \text{Irr}_{\text{fin}} \mathbb{k}G \leq \dim_\mathbb{k} \mathbb{C}(\mathbb{k}G)$ by Theorem 1.44 and $\mathbb{C}(\mathbb{k}G)$ is a subspace of $(\mathbb{k}G)^*_\text{trace} \equiv \text{cf}_\mathbb{k}(G)$, the bound for $\# \text{Irr}_{\text{fin}} \mathbb{k}G$ follows. \qed
We will apply the proposition to finite groups only, but we remark that there are infinite groups with finitely many conjugacy classes; in fact, every torsion free group embeds into a group with exactly two conjugacy classes [176, Exercise 11.78]. If the set $\mathcal{C}(G)$ of conjugacy classes of $G$ is finite, then the foregoing allows us to identify the space of class functions $cf_\kappa(G) \cong \kappa^{\mathcal{C}(G)}$ with its dual, $\text{Tr} \kappa G = \kappa G/[\kappa G, \kappa G]$, by means of the isomorphism

$$\text{Tr} \kappa G \overset{\sim}{\longrightarrow} cf_\kappa(G)$$

$$\sum_{x \in G} a_x x + [\kappa G, \kappa G] \longmapsto \left( \sum_{x \in C} a_x \right)_{C \in \mathcal{C}(G)}$$

(3.12)

**Character Tables.** Important representation theoretic information for a given finite group $G$ is recorded in the **character table** of $G$ over $\kappa$. By definition, this is the matrix whose $(i, j)$-entry is $\chi_i(g_j)$, where $\{\chi_i\}$ are the irreducible characters of $\kappa G$ in some order, traditionally starting with the trivial character, $\chi_1 = 1$, and $\{g_j\}$ is a set of representatives for the conjugacy classes of $G$, generally with $g_1 = 1$. Thus, the first column of the character table gives the degrees, viewed in $\kappa$, of the various irreducible representations of $G$ over $\kappa$. Usually, the sizes of the conjugacy classes of $G$ are also indicated in the character table and other information may be included as well.

<table>
<thead>
<tr>
<th>classes</th>
<th>$1$</th>
<th>$g_j$</th>
<th>$\ldots$</th>
</tr>
</thead>
<tbody>
<tr>
<td>sizes</td>
<td>$1$</td>
<td>$\lvert G g_j \rvert$</td>
<td>$\ldots$</td>
</tr>
<tr>
<td>$\chi_i = \chi S_i$</td>
<td>$(\dim \kappa S_i) 1_\kappa$</td>
<td>$\ldots$</td>
<td>$\chi_i(g_j)$</td>
</tr>
<tr>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
</tr>
</tbody>
</table>

**Conjugacy Classes of $p$-Regular Elements.** The only substantial result on modular group representations that we shall offer is the theorem below, which is due to Brauer [31]. To state it, let $p$ be a prime number. An element $g \in G$ is called **$p$-regular** if its order is not divisible by $p$. Since conjugate elements have the same order, we may also speak of **$p$-regular conjugacy classes** of $G$.

**Theorem 3.7.** Let $G$ be a finite group and assume that $\text{char} \kappa = p > 0$. Then $\# \text{Irr} \kappa G \leq \# \{ p\text{-regular conjugacy classes of } G \}$. Equality holds if $\kappa$ is a splitting field for $G$.

**Proof.** By Proposition 1.46 and Lemma 1.48, it suffices to consider the case where $\kappa$ is a splitting field. Each $g \in G$ can be uniquely written as $g = g_p g'_p = g_p g'_p$,
where \( g_{p'} \in G \) is \( p \)-regular and the order of \( g_p \) is a power of \( p \).\(^4\) We will prove that the isomorphism \( (\mathbb{k}G)^* \cong \text{cf}_\mathbb{k}(G) \) of (3.11) restricts to an isomorphism
\[
C(\mathbb{k}G) \cong \left\{ f \in \text{cf}_\mathbb{k}(G) \mid f(g) = f(g_{p'}) \text{ for all } g \in G \right\}.
\]
and show that \( \dim_\mathbb{k} C(\mathbb{k}G) \) is equal to the number of \( p \)-regular conjugacy classes of \( G \). In view of Theorem 1.44, this will imply the statement about \( \# \text{Irr} \mathbb{k}G \).

Recall that the space \([\mathbb{k}G, \mathbb{k}G]\) of Lie commutators is stable under the \( p \)-th power map and that \((a + b)^p \equiv a^p + b^p \mod [\mathbb{k}G, \mathbb{k}G] \) for all \( a, b \in \mathbb{k}G \) (Lemma 1.42). By Proposition 1.43, \( C(\mathbb{k}G) \) consists of exactly those linear forms on \( \mathbb{k}G \) that vanish on the subspace \( T = \{ a \in \mathbb{k}G \mid a^q \in [\mathbb{k}G, \mathbb{k}G] \text{ for some } q = p^n, n \in \mathbb{Z}_+ \} \).

Write a given \( g \in G \) as \( g = g_{p'} g_p \) as above, with \( g_p^q = 1 \) for \( q = p^n \). Then \((g - g_{p'})^q = g^q - g_{p'}^q = 0\); so \( g \equiv g_{p'} \mod T \). Thus, fixing representatives \( x_i \) for the \( p \)-regular conjugacy classes of \( G \), the family \( (x_i + T) \) spans \( \mathbb{k}G/T \) and it suffices to prove linear independence. So assume that \( \sum_i \lambda_i x_i \in T \) with \( \lambda_i \in \mathbb{k} \). Then \((\sum_i \lambda_i x_i)^q \in [\mathbb{k}G, \mathbb{k}G] \) for all sufficiently large \( q = p^n \). Writing \(|G| = p^m \) with \((p, m) = 1\) and choosing a large \( n \) so that \( q \equiv 1 \mod m \), we also have \( x_i^q = x_i \) for all \( i \). Since the \( p \)-th power map yields an additive endomorphism of \([\mathbb{k}G]/[\mathbb{k}G, \mathbb{k}G]\), we obtain
\[
0 \equiv \left( \sum_i \lambda_i x_i \right)^q = \sum_i \lambda_i x_i^q = \sum_i \lambda_i^q x_i \mod [\mathbb{k}G, \mathbb{k}G].
\]
Finally, non-conjugate elements of \( G \) are linearly independent modulo \([\mathbb{k}G, \mathbb{k}G]\) by (3.12), whence \( \lambda_i^q = 0 \) for all \( i \) and so \( \lambda_i = 0 \) as desired. \( \Box \)

The result remains true for \( \text{char} \mathbb{k} = 0 \) with the understanding that all conjugacy classes of a finite group \( G \) are 0-regular. Indeed, by Proposition 3.6, we already know that \( \# \text{Irr} \mathbb{k}G \) is bounded above by the number of conjugacy classes of \( G \), and we shall see in Corollary 3.21 that equality holds if \( \mathbb{k} \) is a splitting field for \( G \).

### 3.1.6. Finite Group Algebras as Symmetric Algebras

We mention here in passing that group algebras of finite groups are symmetric algebras. As we shall see later (§3.6.1), certain properties of finite group algebras are in fact most conveniently derived in this more general ring theoretic setting.

In detail, the map \( \pi_1 \) in (3.6) is a trace form for any group algebra \( \mathbb{k}G \), even if \( G \) is infinite. Denoting \( \pi_1 \) by \( \lambda \) as in Section 2.2 and using \( \langle \cdot, \cdot \rangle \) for evaluation of linear forms, we have
\[
\langle \lambda, \sum_{x \in G} \alpha_x x \rangle = \alpha_1.
\]
If \( 0 \neq a = \sum_{x \in G} \alpha_x x \in \mathbb{k}G \), then \( \langle \lambda, ax^{-1} \rangle = \alpha_x \neq 0 \) for some \( x \in G \). Thus, if the group \( G \) is finite, then \( \lambda \) is a Frobenius form for \( \mathbb{k}G \). Since \( \langle \lambda, xy \rangle = \delta_{x,y}^{-1} \) for

\(^4\)If \( G \) has order \( p^n m \) with \( p \nmid m \), then write \( 1 = p^n a + mb \) with \( a, b \in \mathbb{Z} \) and put \( g_{p'} = g^{p^n a} \) and \( g_p = g^{mb} \). The factor \( g_{p'} \) is often called the \( p \)-regular part of \( g \).
x, y ∈ G, the Casimir element and Casimir trace are given by

\[ c_\lambda = \sum_{g \in G} g \otimes g^{-1} \quad \text{and} \quad \gamma_\lambda(a) = \sum_{g \in G} gag^{-1}. \]

In particular, \( \gamma_\lambda(1) = |G| \). Thus, if \( \text{char } k \nmid |G| \), then \( 1 = \gamma_\lambda(|G|^{-1} 1) \) belongs to the Higman ideal \( \Gamma(kG) \) and so \( kG \) is semisimple by Theorem 2.21. The converse is also true and is in fact easier. Later in this chapter, we will give a direct proof of both directions that is independent of the material on symmetric algebras; see Maschke’s Theorem (§3.4.1).

**Exercises for Section 3.1**

3.1.1 (Representations are functors). For a given group \( G \), let \( G \) denote the category with one object, \( * \), and with \( \text{Hom}_G(*, *) = G \). The binary operation of \( G \) is the composition in \( \text{Hom}_G(*, *) \) and the identity element of \( G \) is the identity morphism \( 1_* : * \to * \). Show:

(a) Any \( V \in \text{Rep}_kG \) gives a functor \( F_V : G \to \text{Vect}_k \) and conversely.

(b) A map \( f : V \to W \) in \( \text{Rep}_kG \) amounts to a natural transformation \( \phi : F_V \Rightarrow F_W \); and \( f \) is an isomorphism in \( \text{Rep}_kG \) if and only if \( \phi \) is an isomorphism of functors (§A.3.2).

3.1.2 (Some group algebra isomorphisms). Establish the following isomorphisms in \( \text{Alg}_k \), for arbitrary groups \( G \) and \( H \). Here, \( \otimes = \otimes_k \) as usual.

(a) \( k[G \times H] \cong kG \otimes kH \).

(b) \( k[G^{\text{op}}] \cong (kG)^{\text{op}} \). Here \( G^{\text{op}} \) is the opposite group: \( G^{\text{op}} = G \) as sets, but with new group operation \( * \) given by \( x * y = yx \).

(c) \( K \otimes kG \cong KG \) for any field extension \( K/k \). More generally, \( K \) can be any \( k \)-algebra here.

3.1.3 (Induced representations). Let \( H \) be a subgroup of \( G \).

(a) Show that \( kG \) is free as a left (and right) module over \( kH \) by multiplication: any set of right (resp., left) coset representatives for \( H \) in \( G \) provides a basis.

(b) Conclude from (a) and Exercise 1.2.1 that, for any \( W \) in \( \text{Rep}_kH \), we have \( \text{Ker}(W) \subseteq kG \text{ Ker } W \), the largest ideal of \( kG \) that is contained in the left ideal \( kG \text{ Ker } W \).

(c) Let \( W \) be a finite-dimensional representation of \( kH \). Use Proposition 3.4 to show that the character of the induced representation \( W \uparrow^G \) is given by

\[ \chi_{W \uparrow^G}(x) = \sum_{g \in G/H} \chi_W(g^{-1} x g) \]
(d) Use Proposition 3.4 to show that the induction functor \( \uparrow^G \) is exact: for any short exact sequence \( 0 \rightarrow W' \rightarrow W \rightarrow W'' \rightarrow 0 \) in \( \text{Rep}_{k}H \), the sequence \( 0 \rightarrow W' \uparrow^G \rightarrow W \uparrow^G \rightarrow W'' \uparrow^G \rightarrow 0 \) is exact in \( \text{Rep}_{k}G \).

3.1.4 (An irreducibility criterion). Let \( H \) be a subgroup of \( G \). Assume that \( V \in \text{Rep}_{k}G \) is such that \( V|_{H} = \bigoplus_{i \in I} W_{i} \) for pairwise non-isomorphic \( W_{i} \in \text{Irr}_{k}H \) such that \( kG.W_{i} = V \) for all \( i \). Show that \( V \) is irreducible.

3.1.5 (Splitting fields in positive characteristic). Let \( G \) be finite and let \( e = \exp G \) denote its exponent, that is, the least common multiple of the orders of all elements of \( G \). Let \( k \) be a field with \( \text{char } k = p > 0 \) and assume that \( k \) contains all \( e^{th} \) roots of unity in some fixed algebraic closure of \( k \). Use Exercise 1.5.8 to show that \( k \) is a splitting field for \( G \).

3.1.6 (Hattori’s Lemma). Let \( G \) be finite and let \( P \in \text{Proj}_{\text{fin}} kG \). Using the isomorphism (3.12), we may view the Hattori-Stallings rank (2.11) as a function \( \chi_{P}(g) = |C_{G}(g)| \text{rank}(P)(g^{-1}) \) for \( g \in G \).

3.2. First Examples

Thus far, we have only mentioned a single example of a group representation: the trivial representation \( 1 \) in (3.5). This section adds to our cast of characters, and we will also determine some explicit character tables.

3.2.1. Finite Abelian Groups

Let \( G \) be finite abelian. Then the group algebra \( kG \) is a finite-dimensional commutative algebra and (1.37) gives a bijection \( \text{MaxSpec } kG \sim \text{Irr } kG \)

\[
P \mapsto kG/P
\]

The Schur division algebra of the irreducible representation \( S = kG/P \) is given by \( D(S) = \text{End}_{kG}(kG/P) \cong kG/P \). Let \( e = \exp G \) denote the exponent of \( G \), that is, the smallest positive integer \( e \) such that \( x^{e} = 1 \) for all \( x \in G \). Fix an algebraic closure \( \overline{k} \) of \( k \) and put \( \mu_{e} = \{ \xi \in \overline{k} | \xi^{e} = 1 \} \), \( K = k(\mu_{e}) \subseteq \overline{k} \) and \( \Gamma = \text{Gal}(\overline{k}/k) \).

Consider the group algebra \( KG \). Every \( Q \in \text{MaxSpec } KG \) satisfies \( KG/Q \cong K \), because the images of all elements of \( G \) in the field \( KG/Q \) are \( e^{th} \) roots of unity.

\[5\]The fact stated in this exercise is also true in characteristic 0 by another result of Brauer. For a proof, see [106, Theorem 10.3] for example. (3.16) is an easy special case.
and hence they all belong to $K$. Thus,
\begin{equation}
(3.16) \quad K \text{ is a splitting field for } G.
\end{equation}

For each $P \in \text{MaxSpec } kG$, the field $kG/P$ embeds into $K$; so there is a $k$-algebra map $f: kG \to K$ with Ker $f = P$. Conversely, if $f \in H := \text{Hom}_{\text{Alg}}(kG, K)$, then Ker $f \in \text{MaxSpec } kG$. Consider the action of $\Gamma$ on $H$ that is given by $\gamma \cdot f = \gamma \circ f$ for $\gamma \in \Gamma$ and $f \in H$. By general facts about Galois actions (Exercise 1.4.6), Ker $f = \text{Ker } f'$ for $f, f' \in H$ if and only if $f$ and $f'$ belong to the same $\Gamma$-orbit. Thus, we have a bijection
\[ \text{MaxSpec } kG \leftrightarrow \Gamma \setminus \text{Hom}_{\text{Alg}}(kG, K). \]

By (3.2), we may identify $\text{Hom}_{\text{Alg}}(kG, K)$ with $\text{Hom}_{\text{Groups}}(G, K^x)$. Pointwise multiplication equips $\text{Hom}_{\text{Groups}}(G, K^x)$ with the structure of an abelian group: $(\phi \psi)(x) = \phi(x)\psi(x)$ for $x \in G$. The identity element of $\text{Hom}_{\text{Groups}}(G, K^x)$ is the trivial homomorphism sending all $x \mapsto 1$. In order to further describe this group, put $p = \text{char } k$ ($\geq 0$) and let $G_p$ denote the Sylow $p$-subgroup of $G$ and $G_p'$ the subgroup consisting of the $p$-regular elements of $G$, with the understanding that $G_0 = 1$ and $G_0' = G$. Then $G = G_p \times G_{p'}$, and restriction gives a group isomorphism $\text{Hom}_{\text{Groups}}(G, K^x) \cong \text{Hom}_{\text{Groups}}(G_p', K^x)$, because all homomorphisms $G \to K^x$ are trivial on $G_{p'}$. Finally, it is not hard to show that $\text{Hom}_{\text{Groups}}(G_{p'}, K^x) \cong G_{p'}$ as groups (non-canonically). Example 3.9 below does this for cyclic groups and you are asked to generalize the example in Exercise 3.2.1 using (3.3).

To summarize, $\text{Hom}_{\text{Groups}}(G, K^x) \cong G_{p'}$ as groups and we have a bijection
\begin{equation}
(3.17) \quad \text{Irr } kG \leftrightarrow \Gamma \setminus \text{Hom}_{\text{Groups}}(G, K^x)
\end{equation}

Under this bijection, the (singleton) orbit of the identity element of $\text{Hom}_{\text{Groups}}(G, K^x)$ corresponds to the trivial representation, $1$. For future reference, let us record the following fact, which is of course a very special case of Theorem 3.7 if $p > 0$, but which follows directly from the foregoing in any characteristic.

**Proposition 3.8** (notation as above). Assume that $K = k$. Then $\text{Irr } kG = |G_{p'}|$.

**Example 3.9** ($C_n$ over $\mathbb{Q}$). Taking $k = \mathbb{Q}$ and $G = C_n$, the cyclic group of order $n$, we obtain $K = \mathbb{Q}(\zeta_n)$, with $\zeta_n = e^{2\pi i/n} \in \mathbb{C}$, and $\Gamma = \mathbb{Z}/n\mathbb{Z}^\times$. The group $\text{Hom}_{\text{Groups}}(G, K^x)$ is isomorphic to $G$: fixing a generator $x$ for $G$, the group homomorphism $\phi$ that is determined by $\phi(x) = \zeta_n$ will serve as a generator for $\text{Hom}_{\text{Groups}}(G, K^x)$. Explicitly, $\text{Hom}_{\text{Groups}}(G, K^x)$ consists of the maps $\phi^k: x \mapsto \zeta_n^k$ with $0 \leq k \leq n - 1$. Moreover, $\phi^k$ and $\phi^l$ belong to the same $\Gamma$-orbit if and only if the roots of unity $\zeta_n^k$ and $\zeta_n^l$ have the same order. Thus,

$$\text{Irr } \mathbb{Q}C_n = \# \{ \text{ divisors of } n \}.$$
This could also have been obtained from the Chinese Remainder Theorem together with the familiar decomposition \( x^n - 1 = \prod_{m|n} \Phi_m \), where \( \Phi_m \) is the \( m \)th cyclotomic polynomial (which is irreducible over \( \mathbb{Q} \)):

\[
\mathbb{Q}C_n \cong \mathbb{Q}[x]/(x^n - 1) \cong \prod_{m|n} \mathbb{Q}[x]/(\Phi_m).
\]

### 3.2.2. Degree-1 Representations

Recall from (1.36) that, for any \( A \in \text{Alg}_k \), the equivalence classes of degree-1 representations form a subset of \( \text{Irr}_A \) that is in natural 1-1 correspondence with \( \text{Hom}_{\text{Alg}_k}(A, k) \). Since \( \text{Hom}_{\text{Alg}_k}(kG, k) \cong \text{Hom}_{\text{Groups}}(G, k^\times) \) by (3.2), the bijection takes the following form for \( A = kG \):

\[
\text{Hom}_{\text{Groups}}(G, k^\times) \cong \left\{ \text{equivalence classes of degree-1 representations of } kG \right\} \subseteq \text{Irr } kG
\]

Note that \( k\phi \) has character \( \chi_{k\phi} = \phi \). Occasionally, we will simply write \( \phi \) for \( k\phi \).

Recall from §3.2.1 that the group structure of \( k^\times \) makes \( \text{Hom}_{\text{Groups}}(G, k^\times) \) an abelian group with identity element 1. The canonical bijection \( \text{Hom}_{\text{Groups}}(G, k^\times) \cong \text{Hom}_{\text{Groups}}(G^{ab}, k^\times) \), where \( G^{ab} = G/[G, G] \) denotes the abelianization of \( G \) (Example A.4), is in fact an isomorphism of groups. Proposition 3.8 has the following

**Corollary 3.10.** Assume that \( G^{ab} \) is finite with exponent \( e \). If \( \text{char } k \nmid e \) and \( k \) contains all \( e \)th roots of unity (in some algebraic closure of \( k \)), then the number of non-equivalent degree-1 representations of \( G \) is equal to \( |G^{ab}| \).

For representations of finite \( p \)-groups in characteristic \( p > 0 \), we have the following important fact, which is once again immediate from Theorem 3.7. However, we give a simple direct argument below.

**Proposition 3.11.** If \( \text{char } k = p > 0 \) and \( G \) is a finite \( p \)-group, then 1 is the only irreducible representation of \( kG \) up to equivalence.

**Proof.** The case \( G = 1 \) being obvious, assume that \( G \neq 1 \) and let \( S \in \text{Irr } kG \). Our hypotheses on \( k \) and \( G \) imply that \( S \) is finite dimensional and that 1 is the only eigenvalue of \( g_S \) for all \( g \in G \). Choosing \( 1 \neq g \in \mathbb{F}G \), the eigenspace of \( g_S \) is a nonzero subrepresentation of \( S \), and hence it must be equal to \( S \). Thus, \( S \) is a representation of \( k[G/\langle g \rangle] \), clearly irreducible, and so \( S \cong 1 \) by induction on the order of \( G \). \( \square \)
3.2. The Dihedral Group $D_4$

The dihedral group $D_n$ is given by the presentation

\[ D_n \overset{\text{def}}{=} \langle x, y \mid x^n = 1, y^2 = xyx^{-1} \rangle. \]

Geometrically, $D_n$ can be described as the symmetry group of the regular $n$-gon in $\mathbb{R}^2$. The order of $D_n$ is $2n$ and $D_n$ has the structure of a semidirect product:

\[ D_n = \langle x \rangle \rtimes \langle y \rangle \cong C_n \rtimes C_2 \]

Since $x^2 = yxyx^{-1} \in [D_n, D_n]$, it is easy to see that $D_n^{ab} \cong C_2$ for odd $n$ and $D_n^{ab} \cong C_2 \times C_2$ for even $n$.

Let us now focus on $D_4$ and work over a base field $k$ with $\text{char} k \neq 2$. Since $D_4^{ab} \cong C_2 \times C_2$, we know by Corollary 3.10 that $D_4$ has four degree-1 representations: they are given by $\phi_{\pm, \pm}: x \mapsto \pm 1, y \mapsto \pm 1$; so $\phi_{+, +} = 1$. Another representation arises from the realization of $D_4$ as the symmetry group of the square: $x$ acts as the counterclockwise rotation by $\pi/2$ and $y$ as the reflection across the vertical axis of symmetry; see the picture on the right. With respect to the indicated basis $v_1, v_2$, the matrix of $x$ is $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ and $y$ has matrix $\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$. These matrices make sense in $\text{Mat}_2(k)$ and they satisfy the defining relations (3.18) of $D_4$; hence, they yield a representation of $kD_4$. Let us call this representation $S$. Since the matrices $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ and $\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$ have no common eigenvector, $S$ is irreducible. Furthermore, $D(S) = k$: only the scalars in $\text{Mat}_2(k)$ commute with both matrices. It is easy to check that the two matrices generate the algebra $\text{Mat}_2(k)$; so $kD_4/\text{Ker} S \cong \text{Mat}_2(k)$—this also follows from Burnside’s Theorem (§1.4.6). To summarize, we have constructed five non-equivalent irreducible representations of $kD_4$. Since $D_4$ has five conjugacy classes, represented by $1, x^2, x, y$ and $xy$, the list is complete by Proposition 3.6: $\text{Irr} kD_4 = \{1, \phi_{+, -, \pm, \mp, \pm, \mp}, S\}$. Alternatively, the kernels are distinct maximal ideals of $kD_4$, with factors isomorphic to $k$ for all $\phi_{\pm, \pm}$ and to $\text{Mat}_2(k)$ for $S$. Therefore, the Chinese Remainder Theorem yields a surjective map of $k$-algebras, $\mathbb{Z}[D_4] \to k \times k \times k \times k \times \text{Mat}_2(k)$, which must be an isomorphism for dimension reasons. Thus:

\[ kD_4 = k \times k \times k \times k \times \text{Mat}_2(k). \]

In particular, $kD_4$ is split semisimple. We record the character table of $kD_4$; all values in this table have to be interpreted in $k$.

3.2.4. Some Representations of the Symmetric Group $S_n$

Let $S_n$ denote the group of all permutations of the set $[n] = \{1, 2, \ldots, n\}$ and assume that $n \geq 2$. Then $S_n^{ab} = S_n/\mathcal{A}_n \cong C_2$, where $\mathcal{A}_n$ is the alternating group consisting of the even permutations in $S_n$. Thus, besides the trivial representation
There is only one other degree-1 representation, up to equivalence, and only if \( \text{char } k \neq 2 \): this is the so-called **sign representation**,

\[
\text{sgn}: S_n \rightarrow S_n^{ab} \cong \{ \pm 1 \} \subseteq k^x.
\]

In order to find additional irreducible representation of \( S_n \), we use the action of \( S_n \) on the set \([n]\), which we will write as \([n] = \{ b_1, b_2, \ldots, b_n \} \) so as to not confuse its elements with scalars from \( k \). Let \( M_n = k[n] \) denote the \( k \)-vector space with basis \([n]\). The **standard permutation representation** of \( S_n \) is defined by

\[
M_n \overset{\text{def}}{=} \bigoplus_{i=1}^{n} k b_i \quad \text{with} \quad s.b_i = b_{s(i)} \quad (s \in S_n).
\]

In terms of the isomorphism \( \text{GL}(M_n) \cong \text{GL}_n(k) \) that is provided by the given basis of \( M_n \), the image of the homomorphisms \( S_n \rightarrow \text{GL}_n(k) \) consists exactly of the permutation matrices, having one entry equal to 1 in each row and column with all other entries being 0. Note that \( M_n \) is not irreducible: the 1-dimensional subspace spanned by the vector \( \sum_i b_i \in M_n \) is a proper subrepresentation of \( M_n \) that is equivalent to the trivial representation \( \mathbb{1} \). Also, the map

\[
\pi: M_n \rightarrow k = \mathbb{1}
\]

\[
\psi
\sum_i A_i b_i \rightarrow \sum_i A_i
\]

is easily seen to be an epimorphism of representations. Therefore, we obtain a representation of degree \( n - 1 \) by putting

\[
V_{n-1} \overset{\text{def}}{=} \text{Ker } \pi
\]

This is called the **standard representation** of \( S_n \). It is not hard to show that \( V_{n-1} \) is irreducible if and only if either \( n = 2 \) or \( n > 2 \) and \( \text{char } k \nmid n \) and that

---

\*\( V_{n-1} \) is also called the **deleted permutation representation** of \( S_n \).
one always has $\text{End}_k \mathbb{S}_n(V_{n-1}) = k$ (Exercise 3.2.3). Consequently, if $n > 2$ and $\text{char} k \nmid n$, then Burnside’s Theorem ($\S$1.4.6) gives a surjective map of algebras $k \mathbb{S}_n \to \text{BiEnd}_k \mathbb{S}_n(V_{n-1}) \cong \text{Mat}_{n-1}(k)$.

**Example 3.12** (The structure of $k \mathbb{S}_3$). Assume that $\text{char} k \neq 2, 3$. (See Exercise 3.4.1 for characteristics 2 and 3.) Then the foregoing provides us with three non-equivalent irreducible representations for $\mathbb{S}_3$ over $k$: $1$, sgn and $V_2$. Their kernels are three distinct maximal ideals of $k \mathbb{S}_3$, with factors $k$, $k$ and $\text{Mat}_2(k)$, respectively. Exactly as for $k \mathbb{D}_4$ above, we obtain an isomorphism of $k$-algebras,

$$k \mathbb{S}_3 \cong k \times k \times \text{Mat}_2(k).$$

Thus, $k \mathbb{S}_3$ is split semisimple and $\text{Irr} k \mathbb{S}_3 = \{1, \text{sgn}, V_2\}$. Note also that $\mathbb{S}_3$ has three conjugacy classes, with representatives $(1)$, $(1 \ 2)$ and $(1 \ 2 \ 3)$. With respect to the basis $(b_1 - b_2, b_2 - b_3)$ of $V_2$, the operators $(1 \ 2)V_2$ and $(1 \ 2 \ 3)V_2$ have matrices $\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$ and $\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, respectively. Here is the character table of $k \mathbb{S}_3$:

<table>
<thead>
<tr>
<th>classes</th>
<th>sizes</th>
<th>$(1)$</th>
<th>$(1 \ 2)$</th>
<th>$(1 \ 2 \ 3)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$1$</td>
<td>$1$</td>
<td>$1$</td>
<td>$1$</td>
<td></td>
</tr>
<tr>
<td>sgn</td>
<td></td>
<td>$1$</td>
<td>$-1$</td>
<td>$1$</td>
</tr>
<tr>
<td>$\chi_{V_2}$</td>
<td>$2$</td>
<td>$0$</td>
<td>$-1$</td>
<td></td>
</tr>
</tbody>
</table>

*Table 3.2. Character table of $S_3$ (char $k \neq 2, 3$)*

We remark that $\mathbb{S}_3$ is isomorphic to the dihedral group $\mathbb{D}_3$, the group of symmetries of a unilateral triangle, by sending $(1 \ 2)$ to the reflection across the vertical line of symmetry and $(1 \ 2 \ 3)$ to counterclockwise rotation by $2\pi/3$ as in the picture on the right. If $k = \mathbb{R}$, then we may regard $V_2 \cong \mathbb{R}^2$ as the Euclidean plane. Using the basis consisting of $v_1 = \sqrt{3}(b_1 - b_2)$ and $v_2 = b_1 + b_2 - 2b_3$, the matrices of $(1 \ 2)V_2$ and $(1 \ 2 \ 3)V_2$ are $\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$ and $\begin{pmatrix} \cos 2\pi/3 - \sin 2\pi/3 \\ \sin 2\pi/3 \cos 2\pi/3 \end{pmatrix}$, respectively. Thus, over $\mathbb{R}$, the representation $V_2$ also arises from the realization of $\mathbb{S}_3$ as the group of symmetries of the triangle.
3.2.5. Permutation Representations

Returning to the case of an arbitrary group $G$, let us now consider a $G$-set, that is, a set $X$ with a $G$-action

$$G \times X \longrightarrow X$$

$$\omega \quad \omega$$

$$(g, x) \longmapsto g.x$$

satisfying the usual axioms: $1.x = x$ and $g.(g'.x) = (gg').x$ for all $g, g' \in G$ and $x \in X$. We will write $G \subseteq X$ to indicate a $G$-action on $X$. Such an action extends uniquely to an action of $G$ by $\mathbb{k}$-linear automorphisms on the vector space $\mathbb{k}X$ of all formal $\mathbb{k}$-linear combinations of the elements of $X$ (Example A.5), thereby giving rise to a representation

$$\rho_X : G \to \text{GL}(\mathbb{k}X).$$

Representations $V \in \text{Rep}_\mathbb{k}G$ that are equivalent to a representation of this form are called permutation representations of $G$; they are characterized by the fact that the action of $G$ on $V$ stabilizes some $\mathbb{k}$-basis of $V$.

If the set $X$ is finite, then we can consider the character $\chi_{\mathbb{k}X} : G \to \mathbb{k}$. Letting $\text{Fix}_X(g) = \{x \in X \mid g.x = x\}$ denote the set of all fixed points of $g$ in $X$, we evidently have

$$\chi_{\mathbb{k}X}(g) = \#\text{Fix}_X(g)1_\mathbb{k} \quad (g \in G)$$

(3.21)

**Examples 3.13** (Some permutation representations). (a) If $|X| = 1$, then $\mathbb{k}X \cong \mathbb{k}$ and $\chi_\mathbb{k} = 1$.

(b) Taking $X = G$, with $G$ acting on itself by left multiplication, we obtain the regular representation $\rho_G = \rho_{\text{reg}}$ of $\mathbb{k}G$. As elsewhere in this book, we will write $(\mathbb{k}G)_{\text{reg}}$ for this representation. Evidently, $\text{Fix}_G(1) = G$ and $\text{Fix}_G(g) = \emptyset$ if $g \neq 1$. Thus, if $G$ is finite, then the regular character of $\mathbb{k}G$ is given by

$$\chi_{\text{reg}}(g) = \begin{cases} |G|1_\mathbb{k} & \text{for } g = 1 \\ 0 & \text{otherwise} \end{cases}$$

or $\chi_{\text{reg}}(\sum_{g \in G} \alpha_g g) = |G|\alpha_1$. Viewing $\mathbb{k}G$ as a symmetric algebra as in §3.1.6, this formula is identical to (2.30).

(c) We can also let $G$ act on itself by conjugation as in (3.10). The resulting permutation representation is called the adjoint representation of $\mathbb{k}G$; it will be denoted by $(\mathbb{k}G)_{\text{ad}}$. Now we have $\text{Fix}_G(g) = C_G(g)$, the centralizer of $g$ in $G$. Hence, for finite $G$, the character of the adjoint representation is given by

$$\chi_{\text{ad}}(g) = |C_G(g)|1_\mathbb{k}.$$

(d) With $G = S_n$ acting as usual on $X = [n]$, we recover the standard permutation representation $M_n$ of $S_n$. Here, $\text{Fix}_{[n]}(s)$ is the number of 1-cycles in the
disjoint cycle decomposition of \( s \in S_n \). Writing this number as \( \text{Fix}(s) \), we obtain

\[
\chi_{M_n}(s) = \# \text{Fix}(s)_1 \quad (s \in S_n).
\]

Recall from (3.20) that there is a short exact sequence \( 0 \to V_{n-1} \to M_n \to 1 \to 0 \) in \( \text{Rep} \mathbb{K}S_n \). Thus, Lemma 1.41 gives

\[
\chi_{V_{n-1}}(s) = \# \text{Fix}(s)_1 \text{k} - 1 \text{k}.
\]

### Exercises for Section 3.2

**3.2.1 (Dual group).** Let \( G \) be finite abelian and assume that \( \text{char} \mathbb{K} \nmid |G| \). Put \( e = \exp G, \mu_e = \{ \zeta \in \mathbb{K} \mid \zeta^e = 1 \} \) and \( K = \mathbb{K}(\mu_e) \subseteq \mathbb{K} \), a fixed algebraic closure of \( \mathbb{K} \). Use (3.3) to show that \( G \simeq \text{Hom}_{\text{Groups}}(G, K^X) \) as groups.

**3.2.2 (Splitting fields).** Show:

(a) If \( \mathbb{K} \) is a splitting field for \( G \), then \( \mathbb{K} \) is also a splitting field for all homomorphic images of \( G \).

(b) Assume that \( \mathbb{K} \) is a splitting field for \( G \) and that \( G^{ab} \) is finite. Show that \( \mu_e \subseteq \mathbb{K} \), where \( e = \exp(G^{ab}) \) and \( \mu_e \) is as in Exercise 3.2.1.

(c) Give an example showing that if \( \mathbb{K} \) is a splitting field for \( G \), then \( \mathbb{K} \) need not be a splitting field for all subgroups of \( G \).

**3.2.3 (Standard representation of \( S_n \)).** Let \( V_{n-1} (n \geq 2) \) be the standard representation of the symmetric group \( S_n \). Show:

(a) \( V_{n-1} \) is irreducible if and only if \( n = 2 \) or \( \text{char} \mathbb{K} \nmid n \);

(b) \( \text{End}_{\mathbb{K}S_n}(V_{n-1}) = \mathbb{K} \).

**3.2.4 (Deleted permutation representation).** Let \( X \) be a \( G \)-set with \( |X| \geq 2 \) and let \( \pi : \mathbb{K}X \to 1 \) be defined by \( \pi(\sum x \in X \lambda_x x) = \sum x \in X \lambda_x x \). The kernel \( V = \text{Ker} \pi \in \text{Rep} \mathbb{K}G \) is called a deleted permutation representation. Assume that \( G \) is finite with \( \text{char} \mathbb{K} \nmid |G| \) and that the action \( G \acts X \) is doubly transitive, that is, the \( G \)-action on \( \{(x, y) \in X \times X \mid x \neq y\} \) that is given by \( g.(x, y) = (g.x, g.y) \) is transitive. Show that \( V \) is irreducible.

**3.2.5 (The character table does not determine the group).** Consider the real quaternions, \( \mathbb{H} = \mathbb{R} \oplus \mathbb{R}i \oplus \mathbb{R}j \oplus \mathbb{R}k \) with \( i^2 = j^2 = k^2 = ijk = -1 \), and the quaternion group \( Q_8 = \langle i, j \rangle = \{ \pm 1, \pm i, \pm j, \pm k \} \leq \mathbb{H}^X \). Show that \( Q_8 \) has the same character table over any field \( \mathbb{K} \) with \( \text{char} \mathbb{K} \neq 2 \) as the dihedral group \( D_4 \) (Table 3.1), even though \( Q_8 \neq D_4 \).

### 3.3. More Structure

Returning to the general development of the theory of group representations, this section applies the group algebra functor \( \mathbb{K} \cdot \) : \( \text{Groups} \to \text{Alg} \mathbb{K} \) to construct some
algebra maps for the group algebra \( \k G \) that add important structure to \( \k G \) and its category of representations.

### 3.3.1. Invariants

**The Augmentation.** The trivial group homomorphism \( G \to \{1\} \) gives rise to the following algebra map, called the **augmentation map** or **counit** of \( \k G \):

\[
\begin{align*}
\varepsilon: \k G & \longrightarrow \k \\
& \sum_x a_x x & \longmapsto \sum_x a_x
\end{align*}
\]

(3.22)

The map \( \varepsilon \) is the \( \k \)-linear extension of the trivial representation \( \mathbb{1} \) in (3.5) from \( G \) to \( \k G \), and we can also think of it as a map \( \varepsilon: (\k G)_{\text{reg}} \to \mathbb{1} \) in \( \text{Rep} \, \k G \). The kernel of \( \varepsilon \) is called the **augmentation ideal** of \( \k G \); we will use the notation

\[
(\k G)^{+} \overset{\text{def}}{=} \text{Ker} \, \varepsilon.
\]

Clearly, \((\k G)^{+}\) is the \( \k \)-subspace of \( \k G \) that is generated by the subset \( \{g-1 \mid g \in G\} \).

**Invariants and Weight Spaces.** For any \( V \in \text{Rep} \, \k G \), the \( \k \)-subspace of \( G \)-invariants in \( V \) is defined by

\[
V^{G} \overset{\text{def}}{=} \{ v \in V \mid \forall g \in G, g \cdot v = v \}
\]

The invariants can also be described as the common kernel of the operators \( a_v \) with \( a \in (\k G)^{+} \) or, alternatively, as the \( \mathbb{1} \)-homogeneous component \( V(\mathbb{1}) \) of \( V \):

\[
V^{G} = \{ v \in V \mid (\k G)^{+}.v = 0 \} = \{ v \in V \mid a \cdot v = \varepsilon(a) v \text{ for all } a \in \k G \}.
\]

More generally, if \( \k_{\phi} \) is any degree-1 representation of \( G \), given by a group homomorphism \( \phi: G \to \k^{\times} \) (§3.2.2), then the homogeneous component \( V(\k_{\phi}) \) will be written as \( V_{\phi} \) and, if nonzero, referred to as a **weight space** of \( V \) as in Example 1.30:

\[
V_{\phi} \overset{\text{def}}{=} \{ v \in V \mid \forall g \in G, g \cdot v = \phi(g) v \}
\]

The elements of \( V_{\phi} \) are called **weight vectors** or **semi-invariants**.

**Invariants of Permutation Representations.** Let \( X \) be a \( G \)-set and let \( V = \k X \) be the associated permutation representation of \( \k G \) (§3.2.5). An element \( v = \sum_{x \in X} \lambda_{x} x \in \k X \) belongs to \((\k X)^{G}\) if and only if \( \lambda_{g \cdot x} = \lambda_{x} \) for all \( x \in X \) and \( g \in G \); in other words, the function \( X \to \k, x \mapsto \lambda_{x} \), is constant on all \( G \)-orbits \( G.x \subseteq X \). Since
\( \lambda_x = 0 \) for almost all \( x \in X \), we conclude that if \( \lambda_x \neq 0 \) then \( x \) must belong to the following \( G \)-subset of \( X \):

\[
X_{\text{fin}} = \{ x \in X \mid \text{the orbit } G.x \text{ is finite} \}.
\]

For each orbit \( O \in G \setminus X_{\text{fin}} \), we may define the orbit sum

\[
\sigma_O \overset{\text{def}}{=} \sum_{x \in O} x \in (kX)^G
\]

Since distinct orbits are disjoint, the various orbit sums are \( k \)-linearly independent. The orbit sums also span \( (kX)^G \). For any \( v = \sum_{x \in X} \lambda_x x \in (kX)^G \) can be written as \( v = \sum_{O \in G \setminus X_{\text{fin}}} \lambda_O \sigma_O \), where \( \lambda_O \) denotes the common value of all \( \lambda_x \) with \( x \in O \).

To summarize,

\[
(\mathbb{Z}) = \bigoplus_{O \in G \setminus X_{\text{fin}}} k \sigma_O.
\]

The foregoing applies verbatim to any commutative ring \( k \) rather than a field.

**Example 3.14** (Invariants of the adjoint representation). Since \( G \) generates the group algebra \( kG \), the invariants of the adjoint representation (Example 3.13) coincide with the center of \( kG \):

\[
(kG)^G_{\text{ad}} = \{ a \in kG \mid gag^{-1} = a \text{ for all } g \in G \} = \mathcal{Z}(kG).
\]

The set \( G_{\text{fin}} \) for the conjugation action \( G \subset G \) consists of the finite conjugacy classes of \( G \). The corresponding orbit sums are also called the class sums of \( G \); they form a \( k \)-basis of \( \mathcal{Z}(kG) \) by (3.24).

**Example 3.15** (Invariants of the regular representation). Applying (3.24) to the regular representation \( (kG)^G_{\text{reg}} \) and noting that \( X = G \) consists of just one \( G \)-orbit in this case, we obtain

\[
(kG)^G_{\text{reg}} = \begin{cases} 
0 & \text{if } G \text{ is infinite} \\
 k \sigma_G & \text{if } G \text{ is finite, with } \sigma_G = \sum_{g \in G} g
\end{cases}.
\]

Focusing on the case where \( G \) is finite, we have \( a^2 = \epsilon(a)a \) for any \( a \in (kG)^G_{\text{reg}} \) and \( \epsilon(a) \in \epsilon(\sigma_G) k = |G| k \). Thus, \( \epsilon \) is nonzero on \( (kG)^G_{\text{reg}} \) if and only if the group \( G \) is finite and \( \text{char } k \nmid |G| \). In this case, the unique element \( e \in (kG)^G_{\text{reg}} \) satisfying \( \epsilon(e) = 1 \) or, equivalently, \( 0 \neq e = e^2 \) is given by

\[
e = \frac{1}{|G|} \sigma_G = \frac{1}{|G|} \sum_{g \in G} g.
\]
Invariants for Finite Groups: Averaging. For a finite group $G$ and an arbitrary $V \in \text{Rep}_k G$, the operator $(\sigma_G)_V : V \to V$, $v \mapsto \sum_{g \in G} g \cdot v$, clearly has image in $V^G$. If the order of $G$ is invertible in $k$, then the following proposition shows that all $G$-invariants in $V$ are obtained in this way.

**Proposition 3.16.** Let $G$ be finite with $\text{char} k \nmid |G|$ and let $e$ be as in (3.25). Then, for every $V \in \text{Rep}_k G$, the following “averaging operator” is a projection onto $V^G$:

$$e_V : V \to V, \quad v \mapsto \frac{1}{|G|} \sum_{g \in G} g \cdot v$$

If $V$ is finite dimensional, then $\dim_k V^G \cdot 1_k = \chi_V(e) = \frac{1}{|G|} \sum_{g \in G} \chi_V(g)$.

**Proof.** Clearly, $\text{Im} e_V \subseteq V^G$. If $v \in V^G$, then $e_V \cdot v = e(e)v = v$, because $e(e) = 1$. Thus, $e_V$ is a projection onto $V$. With respect to a $k$-basis of $V = e \cdot V \oplus (1 - e) \cdot V = V^G \oplus (1 - e) \cdot V$ that is the union of bases of $V^G$ and $(1 - e) \cdot V$, the matrix of the projection $e_V$ has the form

$$
\begin{pmatrix}
\text{Id}_{V^G} & 0 \\
0 & 0
\end{pmatrix}
$$

Therefore, $\chi_V(e) = \text{trace}(e_V) = \dim_k V^G \cdot 1_k$, which completes the proof. □

The following corollary is variously referred to as Burnside’s Lemma or the Cauchy-Frobenius Lemma, the latter attribution being historically more correct.

**Corollary 3.17.** If a finite group $G$ acts on a finite set $X$, then the number of $G$-orbits in $X$ is equal to the average number of fixed points of elements of $G$:

$$\# G \backslash X = \frac{1}{|G|} \sum_{g \in G} \# \text{Fix}_X(g).$$

**Proof.** By (3.24) we know that $\dim_k (\mathbb{Q} X)^G = \# G \backslash X$ and (3.21) tells us that the character of $\mathbb{Q} X$ is given by $\chi_{\mathbb{Q} X}(g) = \# \text{Fix}_X(g)$ for $g \in G$. The corollary therefore follows from the dimension formula for invariants in Proposition 3.16. □

3.3.2. Comultiplication and Antipode

In this subsection, we construct two further maps for the group algebra $k G$, the comultiplication and the antipode. Equipped with these new maps and the augmentation (counit), $k G$ becomes a first example of a **Hopf algebra**.
**Comultiplication.** Applying the group algebra functor \( k \cdot \) : Groups \( \rightarrow \) \( \text{Alg}_k \) to the diagonal group homomorphism \( G \rightarrow G \times G, x \mapsto (x, x) \), and using the isomorphism \( k[G \times G] \xrightarrow{\sim} kG \otimes kG \) that is given by \( (x, y) \mapsto x \otimes y \) for \( x, y \in G \) (Exercise 3.1.2), we obtain the algebra map

\[
\Delta : \quad kG \longrightarrow kG \otimes kG
\]

\[
\sum_x \alpha_x x \longmapsto \sum_x \alpha_x (x \otimes x)
\]

This map is called the **comultiplication** of \( kG \). The nomenclature “comultiplication” and “counit” derives from the fact that these maps fit into commutative diagrams resembling the diagrams (1.1) for the multiplication and unit maps, except that all arrows now point in the opposite direction:

\[
\begin{array}{ccc}
\Delta \otimes \text{Id} & \Delta & \text{Id} \otimes \Delta \\
\text{Id} & \Delta & \Delta \\
kG \otimes kG & kG \otimes kG & kG \otimes kG
\end{array}
\]

Both diagrams are manifestly commutative. The property of \( \Delta \) that is expressed by the diagram on the left is called **coassociativity**. Another notable property of the comultiplication \( \Delta \) is its **cocommutativity**: Letting \( \tau : kG \otimes kG \rightarrow kG \otimes kG \) denote the map given by \( \tau(a \otimes b) = b \otimes a \), we have

\[
(3.27) \quad \Delta = \tau \circ \Delta
\]

Again, this concept is “dual” to commutativity: recall that an algebra \( A \) with multiplication \( m : A \otimes A \rightarrow A \) is commutative if and only if \( m = m \circ \tau \).

**Antipode.** Inversion gives a group isomorphism \( G \xrightarrow{\sim} G^{\text{op}}, x \mapsto x^{-1} \). Here \( G^{\text{op}} \) denotes the opposite group: \( G^{\text{op}} = G \) as sets, but with new group operation \( * \) given by \( x * y = yx \). We obtain a \( k \)-linear map,

\[
(3.28) \quad S : kG \longrightarrow kG
\]

\[
\sum_x \alpha_x x \longmapsto \sum_x \alpha_x x^{-1}
\]

satisfying \( S(ab) = S(b)S(a) \) for all \( a, b \in kG \) and \( S^2 = \text{Id} \). The map \( S \) is called the **standard involution** or the **antipode** of \( kG \). We can also think of \( S \) as an isomorphism \( S : kG \xrightarrow{\sim} k[G^{\text{op}}] \equiv (kG)^{\text{op}} \) in \( \text{Alg}_k \) (Exercise 3.1.2).
3.3.3. A Plethora of Representations

The structure maps in §§3.3.1, 3.3.2 allow us to construct many new representations of \( kG \) from given representations. This is sometimes referred to under the moniker “plethysm.” Analogous constructions will later be carried out also in the context of Lie algebras and, more generally, Hopf algebras. We will then refer to some of the explanations below.

**Homomorphisms.** For given \( V, W \in \text{Rep} \ kG \), the \( k \)-vector space \( \text{Hom}_k(V, W) \) can be made into a representation of \( kG \) by defining

\[
(g.f)(v) \overset{\text{def}}{=} g.f(g^{-1}.v) \quad (g \in G, v \in V, f \in \text{Hom}_k(V, W))
\]

Even though it is straightforward to verify that this rule does indeed define a representation of \( G \), let us place it in a more conceptual framework. If \( V \) and \( W \) are representations of arbitrary algebras \( B \) and \( A \), resp., then \( \text{Hom}_k(V, W) \) becomes a representation of the algebra \( A \otimes B^{\text{op}} \) as in Example 1.3:

\[
(a \otimes b^{\text{op}}).f = a_W \circ f \circ b_V \quad a \in A, b \in B, f \in \text{Hom}_k(V, W).
\]

Thus, we have a map \( A \otimes B^{\text{op}} \to \text{End}_k(\text{Hom}_k(V, W)) \) in \( \text{Alg}_k \). For \( A = B = kG \), we also have the map \( (\text{Id} \otimes \Delta) \circ \Delta : kG \to kG \otimes kG \to kG \otimes (kG)^{\text{op}} \). Restricting the composite of these two algebra maps to \( G \) leads to (3.29).

The bifunctor \( \text{Hom}_k \) for \( k \)-vector spaces (§B.3.2) restricts to a bifunctor

\[
\text{Hom}_k(\ , \ , ) : (\text{Rep} \ kG)^{\text{op}} \times \text{Rep} \ kG \to \text{Rep} \ kG.
\]

Here, we use \(^{\text{op}}\) for the first variable, because \( \text{Hom}_k \) is contravariant in this variable whereas \( \text{Hom}_k \) is covariant in the second variable: for any map \( f : W \to W' \) in \( \text{Rep} \ kG \), we have \( \text{Hom}_k(f, V) = f^* = \cdot \circ f : \text{Hom}_k(V, W') \to \text{Hom}_k(V, W) \) but \( \text{Hom}_k(V, f) = f_* = f \circ \cdot : \text{Hom}_k(V, W) \to \text{Hom}_k(V, W') \). It is readily verified that \( f^* \) and \( f_* \) are indeed morphisms in \( \text{Rep} \ kG \). Recall also that \( \text{Hom}_k \) is exact in either argument (§B.3.2).

Evidently, \( g.f = f \) holds for all \( g \in G \) if and only if \( f(g.v) = g.f(v) \) for all \( g \in G \) and \( v \in V \), and the latter condition in turn is equivalent to \( f(a.v) = a.f(v) \) for all \( a \in kG \) and \( v \in V \). Thus, the \( G \)-invariants of \( \text{Hom}_k(V, W) \) are exactly the homomorphism \( V \to W \) in \( \text{Rep} \ kG \):

\[
\text{Hom}_k(V, W)^G = \text{Hom}_k(V, W)
\]

\(^7\)This term originated in the theory of symmetric functions; see Littlewood [133, p. 289].
Example 3.18. The following map is easily seen to be an isomorphism in $\text{Rep } kG$:

$$\text{Hom}_k(\mathbb{1}, V) \xrightarrow{\sim} V$$

$$\text{Hom}_k(\mathbb{1}, V) \xrightarrow{\sim} V$$

$$\text{Hom}_k(\mathbb{1}, V) \xrightarrow{\sim} V$$

$$\text{Hom}_k(\mathbb{1}, V) \xrightarrow{\sim} V$$

$$f \mathrel{\mapsto} f(1)$$

By (3.30), this map restricts to an isomorphism $\text{Hom}_k(\mathbb{1}, V) \xrightarrow{\sim} V^G$.

Duality. Taking $W = \mathbb{1} = k_\mathbb{1}$ in the preceding paragraph, the dual vector space $V^* = \text{Hom}_k(V, k)$ becomes a representation of $kG$. By (3.29), the $G$-action on $V^*$ is given by $\langle g.f, v \rangle = \langle f, g^{-1}.v \rangle$ for $g \in G$, $v \in V$ and $f \in V^*$ or, equivalently,

$$a.f = f \circ S(a)_V \quad (a \in kG, f \in V^*)$$

where $S$ is the antipode (3.28). By our remarks about $\text{Hom}_k$, duality gives an exact contravariant functor

$$\cdot^*: \text{Rep } kG \to \text{Rep } kG$$

A representation $V$ is called self-dual if $V \cong V^*$ in $\text{Rep } kG$. Note that this forces $V$ to be finite dimensional, because otherwise $\dim_k V^* > \dim_k V$ (§B.3.2). The lemma below shows that finite-dimensional permutation representations are self-dual; further self-dual representations can be constructed with the aid of Exercise 3.3.10.

Lemma 3.19. The permutation representation $kX$ for a finite $G$-set $X$ is self-dual.

Proof. Let $(\delta_x)_{x \in X} \in (kX)^*$ be the dual basis for the basis $X$ of $kX$; so $\langle \delta_x, y \rangle = \delta_{x,y} 1_k$ for $x, y \in X$. Then $x \mapsto \delta_x$ defines a $k$-linear isomorphism $\delta: kX \xrightarrow{\sim} (kX)^*$. We claim that this is in fact an isomorphism in $\text{Rep } kG$, that is, $\delta(a.v) = a.\delta(v)$ holds for all $a \in kG$ and $v \in kX$. By linearity, we may assume that $a = g \in G$ and $v = x \in X$. The following calculation, for any $y \in X$, shows that $\delta(g.x) = g.\delta(x)$:

$$\langle \delta_{g.x}, y \rangle = \delta_{g.x,y} 1_k = \delta_{x,g^{-1}y} 1_k = \langle \delta_x, g^{-1}.y \rangle = \langle g, \delta_x, y \rangle.$$  

$\square$

Tensor Products. Let $V, W \in \text{Rep } kG$ be given. Then the tensor product $V \otimes W$ becomes a representation of $kG$ via the “diagonal action” $g_{V \otimes W} = g_V \otimes g_W$ for $g \in G$, or

$$g.(v \otimes w) \overset{\text{def}}{=} g.v \otimes g.w \quad (g \in G, v \in V, w \in W)$$

(3.31)

The switch map gives an isomorphism $\tau: V \otimes W \xrightarrow{\sim} W \otimes V$, $v \otimes w \mapsto w \otimes v$ and it is also clear that $V \otimes \mathbb{1} \cong V$ in $\text{Rep } kG$. Finally, the $G$-action (3.31) clearly is compatible with the standard associativity isomorphism for tensor products; so the tensor product in $\text{Rep } kG$ is associative. The tensor product construction makes $\text{Rep } kG$ an example of a tensor category or monoidal category; see [68].
Again, let us place the action rule (3.31) in a more general context. Recall from (1.51) that the outer tensor product of representations \( V \in \text{Rep} A \) and \( W \in \text{Rep} B \), for arbitrary algebras \( A \) and \( B \), is a representation of the algebra \( A \otimes B \): the algebra map \( A \otimes B \to \text{End}_k(V \otimes W) \) is given by \( a \otimes b \mapsto a_V \otimes b_W \). If \( A = B = kG \), then we also have the comultiplication \( \Delta : kG \to kG \otimes kG \). The composite with the previous map is an algebra map \( kG \to \text{End}_k(V \otimes W) \) that gives the diagonal \( G \)-action (3.31).

**Tensor, Symmetric and Exterior Powers.** The action (3.31) inductively gives diagonal \( G \)-actions on all tensor powers \( V^{\otimes k} \) of a given \( V \in \text{Rep} kG \), with \( g \in G \) acting on \( V^{\otimes k} \) by the \( k \)-linear automorphisms \( g_{V^k} \) (§B.1.3):

\[
g(V_1 \otimes V_2 \otimes \cdots \otimes V_k) = g.V_1 \otimes g.V_2 \otimes \cdots \otimes g.V_k.\]

Thus, \( V^{\otimes k} \in \text{Rep} kG \). Defining \( V^{\otimes 0} \) to be the trivial representation, \( 1 \), the tensor algebra \( TV = \bigoplus_{k \geq 0} V^{\otimes k} \) becomes a \( k\)-\( G \)-representation as well. An element \( g \in G \) acts on \( TV \) by the \( k \)-algebra automorphism \( Tg_V = \bigoplus_k g_{V^k} \) that comes from the functor \( T : \text{Vec}_k \to \text{Alg}_k \) (§1.1.2).

Similarly, the symmetric algebra \( \text{Sym} V \) and the exterior algebra \( \wedge V \) become representations of \( kG \) via the functors \( \text{Sym} : \text{Vec}_k \to \text{CommAlg}_k \) and \( \wedge : \text{Vec}_k \to \text{Alg}_k \), with \( g \in G \) acting by the graded \( k \)-algebra automorphisms \( \text{Sym} g_V \) and \( \wedge g_V \), respectively. Since the homogeneous components \( \text{Sym}^k V \) and \( \wedge^k V \) are stable under these actions, we also obtain \( \text{Sym}^k V, \wedge^k V \in \text{Rep} kG \). The canonical epimorphisms \( V^{\otimes k} \to \text{Sym}^k V \) and \( V^{\otimes k} \to \wedge^k V \) are maps in \( \text{Rep} kG \). On \( \wedge^k V \), for example, an element \( g \in G \) acts via the map \( \wedge^k g_V \) (§1.1.2):

\[
g(V_1 \wedge V_2 \wedge \cdots \wedge V_k) = g.V_1 \wedge g.V_2 \wedge \cdots \wedge g.V_k.\]

If \( \dim_k V = n < \infty \), then \( \wedge^n V \) is a degree-1 representation that is given by the group homomorphism \( \det : G \to k^\times \), \( g \mapsto \det(g_V) \) by (1.14):

\[
\wedge^n V \cong k_{\det V} .
\]

**\( G \)-Algebras.** As we have seen, the tensor, symmetric and exterior algebras of a given \( V \in \text{Rep} kG \) all become representations of \( kG \), with \( G \) acting by algebra automorphisms. More generally, any \( A \in \text{Alg}_k \) that is equipped with an action \( G \subseteq \text{Aut}_k A \) by \( k \)-algebra automorphisms is called a **\( G \)-algebra** in the literature (e.g., [197]). Thus, \( A \in \text{Rep} kG \) by virtue of the given \( G \)-action. The conditions

\[
g(ab) = (g.a)(g.b) \quad \text{and} \quad g.1 = 1 \quad (g \in G, a, b \in A)
\]

state, respectively, that the multiplication \( m : A \otimes A \to A \) and the unit \( u : k = 1 \to A \) are maps in \( \text{Rep} kG \). Thus, \( G \)-algebras can be described concisely as “algebras in the category \( \text{Rep} kG \)”: objects \( A \in \text{Rep} kG \) that are equipped with two maps in \( \text{Rep} kG \), the multiplication \( m : A \otimes A \to A \) and the unit map \( u : 1 \to A \), such that the algebra axioms (1.1) are satisfied. Thus, ordinary \( k \)-algebras are algebras in
V \text{ect}_k$. Morphisms of $G$-algebras, by definition, are maps that are simultaneously maps in $\text{Rep} \ kG$ and $\text{Alg}_k$, that is, $G$-equivariant algebra maps. With this, we obtain a category, 

$G\text{Alg}_k$.

We shall later meet some variants and generalizations of the concept of a $G$-algebra (§§5.5.5 and 10.4.1).

**Canonical Isomorphisms and Characters.** The standard maps in $\text{Vect}_k$ discussed in Appendix B all restrict to morphisms in $\text{Rep} \ kG$; Exercises 3.3.9 and 3.3.12 ask the reader to check this. Specifically, for $U, V, W \in \text{Rep} kG$, the Hom-$\otimes$ adjunction isomorphism (B.15) is an isomorphism in $\text{Rep} kG$:

(3.33) \[ \text{Hom}_k(U \otimes V, W) \cong \text{Hom}_k(U, \text{Hom}_k(V, W)). \]

Similarly the canonical monomorphisms $W \otimes V^* \hookrightarrow \text{Hom}_k(V, W)$ and $V \hookrightarrow V^{**}$ in (B.18) and (B.22) are morphisms in $\text{Rep} kG$, and so is the trace map $\text{End}_k(V) \to k$ for $V \in \text{Rep}_\text{fin} kG$ when $k$ is viewed as the trivial representation, $k = 1$. Thus, we have the following isomorphisms in $\text{Rep} kG$:

(3.34) \[ W \otimes V^* \cong \text{Hom}_k(V, W). \]

provided at least one of $V, W$ is finite dimensional. In this case, (3.33) and (3.34) give the following isomorphism, for any $U \in \text{Rep} kG$,

(3.35) \[ \text{Hom}_k(U \otimes V, W) \cong \text{Hom}_k(U, W \otimes V^*). \]

Finally, for any $V \in \text{Rep}_\text{fin} kG$,

(3.36) \[ V \cong V^{**}. \]

**Lemma 3.20.** Let $V, W \in \text{Rep}_\text{fin} kG$.

(a) The characters of the representations $V^*, V \otimes W$ and $\text{Hom}_k(V, W)$ are given, for $g \in G$, by

(i) \[ \chi_{V^*}(g) = \chi_V(g^{-1}); \]

(ii) \[ \chi_{V \otimes W}(g) = \chi_V(g) \chi_W(g); \]

(iii) \[ \chi_{\text{Hom}_k(V, W)}(g) = \chi_W(g) \chi_V(g^{-1}). \]

(b) If $G$ is finite with $\text{char } k \nmid |G|$, then

\[ \dim_k \text{Hom}_{kG}(V, W) \cdot 1_k = \frac{1}{|G|} \sum_{g \in G} \chi_W(g) \chi_V(g^{-1}). \]

**Proof.** (a) The formula $a. f = f \circ S(a)_V$ for $a \in kG$ and $f \in V^*$ can be written as

$\alpha_{V^*} = S(a)_V^*$,

where $S(a)_V^*$ is the transpose of the operator $S(a)_V$ (§B.3.2). Since $\text{trace}(S(a)_V) = \text{trace}(S(a)_V^*)$ by (B.25), we obtain $\chi_{V^*}(a) = \chi_V(S(a))$. Formula (i) follows, because $S(g) = g^{-1}$ for $g \in G$. For (ii), recall that $g_{V \otimes W} = g_V \otimes g_W$. Thus, (ii) is a special
Exercises for Section 3.3

3.3.1 (The adjoint representation). Consider the adjoint representation of $(\mathbb{k}G)_{\text{ad}}$ of group $G$ as in Example 3.13(c).

(a) Show that $(\mathbb{k}G)_{\text{ad}} \cong \bigoplus_x \mathbb{k}[G/C_G(x)]$, where $x$ runs over a set of representatives of the conjugacy classes of $G$, and $\text{Ker}(\mathbb{k}G)_{\text{ad}} = \bigcap_{g \in G} \mathbb{k}G(\mathbb{k}C_G(g))^\ast$. (Use Exercise 3.3.6.)

(b) For $G = S_3$, show that $\text{Ker}(\mathbb{k}G)_{\text{ad}} = 0$ if and only if $\text{char} \mathbb{k} \neq 3$.

3.3.2 (Invariants of outer tensor products). Let $G$ and $H$ be arbitrary groups and let $V \in \text{Rep} \mathbb{k}G$ and $W \in \text{Rep} \mathbb{k}H$. Identifying $\mathbb{k}[G \times H]$ with $\mathbb{k}G \otimes \mathbb{k}H$ (Exercise 3.1.2), consider the outer tensor product $V \boxtimes W \in \text{Rep} \mathbb{k}[G \times H]$ as in (1.51). Show that $(V \boxtimes W)^{G \times H} \cong V^G \otimes W^H$.

3.3.3 (Coinvariants). Let $V \in \text{Rep} \mathbb{k}G$. Dually to the definition of the $G$-invariants $V^G$ (§3.3.1), the $G$-coinvariants in $V$ are defined by

$$V_G \overset{\text{def}}{=} V / (\mathbb{k}G)^\ast V = V / \sum_{g \in G} (g - 1) V.$$ 

Thus, $V_G \cong 1' \otimes_{\mathbb{k}G} V$, where $1'$ is the right $\mathbb{k}G$-module $\mathbb{k}$ with trivial $G$-action.

Let $G$ be finite and let $\sigma_G = \sum_{g \in G} g \in \mathbb{k}G$. Show that the operator $(\sigma_G)_V \in \text{End}_k(V)$ yields a well-defined $\mathbb{k}$-linear map $V_G \rightarrow V^G$ and that this map yields a natural equivalence of functors $\cdot_G \cong \cdot^G : \text{Rep} \mathbb{k}G \rightarrow \text{Vect}_k$ if $\text{char} \mathbb{k} \nmid |G|$.

3.3.4 (Representations as functors: limits and colimits). This exercise assumes familiarity with limits and colimits of functors; see [140, III.3 and III.4]. Let $G$ denote the category with one object, $\ast$, and with $\text{Hom}_G(\ast, \ast) = G$ as in Exercise 3.1.1. Recall that any $V \in \text{Rep} \mathbb{k}G$ gives a functor $F_V : G \rightarrow \text{Vect}_k$ and conversely. Show that $\text{lim} F_V \cong V^G$ and $\text{colim} F_V \cong V_G$.

3.3.5 (Permutation representations). Let $G$-Sets denote the category with objects the $G$-sets (§3.2.5) and morphisms the $G$-equivariant functions, that is, functions $f : X \rightarrow Y$ for $X, Y \in G$-Sets such that $f(g \cdot x) = g \cdot f(x)$ for $g \in G, x \in X$.

(a) Show that $X \mapsto \mathbb{k}X$ gives a functor $G$-Sets $\rightarrow \text{Rep} \mathbb{k}G$ satisfying $\mathbb{k}X \cong \bigoplus_X \mathbb{k}X_g$ for the disjoint union $X = \bigsqcup_X \mathbb{X}_g$ of a family $X_g \in G$-Sets. Furthermore, equipping the cartesian product $X \times Y$ of $X, Y \in G$-Sets with the $G$-action $g_{\cdot}(x, y) = (g \cdot x, g \cdot y)$, we have the isomorphism $\mathbb{k}[X \times Y] \cong (\mathbb{k}X) \otimes (\mathbb{k}Y)$ in $\text{Rep} \mathbb{k}G$. 

3.3.6 (Natural equivalence of functors). Let $\text{Rep} \mathbb{k}G$ and $\text{Rep} \mathbb{k}H$ be categories with one object, $\ast$, and with morphisms $\text{Hom}_{\text{Rep} \mathbb{k}G}(\ast, \ast) = G$ and $\text{Hom}_{\text{Rep} \mathbb{k}H}(\ast, \ast) = H$. Consider the outer tensor product $\mathbb{k}G \otimes \mathbb{k}H$, which is a category with one object $\ast$, and with morphisms $\mathbb{k}G \otimes \mathbb{k}H(\ast, \ast') = G \times H$. Show that $\text{Rep} \mathbb{k}G \otimes \text{Rep} \mathbb{k}H$ is equivalent to $\text{Rep} \mathbb{k}[G \times H]$.
(b) Let $H$ be a subgroup of $G$. For any $X \in H$-Sets, let $H$ act on the cartesian product $G \times X$ by $h.(g, x) := (gh^{-1}, hx)$ and let $G \times_H X := H \setminus (G \times X)$ denote the set of orbits under this action. Writing elements of $G \times_H X$ as $[g, x]$, show that $G \times_H X \in G$-Sets via the $G$-action $g.[g', x] := [gg', x]$. Moreover, show that $\mathbb{k}[G \times_H X] \cong (\mathbb{k}X)^G_H$ in $\text{Rep} \mathbb{k}G$. Conclude in particular that $1_H^G \cong \mathbb{k}[G/H]$.

(c) Let $X \in G$-Sets be transitive; so $X \cong G/H$, where $H = \{g \in G \mid g.x = x\}$ is the isotropy group of some $x \in X$. Show that a $\mathbb{k}$-basis of $\text{End}_{\mathbb{k}G}(\mathbb{k}X)$ is given by the endomorphisms $\phi_O$ that are defined by $\phi_O(x) = \sigma_O$ for $O \in H \setminus X$.

3.3.6 (Relative augmentation ideals). Let $H$ be a subgroup of $G$. Consider the map

$$
eu H^G : \mathbb{k}G \cong (\mathbb{k}H)^G_H \to 1_H^G \cong \mathbb{k}[G/H].$$

Show that $\ker e^G_H$ is the left ideal $\mathbb{k}G(\mathbb{k}H)^+$ of $\mathbb{k}G$ and that this left ideal is generated by the elements $h_i - 1$, where $\{h_i\}$ is any generating set of the subgroup $H$. If $H$ is normal in $G$, then show that $\mathbb{k}G(\mathbb{k}H)^+ = (\mathbb{k}H)^+\mathbb{k}G$ is an ideal of $\mathbb{k}G$.

3.3.7 (Coinvariants of permutation representations). Let $X$ be a $G$-set. For each $x \in X$, let $G \cdot x$ denote the $G$-orbit in $X$ containing $x$ and, for each $O \in G \cdot X_{\text{fin}}$, let $\sigma_O \in \mathbb{k}G$ denote the orbit sum (3.24).

(a) Show that the orbit map $\mathbb{k}X \to \mathbb{k}[G \cdot X], x \mapsto G \cdot x$, descends to $(\mathbb{k}X)_G$ and yields an isomorphism $(\mathbb{k}X)_G \cong \mathbb{k}[G \cdot X]$ in $\text{Vect}_\mathbb{k}$. Furthermore, show that the image of $(\mathbb{k}X)^G$ under the orbit map consists of the $\mathbb{k}$-linear span of all finite orbits whose size is not divisible by char $\mathbb{k}$.

(b) Show that $(\mathbb{k}X)^G \hookrightarrow \mathbb{k}[G \cdot X]$ in $\text{Vect}_\mathbb{k}$ via $\sigma_O \mapsto O$.

3.3.8 (Complete reducibility of permutation representations). (a) Let $X$ be a $G$-set. Show that if the permutation representation $\mathbb{k}X$ is completely reducible, then all $G$-orbits in $X$ are finite and have size not divisible by char $\mathbb{k}$.

(b) Let $\text{char} \mathbb{k} = 3$ and let $X$ denote the collection of all 2-element subsets of $\{1, 2, \ldots, 5\}$ with the natural action by $G = S_5$. So $X \cong G/H$ as $G$-sets, where $H = S_2 \times S_3 \leq S_5$, and $|X| = 10$. Use Exercise 3.3.5(c) to show that $\text{End}_{\mathbb{k}G}(\mathbb{k}X)$ is a 3-dimensional commutative $\mathbb{k}$-algebra that has nonzero nilpotent elements. Conclude from Proposition 1.33 that $\mathbb{k}X$ is not completely reducible. Thus, the converse of (a) fails. (This example was communicated to me by Karin Erdmann.)

3.3.9 (Duality). For $V \in \text{Rep} \mathbb{k}G$, show:

(a) The canonical $\mathbb{k}$-linear injection $V \hookrightarrow V^{**}$ in (B.22) is a morphism in $\text{Rep} \mathbb{k}G$. In particular, if $V$ is finite dimensional, then $V \cong V^{**}$ in $\text{Rep} \mathbb{k}G$.

(b) Conclude from exactness of the contravariant functor $^* : \text{Rep} \mathbb{k}G \to \text{Rep} \mathbb{k}G$ that irreducibility of $V^*$ forces $V$ to be irreducible. The converse holds if $V$ is finite dimensional but not in general.

(c) The dual $(V_G)^* \hookrightarrow V^*$ of the canonical map $V \to V^*_G$ (Exercise 3.3.3), gives a natural isomorphism $(V_G)^* \cong (V^*)^G$. 

3.3.10 (Duality, induction and coinduction). (a) Let $H \to G$ be a group homomorphism and let $W \in \text{Rep}_k H$. Show that $\text{Coind}^k_{kH} W^* \cong (\text{Ind}^k_{kH} W)^*$ in $\text{Rep}_k G$.

(b) Conclude from (a) and Proposition 3.4 that, for any finite-index subgroup $H \leq G$, dualizing commutes with induction: $\text{Ind}^k_{kH} W^* \cong (\text{Ind}^k_{kH} W)^*$ for all $W \in \text{Rep}_k H$.

3.3.11 (Twisting). Let $V \in \text{Rep}_k G$. Representations of the form $k \phi \otimes V$ with $\phi \in \text{Hom}_{\text{Groups}}(G, k^*)$ are called twists of $V$. Prove:

(a) The map $f \mapsto 1 \otimes f(1)$ is an isomorphism $\text{Hom}_k(k \phi, V) \xrightarrow{\sim} k \phi^{-1} \otimes V$ in $\text{Rep}_k G$. In particular, $(k \phi)^* \cong k \phi^{-1}$.

(b) $(k \phi \otimes V)^G \cong V^G \phi^{-1}$, the $\phi^{-1}$-weight space of $V$ (§3.3.1).

(c) The map $G \to (kG)^*$, $g \mapsto \phi(g)g$, is a group homomorphism that lifts uniquely to an algebra automorphism $\tilde{\phi} \in \text{Aut}_{\text{Alg}_k}(kG)$. The $\tilde{\phi}$-twist (1.24) of $V$ is isomorphic to $k \phi^{-1} \otimes V$.

(d) Twisting gives an action of the group $\text{Hom}_{\text{Groups}}(G, k^*)$ on $\text{Irr}_k G$, on the set of completely reducible representations of $kG$ etc. (See Exercise 1.2.3.)

3.3.12 (Hom-Tensor relations). Let $U, V, W \in \text{Rep}_k G$. Show:

(a) The canonical embedding $W \otimes V^* \hookrightarrow \text{Hom}_k(V, W)$ in (B.18) is a morphism in $\text{Rep}_k G$. In particular, if at least one of $V, W$ is finite dimensional, then $W \otimes V^* \cong \text{Hom}_k(V, W)$ in $\text{Rep}_k G$.

(b) The trace map $\text{End}_k(V) \xrightarrow{\sim} V \otimes V^* \rightarrow k$ in (B.23) is a morphism in $\text{Rep}_k G$ when $k = \mathbb{1}$ is viewed as the trivial representation.

(c) $\text{Hom} \otimes \text{adjunction}$ (B.15) gives an isomorphism $\text{Hom}_k(U \otimes V, W) \cong \text{Hom}_k(U, \text{Hom}_k(V, W))$ in $\text{Rep}_k G$. Conclude that if $V$ or $W$ is finite dimensional, then the isomorphisms (B.21) and (B.20) are isomorphisms $V^* \otimes W^* \cong (W \otimes V)^*$ and $\text{Hom}_k(U \otimes V, W) \cong \text{Hom}_k(U, W \otimes V^*)$ in $\text{Rep}_k G$.

3.3.13 (Tensor product formula). Let $H$ be a subgroup of $G$ and let $V \in \text{Rep}_k G$ and $W \in \text{Rep}_k H$.

(a) Show that $V \otimes (W^G)^* \cong (V^H \otimes W)^G$ in $\text{Rep}_k G$.

(b) Conclude from (a) that $V^H \otimes W^G \cong V \otimes (W^G)^G$ in $\text{Rep}_k G$.

3.3.14 (Symmetric and exterior powers). Let $V, W \in \text{Rep}_k G$. Prove:

(a) The isomorphisms $\text{Sym}(V \oplus W) \cong \text{Sym} V \otimes \text{Sym} W$ and $\Lambda(V \oplus W) \cong \Lambda V \otimes \Lambda W$ of Exercise 1.1.12 are also isomorphisms in $\text{Rep}_k G$.

(b) If $\dim_k V = n$, then $\Lambda^k V \cong \text{Hom}_k(\Lambda^{n-k} V, \Lambda^n V) \cong k \text{det}_V \otimes (\Lambda^{n-k} V)^*$ in $\text{Rep}_k G$.

*Automorphisms of this form are called winding automorphisms; see also Exercise 10.1.6.
3.4. Semisimple Group Algebras

The material set out in Section 3.3 allows for a quick characterization of semisimple group algebras; this is the content of Maschke’s Theorem (§3.4.1). The remainder of this section then concentrates on the main tools of the trade: characters, especially the orthogonality relations.

3.4.1. The Semisimplicity Criterion

The following theorem is a celebrated result due to Heinrich Maschke [146] dating back to 1899.

**Maschke’s Theorem.** The group algebra \( kG \) is semisimple if and only if \( G \) is finite and \( \text{char} \ k \nmid |G| \).

**Proof.** First assume that \( kG \) is semisimple; so \( (kG)_{\text{reg}} \) is the direct sum of its various homogeneous components. By Schur’s Lemma, the counit \( \varepsilon : (kG)_{\text{reg}} \to \mathbb{1} \) vanishes on all homogeneous components except for the \( 1 \)-homogeneous component, \( (kG)_{\text{reg}}^G \). Therefore, \( \varepsilon \) must be nonzero on \( (kG)_{\text{reg}}^G \), which forces \( G \) to be finite with \( \text{char} \ k \nmid |G| \) (Example 3.15).

Conversely, assume that \( G \) is finite and \( \text{char} \ k \nmid |G| \). Then, as we have already pointed out in §3.1.6, semisimplicity of \( kG \) can be established by invoking Theorem 2.21. However, here we offer an alternative argument by showing directly that every \( kG \)-representation \( V \) is completely reducible: every subrepresentation \( U \subseteq V \) has a complement (Theorem 1.28). To this end, we will construct a map \( \pi \in \text{Hom}_{kG}(V, U) \) with \( \pi|_U = \text{Id}_U \); then \( \ker \pi \) will be the desired complement for \( U \) (Exercise 1.1.2). In order to construct \( \pi \), start with a \( k \)-linear projection map \( p : V \to U \) along some vector space complement for \( U \) in \( V \); so \( p \in \text{Hom}_k(V, U) \) and \( p|_U = \text{Id}_U \). With

\[
\pi = e \cdot p \in \text{Hom}_{kG}(V, U)^G = \text{Hom}_{kG}(V, U).
\]

For \( u \in U \), we have

\[
\pi(u) = \frac{1}{|G|} \sum_{g \in G} g \cdot p(g^{-1} \cdot u) = u,
\]

because each \( g^{-1} \cdot u \in U \) and so \( p(g^{-1} \cdot u) = g^{-1} \cdot u \). Thus, \( \pi|_U = \text{Id}_U \) and the proof is complete. \( \square \)

For a completely different argument proving semisimplicity of \( kG \) for a finite group \( G \) and a field \( k \) of characteristic 0, see Exercise 3.4.2. The following corollary specializes some earlier general results about split semisimple algebras to group algebras.

**Corollary 3.21.** Assume that \( G \) is finite with \( \text{char} \ k \nmid |G| \) and that \( k \) is a splitting field for \( G \). Then:

(a) The irreducible characters form a basis of the space \( \text{cf}_k(G) \) of all \( k \)-valued class functions on \( G \). In particular, \( \# \text{Irr} kG = \# \{ \text{conjugacy classes of } G \} \).
(b) \(|G| = \sum_{S \in \text{Irr} \, kG} (\text{dim}_k S)^2\).

(c) \(m(S, (kG)_{\text{reg}}) = \dim_k S\) for all \(S \in \text{Irr} \, kG\).

**Proof.** All parts come straight from the corresponding parts of Corollary 1.35. Part (a) also uses the fact that the irreducible characters form a \(k\)-basis of the space \((kG)^*_{\text{trace}} \cong \text{cf}_k(G)\) by Theorem 1.44 and (3.11). \(\square\)

### 3.4.2. Orthogonality Relations

For the remainder of Section 3.4, we assume \(G\) to be finite with \(\text{char} \, k \nmid |G|\).

#### An Inner Product for Characters.

For \(\phi, \psi \in \text{cf}_k(G)\), we define

\[(3.37) \quad \langle \phi, \psi \rangle \overset{\text{def}}{=} \frac{1}{|G|} \sum_{g \in G} \phi(g)\psi(g^{-1})\]

This gives a symmetric \(k\)-bilinear form \(\langle \cdot, \cdot \rangle : \text{cf}_k(G) \times \text{cf}_k(G) \rightarrow k\) that is non-degenerate. For, if \((\delta_C, \delta_D)\) is the basis of \(\text{cf}_k(G)\) given by the class functions \(\delta_C\) with \(\delta_C(g) = 1_k\) if \(g\) belongs to the conjugacy class \(C\) and \(\delta_C(g) = 0_k\) otherwise, then \(\langle \delta_C, \delta_D \rangle = \frac{|C|}{|G|}\delta_{C,D}\). We may now restate Lemma 3.20(b) as follows:

\[(3.38) \quad \langle \chi_V, \chi_W \rangle = \langle \chi_W, \chi_V \rangle = \dim_k \text{Hom}_k(G, W) \cdot 1_k\]

In particular, for any subgroup \(H \leq G\) and any \(W \in \text{Rep}_{\text{fin}} \, kH\) and \(V \in \text{Rep}_{\text{fin}} \, kG\),

\[(3.39) \quad \langle \chi_{W^G}, \chi_V \rangle = \langle \chi_W, \chi_{V^H} \rangle .\]

This is the original version of **Frobenius reciprocity**; it follows from (3.38) and the isomorphism (3.8): \(\text{Hom}_{kG}(W^G, V) \cong \text{Hom}_{kH}(W, V^H)\).

**Orthogonality.** We now derive the celebrated orthogonality relations; they also follow from the more general orthogonality relations (2.14).

**Orthogonality Relations.** Assume that \(G\) is finite and that \(\text{char} \, k \nmid |G|\). Then, for \(S, T \in \text{Irr} \, kG\),

\[\langle \chi_S, \chi_T \rangle = \begin{cases} \dim_k D(S) \cdot 1_k & \text{if } S \cong T \\ 0_k & \text{if } S \ncong T \end{cases} .\]

**Proof.** This follows from (3.38) and Schur’s Lemma: \(\text{Hom}_{kG}(S, T) = 0\) if \(S \ncong T\) and \(\text{Hom}_{kG}(S, T) \cong D(S) = \text{End}_{kG}(S)\) if \(S \cong T\). \(\square\)

By the orthogonality relations, the irreducible characters \(\chi_S\) are pairwise orthogonal for the form \(\langle \cdot, \cdot \rangle\) and \(\langle \chi_S, \chi_S \rangle = 1_k\) holds whenever \(S\) is absolutely irreducible (Proposition 1.36). Thus, if \(k\) is a splitting field for \(G\), then \(\langle \chi_S \rangle_{S \in \text{Irr} \, kG}\) is an orthonormal basis of the inner product space \(\text{cf}_k(G)\) by Corollary 3.21(a).
Multiplicities and Irreducibility. The following proposition uses the orthogonality relations to derive information on the multiplicity $m(S, V)$ of $S \in \text{Irr}_k G$ in an arbitrary finite-dimensional representation $V$ and on the dimension of the $S$-homogeneous component $V(S)$. The proposition also gives a criterion for absolute irreducibility.

**Proposition 3.22.** Let $G$ be finite with $\text{char } k \nmid |G|$ and let $V \in \text{Rep}_\text{fin} \ k G$ and $S \in \text{Irr} \ k G$. Then:

(a) $\langle \chi_V, \chi_V \rangle = 1_k$ if $V$ is absolutely irreducible. The converse holds if $\text{char } k = 0$ or $\text{char } k \geq (\dim_k V)^2$.

(b) $\langle \chi_S, \chi_V \rangle = m(S, V) \dim_k D(S) \cdot 1_k$.

(c) $\dim_k V(S) \cdot 1_k = \dim_{D(S)} S \cdot \langle \chi_S, \chi_V \rangle$.

**Proof.** (a) The first assertion is clear from the orthogonality relations, as we have already remarked. For the converse, assume that $\langle \chi_V, \chi_V \rangle = 1_k$ and $\text{char } k = 0$ or $\text{char } k \geq (\dim_k V)^2$. The decomposition $V = \bigoplus_{S \in \text{Irr}_k G} S^\oplus m(S, V)$ implies

$$\chi_V = \sum_{S \in \text{Irr}_k G} m(S, V) \chi_S.$$

Therefore, $1_k = \langle \chi_V, \chi_V \rangle = \sum_{S \in \text{Irr}_k G} m(S, V)^2 \dim_k D(S) \cdot 1_k$ by the orthogonality relations. Since $\dim_k S = \dim_{D(S)} S \cdot \dim_k D(S) \geq \dim_k D(S)$, we have $\sum_{S \in \text{Irr}_k G} m(S, V)^2 \dim_k D(S) \leq (\dim_k V)^2$. In view of our hypothesis on $k$, it follows that the equality $1 = \sum_{S \in \text{Irr}_k G} m(S, V)^2 \dim_k D(S)$ does in fact hold in $\mathbb{Z}$. Therefore, $m(S, V)$ is nonzero for exactly one $S \in \text{Irr}_k G$ and we must also have $\dim_k D(S) = 1 = m(S, V)$. Thus, $V \cong S$ is absolutely irreducible.

(b) The above expression $\chi_V = \sum_{S \in \text{Irr}_k G} m(S, V) \chi_S$ in conjunction with the orthogonality relations gives

$$\langle \chi_S, \chi_V \rangle = \sum_{T \in \text{Irr}_k G} m(T, V) \langle \chi_S, \chi_T \rangle = m(S, V) \dim_k D(S) \cdot 1_k.$$

(c) From $V(S) \cong S^\oplus m(S, V)$ and $\dim_k S = \dim_{D(S)} S \cdot \dim_k D(S)$, we obtain

$$\dim_k V(S) = m(S, V) \dim_{D(S)} S \cdot \dim_k D(S) = \dim_{D(S)} S \cdot \langle \chi_S, \chi_V \rangle.$$

\[3.4.3. \text{ The Case of the Complex Numbers}\]

The inner product $\langle \cdot, \cdot \rangle$ is often replaced by a modified version when the base field $k$ is the field $\mathbb{C}$ of complex numbers. In this case, the formula $\chi_V^*(g) = \chi_V(g^{-1})$ in Lemma 3.20 can also be written as

$$\chi_V^*(g) = \overline{\chi_V(g)} \quad (g \in G)$$

with $\overline{\cdot}$ denoting complex conjugation. Indeed, the Jordan canonical form of the operator $g_V$ is a diagonal matrix having the eigenvalues $\zeta_i \in \mathbb{C}$ of $g_V$ along the
diagonal. The Jordan form of $g^{-1}$ has the inverses $\zeta_i^{-1}$ on the diagonal. Since all $\zeta_i$ are roots of unity, of order dividing the order of $g$, they satisfy $\zeta_i^{-1} = \overline{\zeta_i}$, which implies (3.40). The inner product of characters $\chi_V$ and $\chi_W$ can therefore also be written as follows:

$$\langle \chi_V, \chi_W \rangle = \frac{1}{|G|} \sum_{g \in G} \chi_V(g) \overline{\chi_W(g)}$$

(3.41)

It is common practice to define a form $\langle \cdot, \cdot \rangle : \text{cf}_{\mathbb{C}}(G) \times \text{cf}_{\mathbb{C}}(G) \rightarrow \mathbb{C}$ by

$$\langle \phi, \psi \rangle \overset{\text{def}}{=} \frac{1}{|G|} \sum_{g \in G} \phi(g) \overline{\psi(g)}$$

(3.42)

for $\phi, \psi \in \text{cf}_{\mathbb{C}}(G)$. This form is a Hermitian inner product on $\text{cf}_{\mathbb{C}}(G)$, that is, $\langle \cdot, \cdot \rangle$ is $\mathbb{C}$-linear in the first variable, it satisfies $\langle \phi, \psi \rangle = \overline{\langle \psi, \phi \rangle}$, and it is positive definite: $\langle \phi, \phi \rangle \in \mathbb{R}_{>0}$ for all $0 \neq \phi \in \text{cf}_{\mathbb{C}}(G)$. While $\langle \cdot, \cdot \rangle$ and $\langle \cdot, \cdot \rangle$ are of course different on $\text{cf}_{\mathbb{C}}(G)$, they do coincide on the subgroup that is spanned by the characters, taking integer values there by (3.38).

### 3.4.4. Primitive Central Idempotents of Group Algebras

We continue to assume that $G$ is finite with $\text{char} \mathbb{k} \nmid |G|$. Recall that $\mathbb{k}G$ is a symmetric algebra and, with $\lambda$ chosen as in (3.14), the Casimir element is $c_\lambda = \sum_{g \in G} g \otimes g^{-1}$ and $\gamma_\lambda(1) = |G|1$ (§3.1.6). Thus, if $\mathbb{k}$ is a splitting field for $G$, then Theorem 2.17 gives the formula $e(S) = \frac{\dim S}{|G|} \sum_{g \in G} \chi_S(g^{-1}) g$ for the central primitive idempotent of $S \in \text{Irr} \mathbb{k}G$. Here, we give an independent verification of this formula, without assuming $\mathbb{k}$ to be a splitting field, although our argument will be identical to the one in the proof of Theorem 2.17. Recall that the primitive central idempotents $e(S)$ of a semisimple algebra $A$ are characterized by the conditions

$$(1.47): e(S)_T = \delta_{S,T} \text{Id}_S$$

for $S, T \in \text{Irr} A$.

**Proposition 3.23.** Let $G$ be finite with $\text{char} \mathbb{k} \nmid |G|$ and let $S \in \text{Irr} \mathbb{k}G$. Then

$$e(S) = \frac{\dim D(S)_S}{|G|} \sum_{g \in G} \chi_S(g^{-1}) g .$$

**Proof.** Writing $e(S) = \sum_{g \in G} e_g g$ with $e_g \in \mathbb{k}$, our goal is to prove the equality $e_g = \frac{\dim D(S)_S}{|G|} \chi_S(g^{-1})$. By Example 3.13(b), the character of the regular representation of $\mathbb{k}G$ satisfies $\chi_{\text{reg}}(e(S) g^{-1}) = |G| e_g$, so we need to prove that

$$\chi_{\text{reg}}(e(S) g^{-1}) = \dim D(S)_S \cdot \chi_S(g^{-1}) .$$

But $\mathbb{k}G_{\text{reg}} \cong \bigoplus_{T \in \text{Irr} \mathbb{k}G} T^\oplus \dim D(T)_T$ by Wedderburn’s Structure Theorem and so

$$\chi_{\text{reg}}(e(S) g^{-1}) = \sum_{T \in \text{Irr} \mathbb{k}G} \dim D(T)_T \cdot \chi_T(e(S) g^{-1}) .$$
Finally, \((e(S)g^{-1})_T = \delta_{S,T}g_S^{-1}\) by \((1.47)\), and hence \(\chi_T(e(S)g^{-1}) = \delta_{S,T}\chi_S(g^{-1})\). Therefore, \(\chi_{\text{reg}}(e(S)g^{-1}) = \dim_{D(S)} S \cdot \chi_S(g^{-1})\), as desired. \(\square\)

The idempotent \(e(\mathbb{1})\) is identical to the idempotent \(e\) from Proposition 3.16. The “averaging” projection of a given \(V \in \text{Rep}_kG\) onto the \(G\)-invariants \(V^G\) in Proposition 3.16 generalizes to the projection \((1.49)\) of \(V = \bigoplus_{S \in \text{Irr}_kG} V(S)\) onto the \(S\)-homogeneous component \(V(S)\):

\[
e(S)_V : V \xrightarrow{\omega} V(S)
\]

\[
v \mapsto \frac{\dim_{D(S)} S}{|G|} \sum_{g \in G} \chi_S(g^{-1})g \cdot v
\]

In particular, if \(S = \mathbb{1}\) is a degree-1 representation, then we obtain the following projection of \(V\) onto the weight space \(V_\phi = \{v \in V \mid g \cdot v = \phi(g) v \text{ for all } g \in G\}:\n
\[
V \xrightarrow{\omega} V_\phi
\]

\[
v \mapsto \frac{1}{|G|} \sum_{g \in G} \phi(g^{-1})g \cdot v
\]

**Exercises for Section 3.4**

*Without any mention to the contrary, the group \(G\) and the field \(k\) are arbitrary in these exercises.*

3.4.1 (\(kS_3\) in characteristics 2 and 3). We know by Maschke’s Theorem that the group algebra \(kS_3\) is fails to be semisimple exactly for \(\text{char } k = 2\) and 3. Writing \(S_3 = \langle x, y \mid y^2 = x^3 = 1, xy = yx^2 \rangle\), show:

(a) If \(\text{char } k = 3\) then \(\text{Irr}_kS_3 = \{\mathbb{1}, \text{sgn}\}\) and \(\text{rad}_kS_3 = kS_3(x - 1)\).

(b) If \(\text{char } k = 2\) then \(\text{Irr}_kS_3 = \{\mathbb{1}, V_2\}\), where \(V_2\) is the standard representation of \(S_3\) (see Exercise 3.2.3), and \(\text{rad}_kS_3 = k\sigma_{S_3}\) with \(\sigma_{S_3} = \sum_{g \in S_3} g\).

3.4.2 (Standard involution and semisimplicity). Write the standard involution \((3.28)\) of \(kG\) as \(a^* = S(a)\) for \(a \in kG\) and recall that \((ab)^* = b^*a^*\) and \(a^{**} = a\) for all \(a, b \in kG\). Assuming that \(k\) is a subfield of \(\mathbb{R}\), prove:

(a) \(aa^* = 0\) for \(a \in kG\) implies \(a = 0\). Also, if \(aa^*a = 0\) then \(a = 0\).

(b) All finite-dimensional subalgebras of \(kG\) that are stable under \(^*\) are semisimple. (Use Theorem 1.39.)

Use (b) and the fact that finite-dimensional semisimple algebras over a field of characteristic 0 are separable (Exercises 1.4.10 and Exercises 1.5.6.) to prove:

(c) If \(G\) is finite and \(k\) is any field with \(\text{char } k = 0\), then \(kG\) is semisimple.
3.4.3 (Relative trace map and a relative version of Maschke’s Theorem). Let $H$ be a finite-index subgroup of $G$ such that $\text{char} \, \mathbb{k} \nmid |G : H|$. 

(a) For $V \in \text{Rep} \mathbb{k}G$, define a $\mathbb{k}$-linear map $\tau_H^G : V^H \to V^G$ by

$$\tau_H^G(v) = |G : H|^{-1} \sum_{g \in G/H} g \cdot v \quad (v \in V^H)$$

Show that $\tau_H^G$ is independent of the choice of the transversal for $G/H$, takes values in $V^G$, and is the identity on $V^G$. The map $\tau_H^G$ is called the relative trace map.

(b) Let $0 \to U \to V \to W \to 0$ be a short exact sequence in $\text{Rep} \mathbb{k}G$. Mimic the proof of Maschke’s Theorem to show that if the $0 \to U \downarrow_H \to V \downarrow_H \to W \downarrow_H \to 0$ splits in $\text{Rep} \mathbb{k}H$ (Exercise 1.1.2), then the given sequence splits in $\text{Rep} \mathbb{k}G$.

3.4.4 (Characters and conjugacy). Consider the following statements, with $x, y \in G$:

(i) $G_x = G_y$, that is, $x$ and $y$ are conjugate in $G$;
(ii) $\chi_V(x) = \chi_V(y)$ for all $V \in \text{Rep}_{\text{fin}} \mathbb{k}G$;
(iii) $\chi_S(x) = \chi_S(y)$ for all $S \in \text{Irr}_{\text{fin}} \mathbb{k}G$.

Show that (i) $\iff$ (ii) $\iff$ (iii). For $G$ finite and $\mathbb{k}$ a splitting field for $G$ with $\text{char} \, \mathbb{k} \nmid |G|$, show that all three statements are equivalent.

3.4.5 (Values of complex characters). Let $G$ be finite. A complex character of $G$ is a character $\chi = \chi_V$ for some $V \in \text{Rep}_{\text{fin}} \mathbb{C}G$; if $V$ is irreducible, then $\chi$ is called an irreducible complex character. For $g \in G$, show:

(a) $\chi(g) \in \mathbb{R}$ for every (irreducible) complex character $\chi$ of $G$ if and only if $g$ is conjugate to $g^{-1}$ in $G$.

(b) $\chi(g) \in \mathbb{Q}$ for every (irreducible) complex character $\chi$ if and only if $g$ is conjugate to $g^m$ for every integer $m$ with $(m, |G|) = 1$.

(c) $|\chi(g)| \leq \chi(1)$ for every complex character $\chi = \chi_V$ and equality occurs precisely if $g_V$ is a scalar operator. (Use the triangle inequality.)

(d) If $G$ is non-abelian simple, then $|\chi(g)| < \chi(1)$ for every complex character $\chi = \chi_V$ with $V \neq 1 \oplus_d (d = \dim_{\mathbb{C}} V)$ and every $1 \neq g \in G$.

3.4.6 (The $p$-core of a finite group). Let $G$ be finite and assume that char $\mathbb{k} = p > 0$. The $p$-core $\mathcal{O}_p(G)$ of $G$, by definition, is the unique largest normal $p$-subgroup of $G$ or, equivalently, the intersection of all Sylow $p$-subgroups. Show that $\mathcal{O}_p(G) = \{g \in G \mid g_S = \text{Id}_S \text{ for all } S \in \text{Irr} \mathbb{k}G\} = G \cap (1 + \text{rad} \mathbb{k}G)$.

3.4.7 (Column orthogonality relations). Let $G$ be finite and assume that $\mathbb{k}$ is a splitting field for $G$ with $\text{char} \, \mathbb{k} \nmid |G|$. Prove:

$$\sum_{S \in \text{Irr} \mathbb{k}G} \chi_S(g^{-1}) \chi_S(h) = \begin{cases} |C_G(g)| & \text{if } g, h \in G \text{ are conjugate} \\
0 & \text{otherwise} \end{cases}$$

Here, $C_G(g)$ denotes the centralizer of $g$ in $G$. 

148

3. Groups and Group Algebras
3.4.8 (Generalized orthogonality relations). Let $G$ be a finite group with $\text{char} \neq |G|$. Use the fact that the primitive central idempotents satisfy $e(S)e(T) = \delta_{S,T}e(S)$ to prove the following relations:

$$\frac{1}{|G|} \sum_{g \in G} \chi_S(gh)\chi_T(g^{-1}) = \delta_{S,T} \frac{1}{\dim D(S)} S \chi_S(h)$$

For $h = 1$, this reduces to the ordinary orthogonality relations.

3.4.9 (Irreducibility and inner products). Give an example of a finite group $G$ with $\text{char} \neq |G|$ and a non-irreducible $V \in \text{Rep}_{\text{fin}} kG$ such that $(\chi_V, \chi_V) = 1_k$. Thus, that the hypothesis on $\text{char} k$ in Proposition 3.22(a) cannot be omitted.

3.4.10 (Complete reducibility of the adjoint representation of $kG$). Consider the adjoint representation of $(kG)_{ad}$ of a finite group $G$; see Example 3.13(c). Use Exercise 3.3.8 and Maschke’s Theorem to show that the following are equivalent:

(i) $(kG)_{ad}$ is completely reducible;

(ii) $\text{char} k$ does not divide the order of $G/Z$;

(iii) $\text{char} k$ does not divide the size of any conjugacy class of $G$.

(This exercise was worked out together with Don Passman.)

3.4.11 (Isomorphism of finite $G$-sets and permutation representations). Let $X$ and $Y$ be finite $G$-sets. For each subgroup $H \leq G$, put $X^H = \{ x \in X \mid h.x = x \text{ for all } h \in H \}$ and likewise for $Y^H$.

(a) Show that $X$ and $Y$ are isomorphic in the category $G$-Sets (Exercise 3.3.5) if and only if $\#X^H = \#Y^H$ for all subgroups $H \leq G$.

(b) Assuming that $\text{char} k = 0$ show that $kX \cong kY$ in $\text{Rep} kG$ if and only if $\#X^H = \#Y^H$ for all cyclic subgroups $H \leq G$.

3.4.12 (Hermitian inner products). A \textit{Hermitian inner product} on a $\mathbb{C}$-vector space $V$ is a map $H: V \times V \to \mathbb{C}$ such that

(i) the map $H(\cdot, \cdot): V \to \mathbb{C}$ is a $\mathbb{C}$-linear form for each $v \in V$;

(ii) $H(v, w) = \overline{H(w, v)}$ holds for all $v, w \in V$ (\$ = \text{complex conjugation})$; and

(iii) $H(v, v) > 0$ for all $0 \neq v \in V$.

Note that (ii) implies that $H(v, v) \in \mathbb{R}$ for all $v \in V$; so (iii) makes sense. Now let $G$ be finite and let $V \in \text{Rep}_{\text{fin}} \mathbb{C}G$. Prove:

(a) There exists a Hermitian inner product on $V$ that is $G$-invariant, that is, $H(g.v, g.w) = H(v, w)$ holds for all $g \in G$ and $v, w \in V$.

(b) Let $V$ be irreducible. Then the inner product in (a) is \textit{unique} up to a positive real factor: if $H'$ is another Hermitian inner product on $V$ that is preserved by $G$, then $H' = \lambda H$ for some $\lambda \in \mathbb{R}_{>0}$.
3.5. Further Examples

In this section, we assume that \( \text{char}\ k = 0 \). We will be concerned with certain finite-dimensional representations of the symmetric groups \( S_n \) \((n \geq 3)\) and some of their subgroups. In particular, by Maschke’s Theorem, all representations under consideration will be completely reducible and therefore determined, up to equivalence, by their character (Theorem 1.45).

3.5.1. Exterior Powers of the Standard Representation of \( S_n \)

We have seen (§3.2.4) that \( kS_n \) has two degree-1 representations, \( \mathbb{1} \) and \( \text{sgn} \), and an irreducible representation of degree \( n - 1 \), the standard representation \( V_{n-1} \) (Exercise 3.2.3). Looking for new representations, we may try the “sign twist” of a given representation \( V \):

\[
V^\pm \text{ def } = \text{sgn} \otimes V
\]

Since \( \chi_{V^\pm} = \text{sgn} \chi_V \) (Lemma 3.20), we know that \( V^\pm \not\equiv V \) if and only if \( \chi_V(s) \neq 0 \) for some odd permutation \( s \in S_n \). Furthermore, if \( V \) is irreducible, then it is easy to see that \( V^\pm \) is irreducible as well (Exercise 3.3.11). In principle, we could also consider the dual representation \( V^* \). However, this yields nothing new for the symmetric groups:

**Lemma 3.24.** All finite-dimensional representations of \( kS_n \) are self-dual.

**Proof.** This is a consequence of the fact that each \( s \in S_n \) is conjugate to its inverse, because \( s \) and \( s^{-1} \) have the same cycle type. In view of Lemma 3.20, it follows that \( \chi_V = \chi_{V^*} \) holds for each \( V \in \text{Rep}_\text{fin} kS_n \) and so \( V \equiv V^* \). \hfill \( \Box \)

Our goal in this subsection is to prove the following proposition by an elementary if somewhat lengthy inner product calculation following [78, §3.2].

**Proposition 3.25.** The exterior powers \( \Lambda^k V_{n-1} \) \((0 \leq k \leq n - 1)\) of the standard representation \( V_{n-1} \) are all (absolutely) irreducible and pairwise non-equivalent.

Before proceeding to prove the proposition in general, let us illustrate the result by discussing some special cases. First, \( \Lambda^0 V_{n-1} \equiv \mathbb{1} \) is evidently irreducible and we also know that \( \Lambda^1 V_{n-1} = V_{n-1} \) is irreducible. Next, let \( M_n = \bigoplus_{i=1}^n \mathbb{1} \otimes b_i \) be the standard permutation representation of \( S_n \), with \( s.b_i = b_{s(i)} \) for \( s \in S_n \), and recall that \( M_n/V_{n-1} \equiv \mathbb{1} \) (§3.2.4). By complete reducibility, we obtain the decomposition \( M_n \equiv \mathbb{1} \oplus V_{n-1} \). It is easy to see that \( \det M_n = \text{sgn} \). Therefore, we also have \( \det V_{n-1} = \text{sgn} \) and so \( \Lambda^{n-1} V_{n-1} \equiv \text{sgn} \) by (3.32), which is clearly irreducible. From Exercise 3.3.14(b) and Lemma 3.24, we further obtain

\[
(3.45) \quad \Lambda^{n-1-k} V_{n-1} \equiv (\Lambda^k V_{n-1})^\pm
\]

for all \( k \). In particular, \( \Lambda^{n-2} V_{n-1} \equiv V_{n-1}^\pm \), which is irreducible as well.
Proof of Proposition 3.25. First, we check non-equivalence. The representations $\Lambda^k V_{n-1}$ and $\Lambda^{k'} V_{n-1}$ have the same dimension, $\binom{n-1}{k} = \binom{n-1}{k'}$, if and only if $k = k'$ or $k + k' = n - 1$. In the latter case, $\Lambda^k V_{n-1} \cong (\Lambda^k V_{n-1})^\pm$ by (3.45). Therefore, it suffices to show that $\Lambda^k V_{n-1} \neq (\Lambda^k V_{n-1})^\pm$ for $2k \neq n - 1$. Put

$$\chi_k :=\chi_{\Lambda^k V_{n-1}}.$$

We need to check that $\chi_k(s) \neq 0$ for some odd permutation $s \in S_n$. Let $s$ be a 2-cycle. Then $s$ acts as a reflection on $M_n$: the operator $s M_n$ has a simple eigenvalue $-1$ and the remaining eigenvalues are all 1. From the isomorphism $M_n \cong 1 \oplus V_{n-1}$, we see that the same holds for the operator $s V_{n-1}$. Therefore, we may choose a basis $v_1, \ldots, v_{n-1}$ of $V_{n-1}$ with $s v_i = v_i$ for $1 \leq i \leq n - 2$ but $s v_{n-1} = -v_{n-1}$. By (1.13), a basis of $\Lambda^k V_{n-1}$ is given by the elements $\wedge v_I = v_{i_1} \wedge v_{i_2} \wedge \cdots \wedge v_{i_k}$ with $I = \{i_1, i_2, \ldots, i_k\}$ a $k$-element subset of $[n - 1] = \{1, 2, \ldots, n - 1\}$ in increasing order: $i_1 < i_2 < \cdots < i_k$. Since $s. \wedge v_I = \wedge v_I$ if $n - 1 \notin I$ and $s. \wedge v_I = -\wedge v_I$ otherwise, we obtain

$$\chi_k(s) = \begin{cases} 1 & \text{if } k = 0, \\ \binom{n-2}{k} - \binom{n-2}{k-1} & \text{if } 1 \leq k \leq n - 1. \end{cases}$$

Finally, $\binom{n-2}{k} = \binom{n-2}{k-1}$ if and only if $2k = n - 1$, which we have ruled out. This proves non-equivalence of the various $\Lambda^k V_{n-1}$.

It remains to prove absolute irreducibility of $\Lambda^k V_{n-1}$. By Proposition 3.22(a), this is equivalent to the condition $(\chi_k, \chi_k) = 1$. The case $k = 0$ being trivial, we will assume that $k \geq 1$. Since $M_n \cong 1 \oplus V_{n-1}$, we have $\Lambda M_n \cong \Lambda \otimes \Lambda V_{n-1}$ in $\text{Rep} \mathcal{K} S_n$ (Exercise 3.3.14), and hence

$$\Lambda^k M_n \cong \bigoplus_{r+s=k} \Lambda^r 1 \otimes \Lambda^s V_{n-1} \cong \Lambda^k V_{n-1} \oplus \Lambda^{k-1} V_{n-1}.$$

Putting $\chi := \chi_{\Lambda^k M_n}$, we obtain $(\chi, \chi) = (\chi_{k-1}, \chi_{k-1}) + 2(\chi_{k-1}, \chi_k) + (\chi_k, \chi_k)$. Since the first and last term on the right are positive integers and the middle term is non-negative, our desired conclusion $(\chi_k, \chi_k) = 1$ will follow if we can show that

$$(\chi, \chi) = 2.$$

To compute $\chi$, we use the basis $(\wedge b_I)$ of $\Lambda^k M_n$, where $\wedge b_I = b_{i_1} \wedge \cdots \wedge b_{i_k}$ and $I = \{i_1, \ldots, i_k\}$ is a $k$-element subset of $[n]$ in increasing order. Each $s \in S_n$ permutes the basis $(\wedge b_I)$ up to a $\pm$ sign by (1.12). The diagonal $(I, I)$-entry of the matrix of $s \Lambda^k M_n$ with respect to this basis is given by

$${s}_I := \begin{cases} 0 & \text{if } s(I) \neq I, \\ \text{sgn}(s_I) & \text{if } s(I) = I. \end{cases}$$
Thus, \( \chi(s) = \sum_{I} \{s\}_I = \chi(s^{-1}) \) and so
\[
\langle \chi, \chi \rangle = \frac{1}{n!} \sum_{s \in S_n} \left( \sum_{I} \{s\}_I \right)^2
\]
\[
= \frac{1}{n!} \sum_{s \in S_n} \sum_{I,J} \{s\}_I \{s\}_J
\]
\[
= \frac{1}{n!} \sum_{I,J} \sum_{s \in S_n} \{s\}_I \{s\}_J
\]
\[
= \frac{1}{n!} \sum_{I,J} \sum_{s \in S_n} sgn(s|_I) sgn(s|_J).
\]

Here \( I \) and \( J \) run over the \( k \)-element subsets of \([n]\) and \( Y_{I,J} \) consists of those \( s \in S_n \) that stabilize both \( I \) and \( J \) or, equivalently, all pieces of the partition \([n] = (I \cup J) \setminus I^' \cup J^' \cup (I \cap J)\), where \( \setminus \) denotes the complement and \( .^' = \cdot \setminus (I \cap J) \). Thus, \( Y_{I,J} \) is a subgroup\(^9\) of \( S_n \) with the following structure:
\[
Y_{I,J} \equiv S_{(I \cup J) \setminus .} \times S_{I^'} \times S_{J^'} \times S_{I \cap J}.
\]

Since \( sgn(s|_I) = sgn(s|_{I^'}) sgn(s|_{I \cap J}) \) for \( s \in Y_{I,J} \) and likewise for \( sgn(s|_J) \), we obtain
\[
\langle \chi, \chi \rangle = \frac{1}{n!} \sum_{I,J} \sum_{s \in Y_{I,J}} sgn(s|_{I^'}) sgn(s|_{J^'}) sgn(s|_{I \cap J})^2
\]
\[
= \frac{1}{n!} \sum_{I,J} \sum_{s \in Y_{I,J}} sgn(s|_{I^'}) sgn(s|_{J^'})
\]
\[
= \frac{1}{n!} \sum_{I,J} |S_{(I \cup J) \setminus .}| |S_{I \cap J}| \sum_{\alpha \in S_{I^'}} sgn(\alpha) sgn(\beta)
\]
\[
= \frac{1}{n!} \sum_{I,J} |S_{(I \cup J) \setminus .}| |S_{I \cap J}| \left( \sum_{\alpha \in S_{I^'}} sgn(\alpha) \right)^2.
\]

The last equality above uses the fact that \( I^' \) and \( J^' \) have the same number of elements; so \( S_{I^'} \) and \( S_{J^'} \) are symmetric groups of the same degree and hence \( \sum_{\beta \in S_{I^'}} sgn(\beta) = \sum_{\alpha \in S_{I^'}} sgn(\alpha) \). If \( J^' \) has at least two elements, then this sum is 0; otherwise the sum equals 1. Therefore, the only nonzero contributions to the last expression in (3.46) come from the following two cases.

Case 1: \( I = J \). Then the \((I, J)\)-summand of the last sum in (3.46) is \((n - k)! k!\). Since there are a total of \( \binom{n}{k} \) summands of this type, their combined contribution is
\[
\frac{1}{n!} (n - k)! k! \binom{n}{k} = 1.
\]

\(^9\)Subgroups of this form are called Young subgroups after Alfred Young (1873–1940); they will be considered more systematically later (§3.8.2).
Thus, we finally obtain that these summands is \((\chi, \chi) = 2\) as was our goal.

3.5.2. The Groups \(S_4\) and \(S_5\)

The irreducible representations of \(kS_3\) and the character table have already been determined (Example 3.12). Recall in particular that \(\text{Irr} kS_3 = \{1, \text{sgn}, V_2\}\). Now we shall do the same for \(S_4\) and \(S_5\). Before we enter into the specifics, let us remind ourselves of some basic facts concerning the symmetric groups \(S_n\) in general.

Conjugacy Classes of \(S_n\). The conjugacy classes of \(S_n\) are in one-to-one correspondence with the partitions of \(n\), that is, sequences \(\lambda = (\lambda_1 \geq \lambda_2 \geq \ldots)\) with \(\lambda_i \in \mathbb{Z}_+\) and \(\sum \lambda_i = n\). Specifically, the partition \(\lambda\) corresponds to the conjugacy class consisting of all \(s \in S_n\) whose orbits in \([n]\) have sizes \(\lambda_1, \lambda_2, \ldots\); equivalently, \(s\) is a product of disjoint cycles of lengths \(\lambda_1, \lambda_2, \ldots\). The size of the conjugacy class corresponding to \(\lambda\) is given by

\[
\frac{n!}{\prod_i \lambda_i^{|m_i(\lambda_i)|} m_i(\lambda_i)!}
\]

where \(m_i(\lambda_i) = \#\{j \mid \lambda_j = i\}\) (e.g., [191, Proposition 1.3.2]).

Representations of \(S_4\). We can take advantage of the fact that \(S_4\) has the group theoretical structure of a semidirect product:

\[(3.47) \quad S_4 = V_4 \rtimes S_3\]

with \(V_4 = \{(1), (1 2)(3 4), (1 3)(2 4), (1 4)(2 3)\} \cong C_2 \times C_2\), the Klein 4-group, and with \(S_3\) being identified with the stabilizer of 4 in \(S_4\). Thus, there is a group epimorphism \(f : S_4 \rightarrow S_3\) with \(\text{Ker } f = V_4\) and \(f|_{S_3} = \text{Id}\). Inflation along the algebra map \(\phi = kF : kS_4 \rightarrow kS_3\) allows us to view \(\text{Irr} kS_3 = \{1, \text{sgn}, V_2\}\) as a subset of \(\text{Irr} kS_4\). Besides the obvious degree-1 representations, 1 and sgn, this yields the representation \(V_2\), inflated from \(S_3\) to \(S_4\). We will denote this representation by \(\tilde{V}_2\). By Proposition 3.25, we also have the irreducible representations \(V_3\) and \(\wedge^2 V_3 \cong (V_3)^2\); see (3.45). Thus we have found five non-equivalent absolutely irreducible representations of \(kS_4\), having degrees 1, 1, 2, 3 and 3. Since the squares of these degrees add up to the order of \(S_4\), we know by Wedderburn’s Structure Theorem that there are no further irreducible representations. Alternatively, since \(S_4\) has five conjugacy classes, this could also be deduced from Proposition 3.6. Table 3.3 records the character table; see Example 3.13(d) for the character of the standard representation \(V_3\).
Representations of $S_5$. Unfortunately, no mileage is to be gotten from inflation here due to the scarcity of normal subgroups in $S_5$. However, Proposition 3.25 provides us with the following five non-equivalent absolutely irreducible representations: $1$, $V_4$, $\Lambda^2 V_4$, $\Lambda^3 V_4 = V_4^\pm$ and $\Lambda^4 V_4 = \text{sgn}$. Since the sum of the squares of their degrees is short of the order of $S_5$, there must be further irreducible representations by Wedderburn’s Structure Theorem. We shall later discuss a general result (Theorem 10.13) that, in the case of $S_5$, guarantees that all irreducible representations must occur as constituents of tensor powers of $V_4$. So let us investigate $V_4 \otimes V_4$. First, the isomorphism $(V_4 \otimes V_4)^{S_5} \cong (V_4 \otimes V_4)^{S_5} \cong \text{End}_{S_5}(V_4) \cong \mathbb{k}$ tells us that $1$ is an irreducible constituent of $V_4 \otimes V_4$ with multiplicity $1$. Next, $\Lambda^2 V_4$ is also an irreducible constituent of $V_4 \otimes V_4$ and $\chi_{\Lambda^2 V_4}(s) = \frac{1}{2}(\chi_{V_4}(s)^2 - \chi_{V_4}(s^2))$ for $s \in S_5$. Let us accept these facts for now; they will be proved in much greater generality in (3.63) and (3.67) below. Since $\chi_{V_4}$ is known by Example 3.13(d), we obtain the following table of values for the characters of $V_4$ and $\Lambda^2 V_4$:

<table>
<thead>
<tr>
<th>classes</th>
<th>(1)</th>
<th>(1 2)</th>
<th>(1 2 3)</th>
<th>(1 2 3 4)</th>
<th>(1 2)(3 4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>sizes</td>
<td>1</td>
<td>6</td>
<td>8</td>
<td>6</td>
<td>3</td>
</tr>
<tr>
<td>$1$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$\text{sgn}$</td>
<td>1</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>$\chi_{\Lambda^2 V_4}$</td>
<td>2</td>
<td>0</td>
<td>-1</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>$\chi_{V_4}$</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td>$\chi_{\Lambda^2 V_4}$</td>
<td>3</td>
<td>-1</td>
<td>0</td>
<td>1</td>
<td>-1</td>
</tr>
</tbody>
</table>

Using these values and the fact that $\chi_{V_4} \chi_{V_4} = \chi_{V_4}^2$ (Lemma 3.20), we compute

$$ \left( \chi_{V_4} \cdot \chi_{V_4} \right) = \frac{1}{120} \sum_{s \in S_5} (1 \cdot 4^2 + 10 \cdot 2^3 + 20 \cdot 1^3 + 24 \cdot (-1)^3 + 20 \cdot (-1)^3) = 1. $$

This shows that $V_4$ is a constituent of $V_4^\otimes 2$, with multiplicity $1$ (Proposition 3.22). Letting $W$ denote the sum of the other irreducible constituents, we can write $V_4^\otimes 2 \cong 1 \oplus V_4 \oplus \Lambda^2 V_4 \oplus W$.

The character $\chi_W = \chi_{V_4}^2 - \chi_{V_4} - \chi_{\Lambda^2 V_4}$ along with the character of $W^\pm = \text{sgn} \otimes W$ are given by the following table:

Table 3.3. Character table of $S_4$ (char $\mathbb{k} = 0$)
Further Examples

It is a simple matter to check that \((\chi_W, \chi_W) = (\chi_{W^\pm}, \chi_{W^\pm}) = 1\). Hence \(W\) and \(W^\pm\) are both absolutely irreducible by Proposition 3.22, and they are not equivalent to each other or to any of the prior irreducible representations. Altogether we have now found seven irreducible representations, which are all in fact absolutely irreducible. Since there are also seven conjugacy classes, we have found all irreducible representations of \(kS_5\) by Proposition 3.6. For completeness, we record the entire character table as Table 3.4.

Table 3.4. Character table of \(S_5\) (char \(k = 0\))

<table>
<thead>
<tr>
<th>classes</th>
<th>(1)</th>
<th>(1 2)</th>
<th>(1 2 3)</th>
<th>(1 2 3 4)</th>
<th>(1 2 3 4 5)</th>
<th>(1 2)(3 4)</th>
<th>(1 2 3)(4 5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>sizes</td>
<td>1</td>
<td>10</td>
<td>20</td>
<td>30</td>
<td>24</td>
<td>15</td>
<td>20</td>
</tr>
<tr>
<td>(\chi_W)</td>
<td>5</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>(\chi_{W^\pm})</td>
<td>5</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>-1</td>
</tr>
</tbody>
</table>

3.5.3. The Alternating Groups \(A_4\) and \(A_5\)

It is a notable fact that our arbitrary field \(k\) of characteristic 0 is a splitting field for \(kS_n\); this was observed above for \(n \leq 5\) and it is actually true for all \(n\) as we shall see in Section 4.5. However, the corresponding fact fails to hold for the alternating groups \(A_n\). Indeed, even the group \(A_3 \cong C_3\) requires all third roots of unity to be contained in \(k\) if \(kA_3\) is to be split (§3.2.1). Therefore, we shall assume in this subsection, in addition to char \(k = 0\), that \(k\) is algebraically closed.

Conjugacy Classes of \(A_n\). The conjugacy classes of \(A_n\) are quickly sorted out starting from those of \(S_n\). Clearly, for any permutation \(s \in S_n\), the \(A_n\)-conjugacy class \(A_n s\) is contained in the \(S_n\)-conjugacy class \(S_n s\). By simple general considerations about restricting group actions to subgroups of index 2 (Exercise 3.5.1), there are two possibilities:
If $C_{S^n}(s) \not\subseteq A_n$, then $A^n_s = S^n_s$; otherwise, $S^n_s$ splits into two $A_n$-conjugacy classes of equal size.

It is also easy to see that the first case occurs precisely if $s$ has at least one orbit of even size or at least two orbits of the same size, and the second if the orbit sizes of $s$ are all odd and distinct (Exercise 3.5.2).

**Representations of $A_4$.** The semidirect product decomposition (3.47) of $S_4$ yields the following decomposition of the alternating group $A_4$:

\[ A_4 = V_4 \rtimes C_3, \]

with $V_4 = \{(1), (12)(34), (13)(24), (14)(23)\}$ and $C_3 = A_3 = \langle (123) \rangle$. By inflation from $C_3$, we obtain three degree-1 representations of $A_4$: the trivial representation $1$, the representation $\phi: A_4 \to k^\times$ that sends $(123)$ to a fixed primitive third root of unity $\zeta_3 \in k$, and $\phi^2$. Since $A_4^{ab} \cong C_3$, there are no further degree-1 representations. The squares of the degrees of all irreducible representations need to add up to $|A_4| = 12$ by Corollary 3.21(b); so we need one more irreducible representation, necessarily of degree 3. For this, we try the restriction of the standard $S_4$-representation $V_3$ to $A_4$. While there is no a priori guarantee that the restriction $V_3 \downarrow_{A_4}$ remains irreducible, the following inner product computation shows that this is indeed the case—note that only the classes of $(1)$, $(123)$ and $(12)(34)$ in Table 3.3 give conjugacy classes of $A_4$ and $\chi_{V_3}$ vanishes on $S_4$-conjugacy class of $(123)$, which breaks up into two $A_4$-classes of size 4:

\[
\left( \chi_{V_3} \downarrow_{A_4}, \chi_{V_3} \downarrow_{A_4} \right) = \frac{1}{12} (1 \cdot 3^2 + 4 \cdot 0^2 + 4 \cdot 0^2 + 3 \cdot (-1)^2) = 1.
\]

Thus, we have found all irreducible representations of $kA_4$. The character table (Table 3.5) is easily extracted from the character table of $S_4$ (Table 3.3).

<table>
<thead>
<tr>
<th>classes sizes</th>
<th>(1)</th>
<th>(1 2 3)</th>
<th>(1 3 2)</th>
<th>(1 2)(3 4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$\phi$</td>
<td>1</td>
<td>$\zeta_3$</td>
<td>$\zeta_3^2$</td>
<td>1</td>
</tr>
<tr>
<td>$\phi^2$</td>
<td>1</td>
<td>$\bar{\zeta}_3$</td>
<td>$\zeta_3$</td>
<td>1</td>
</tr>
<tr>
<td>$\chi_{V_3 \downarrow_{A_4}}$</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>-1</td>
</tr>
</tbody>
</table>

Table 3.5. Character table of $A_4$ (char $k = 0$, $\zeta_3 \in k^\times$ a primitive third root of unity)

**Subgroups of Index 2.** By Corollary 3.5(a), all irreducible representations of $A_5$ must arise as constituents of restrictions of suitable irreducible representations of
Since the signed versions of irreducible representations of $S_5$ have the same restrictions to $A_5$, we must look at

$$1\downarrow_{A_5}, V_4\downarrow_{A_5}, W\downarrow_{A_5} \quad \text{and} \quad \lambda^2 V_4\downarrow_{A_5}.$$ 

The following lemma gives a simple criterion for deciding which of these restrictions remain irreducible. The process of restricting irreducible representations to normal subgroups will be addressed in greater generality in Clifford’s Theorem (§3.6.4).

**Lemma 3.26.** Let $G$ be arbitrary and let $H$ be a subgroup of $G$ with $|G : H| = 2$. Then, for every $V \in \text{Irr}_{\text{fin}} kG$, the restriction $V\downarrow_H$ is either irreducible or a direct sum of two irreducible $kH$-representations of equal dimension. The former case happens if and only if $\chi_V$ does not vanish on $G \setminus H$.

**Proof.** Note that, for any $x \in G \setminus H$, we have $G = H \cup xH$ and $xH = Hx$. Now let $W$ be some irreducible subrepresentation of $V\downarrow_H$. Then

$$V = kG.W = (kH + xkH).W = W + x.W$$

Since $Hx = xH$, it follows that $x.W$ is a subrepresentation of $V\downarrow_H$. It is readily seen that $x.W$ is in fact irreducible, because $W$ is so, and $x.W$ clearly has the same dimension as $W$. We conclude that either $V\downarrow_H = W$ is irreducible or $V\downarrow_H = W \oplus x.W$ is the direct sum of two irreducible $kH$-representations of equal dimension. In the latter case, we have $\chi_V(x) = 0$, because the matrix of $x.V$ with respect to a basis of $V$ that is assembled from bases of $W$ and $x.W$ has two blocks of 0-matrices of size $\dim_k W$ along the diagonal. Therefore, $\chi_V$ vanishes on $G \setminus H$ if $V\downarrow_H$ is not irreducible. Conversely, if $\chi_V$ vanishes on $G \setminus H$, then the following computation shows that $V\downarrow_H$ is not irreducible:

$$\left(\chi_{V\downarrow_H}, \chi_{V\downarrow_H}\right) = \frac{1}{|H|} \sum_{h \in H} \chi_V(h)\chi_V(h^{-1}) = \frac{2}{|G|} \sum_{g \in G} \chi_V(g)\chi_V(g^{-1}) = 2\left(\chi_V, \chi_V\right) = 2.$$ 

Note that the above proof only needs $k$ to be algebraically closed of characteristic not dividing $|G|$.

**Representations of $A_5$.** Observe that the characters of the $S_5$-representations $V_4$ and $W$ in Table 3.4 have nonzero value on the transposition $(1 2)$, and hence both representations remain irreducible upon restriction to $A_5$ by Lemma 3.26. The character of $\Lambda^2 V_4$, on the other hand vanishes on $S_5 \setminus A_5$; so we must have

$$\Lambda^2 V_4\downarrow_{A_5} = X \oplus X’$$

for two 3-dimensional irreducible $A_5$-representations $X$ and $X’$. These representations along with $1$, $V_4\downarrow_{A_5}$ and $W\downarrow_{A_5}$ will form a complete set of irreducible
representations of $\mathcal{A}_5$ by Corollary 3.5(a). In order to determine $\chi_X$ and $\chi_{X'}$, we extract the following information from the character table of $S_5$ (Table 3.4):

<table>
<thead>
<tr>
<th>classes</th>
<th>(1)</th>
<th>(1 2 3)</th>
<th>(1 2 3 4 5)</th>
<th>(2 1 3 4 5)</th>
<th>(1 2)(3 4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>sizes</td>
<td>1</td>
<td>20</td>
<td>12</td>
<td>12</td>
<td>15</td>
</tr>
<tr>
<td>$\chi_{V_4 \downarrow A_5}$</td>
<td>4</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>0</td>
</tr>
<tr>
<td>$\chi_{W_4 \downarrow A_5}$</td>
<td>5</td>
<td>-1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>$\chi_{\Lambda^2 V_4 \downarrow A_5}$</td>
<td>6</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>-2</td>
</tr>
<tr>
<td>$\chi_X$</td>
<td>3</td>
<td>$\alpha$</td>
<td>$\beta$</td>
<td>$\gamma$</td>
<td>$\delta$</td>
</tr>
</tbody>
</table>

The orthogonality relations give the following system of equations for the unknowns $\alpha, \beta, \gamma$ and $\delta$:

\[
0 = (\mathbb{1}, \chi_X) = \frac{1}{60}(3 + 20\alpha + 12\beta + 12\gamma + 15\delta)
\]

\[
0 = (\chi_{V_4 \downarrow A_5}, \chi_X) = \frac{1}{60}(3 \cdot 4 + 20\alpha - 12\beta - 12\gamma)
\]

\[
0 = (\chi_{W_4 \downarrow A_5}, \chi_X) = \frac{1}{60}(3 \cdot 5 - 20\alpha + 15\delta)
\]

\[
1 = (\chi_X, \chi_X) = \frac{1}{60}(3^2 + 20\alpha^2 + 12\beta^2 + 12\gamma^2 + 15\delta^2)
\]

Here we have used the fact that each element of $\mathcal{A}_5$ is conjugate to its inverse. The system leads to $\alpha = 0, \beta + \gamma = 1, \delta = -1$ and $\beta^2 - \beta = 1$. Thus, $\beta = \frac{1}{2}(1 \pm \sqrt{5})$. The analogous system of equations also holds with $\chi_{X'}$ in place of $\chi_X$. Let us choose $+$ in $\beta$ for $\chi_X$ and take $-$ for $\chi_{X'}$; this will guarantee that the required equation $\chi_X + \chi_{X'} = \chi_{\Lambda^2 V_4 \downarrow A_5}$ is satisfied. The complete character table of $\mathcal{A}_5$ is given in Table 3.6. We remark that, for $\mathbb{k} = \mathbb{C}$, the representations $X$ and $X'$ arise from identifying $\mathcal{A}_5$ with the group of rotational symmetries of the regular icosahedron; see Example 3.35 below.

<table>
<thead>
<tr>
<th>classes</th>
<th>(1)</th>
<th>(1 2 3)</th>
<th>(1 2 3 4 5)</th>
<th>(2 1 3 4 5)</th>
<th>(1 2)(3 4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>sizes</td>
<td>1</td>
<td>20</td>
<td>12</td>
<td>12</td>
<td>15</td>
</tr>
<tr>
<td>$\mathbb{1}$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$\chi_{V_4 \downarrow A_5}$</td>
<td>4</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>0</td>
</tr>
<tr>
<td>$\chi_{W_4 \downarrow A_5}$</td>
<td>5</td>
<td>-1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>$\chi_X$</td>
<td>3</td>
<td>$\beta$</td>
<td>$\gamma$</td>
<td>-1</td>
<td></td>
</tr>
<tr>
<td>$\chi_{X'}$</td>
<td>3</td>
<td>$\gamma$</td>
<td>$\beta$</td>
<td>-1</td>
<td></td>
</tr>
</tbody>
</table>

Table 3.6. Character table of $\mathcal{A}_5$ (char $\mathbb{k} = 0, \beta = \frac{1}{2}(1 + \sqrt{5}), \gamma = \frac{1}{2}(1 - \sqrt{5})$)
Exercises for Section 3.5

3.5.1 (Restricting group actions to subgroups of index 2). Let $G$ be a group acting on a finite set $X$ (§3.2.5), and let $H$ be a subgroup of $G$ with $|G : H| = 2$. For any $x \in X$, let $G_x = \{g \in G \mid g.x = x\}$ be the isotropy group of $x$. Show:

(a) If $G_x \not\subseteq H$, then the $G$-orbit $G.x$ is identical to the $H$-orbit $H.x$.

(b) If $G_x \subseteq H$, then $G.x$ is the union of two $H$-orbits of equal size.

3.5.2 (Conjugacy in $A_n$). For $s \in S_n$, show that $C_{S_n}(s) \subseteq A_n$ precisely if the orbit sizes of $s$ are all odd and distinct.

3.5.3 (The degree-5 representation of $S_5$). This exercise constructs a 5-dimensional irreducible representation of $S_5$ over any field $\mathbb{k}$ with $\text{char } \mathbb{k} \neq 2, 3$ or 5.

(a) The standard action of $GL_2(F)$ on $F^2$ induces a permutation action of $PGL_2(F) = GL_2(F)/F^X$ on the projective line $P^1(F) = (F^2 \setminus \{0\})/F^X$. Show that this action is faithful, that is, only the identity element of $PGL_2(F)$ fixes all elements of $P^1(F)$, and doubly transitive in the sense of Exercise 3.2.4.10

(b) Let $F = \mathbb{F}_q$ be the field with $q$ elements and assume that $\text{char } \mathbb{k}$ does not divide $(q - 1)q(q + 1)$. Conclude from (a) and Exercise 3.2.4 that the deleted permutation representation over $\mathbb{k}$ for the permutation action $PGL_2(\mathbb{F}_q) \supseteq P^1(\mathbb{F}_q)$ is irreducible of degree $q$.

(c) Conclude from (a) that the action $PGL_2(\mathbb{F}_5) \supseteq P^1(\mathbb{F}_5)$ gives an embedding of $PGL_2(\mathbb{F}_5)$ as a subgroup of index 6 in $S_6$. The standard permutation action of $S_6$ on the set $S_5/PGL_2(\mathbb{F}_5)$ of left cosets of $PGL_2(\mathbb{F}_5)$ gives an automorphism $\phi \in \text{Aut}(S_6)$ such that $S_5 = \phi(PGL_2(\mathbb{F}_5))$. Thus, the deleted permutation representation in (b) gives a 5-dimensional irreducible representation of $S_5$ if $\text{char } \mathbb{k} \neq 2, 3$ or 5.

3.6. Some Classical Theorems

This section is entirely devoted to certain classical results, some of which establish purely group theoretical facts using representation theory as a tool. The reader is referred to Curtis’ Pioneers of Representation Theory: Frobenius, Burnside, Schur and Brauer [50] for a historical account of the formative stages in the development of group representation theory.

3.6.1. Divisibility Theorems of Frobenius, Schur and Itô

We first consider the degrees of irreducible representations of a finite group $G$. Clearly, $\dim_\mathbb{k} S \leq |G|$ for any $S \in \text{Irr } \mathbb{k} G$, because $S$ is an image of the regular representation. In fact, $\dim_\mathbb{k} S \leq [G : A]$ for any abelian subgroup $A \leq G$ provided

10In fact, the action $PGL_2(F) \supseteq P^1(F)$ is sharply 3-transitive: given a set of three distinct points $z_1, z_2, z_3 \in P^1(F)$ and a second set of distinct points $w_1, w_2, w_3$, there exists precisely one $g \in PGL_2(F)$ such that $g.z_i = w_i$ for $i = 1, 2, 3$. 
\( \kappa \) contains all \( e^{th} \) roots of unity, where \( e = \exp A \) is the exponent of \( A \); this follows from Corollary 3.5 and (3.16). Our goal in this subsection is to show that, for a large enough field \( \kappa \) of characteristic 0, the degrees of all \( S \in \text{Irr}_{kG} \) divide the index \( |G : A| \) of any abelian normal subgroup \( A \leq G \) (Ito’s Theorem). We shall repeatedly make use of the standard facts about integrality that were stated in §2.2.7.

Our starting point is a celebrated result of Frobenius from 1896 [74, §12].

**Frobenius’ Divisibility Theorem.** If \( S \) is an absolutely irreducible representation of a finite group \( G \) over a field \( \kappa \) of characteristic 0, then \( \dim_{\kappa} S \) divides \( |G| \).

**Proof.** After a field extension, we may assume that \( \kappa G \) is split semisimple. Choosing \( \lambda \) as in (3.14), we have \( c_\lambda = \sum_{g \in G} g \otimes g^{-1} \) and \( \gamma_A(1) = |G| \in \mathbb{Z} \subseteq \kappa \) by (3.15). Thus, Corollary 2.18 applies and we need to check that the Casimir element \( c_\lambda \) is integral over \( \mathbb{Z} \). But \( c_\lambda \) belongs to the subring \( \mathbb{Z} \otimes_{\mathbb{Z}} \mathbb{Z}G \subseteq \kappa G \otimes \kappa G \), which is finitely generated over \( \mathbb{Z} \), and integrality follows. \( \square \)

We remark that Frobenius’ Divisibility Theorem and the remaining results in this subsection remain valid in positive characteristics as long as \( \text{char} \kappa \nmid |G| \). In fact, the generalized versions follow from the characteristic-0 results proved here (e.g., Serre [181, Section 15.5]). However, Frobenius’ Divisibility Theorem generally no longer holds if \( \text{char} \kappa \) divides \( |G| \) as the following example shows.

**Example 3.27** (Failure of Frobenius’ Divisibility Theorem). Let \( \kappa \) be a field with \( \text{char} \kappa = p > 0 \) and let \( G = \text{SL}_2(\mathbb{F}_p) \), the group of all \( 2 \times 2 \)-matrices over \( \mathbb{F}_p \) having determinant 1. Since \( G \) is the kernel of \( \text{det} : \text{GL}_2(\mathbb{F}_p) \to \mathbb{F}_p^* \), we have \( |G| = \frac{|\text{GL}_2(\mathbb{F}_p)|}{|\mathbb{F}_p^*|} = \frac{|p^2 - 1|}{p - 1} = p(p + 1)(p - 1) \). For the second equality, observe that there are \( p^2 - 1 \) choices for the first column of an invertible \( 2 \times 2 \)-matrix, and then \( p^2 - p \) choices for the second column. Via the embedding \( G \hookrightarrow \text{SL}_2(k) \), the group \( G \) acts naturally on the vector space \( \kappa^2 \), and hence on all its symmetric powers,

\[ V(m) := \text{Sym}^m(\kappa^2) \quad (m \geq 0). \]

The dimension of \( V(m) \) is \( m + 1 \): if \( x, y \) is any basis of \( \kappa^2 \), then the monomials \( x^{m-i}y^i \) \((i = 0, 1, \ldots, m)\) form a basis of \( V(m) \). Moreover, \( V(m) \) is (absolutely) irreducible for \( m = 0, 1, \ldots, p - 1 \) (Exercise 3.6.1). However, \( \dim_{\kappa} V(p - 3) = p - 2 \) does not divide \( |G| \) for \( p \geq 7 \), because \( |G| \equiv 6 \mod p - 2 \).

The first sharpening of Frobenius’ Divisibility Theorem is due to I. Schur [180, Satz VII].

**Proposition 3.28.** Let \( S \) be an absolutely irreducible representation of a finite group \( G \) over a field \( \kappa \) of characteristic 0. Then \( \dim_{\kappa} S \) divides \( |G : \mathcal{Z}G| \).
3.6. Some Classical Theorems

Proof (following J. Tate). By hypothesis, the representation map \( \mathbb{k}G \to \text{End}_\mathbb{k}(S) \) is surjective (Burnside’s Theorem (§1.4.6)). Consequently, for each positive integer \( m \), we have a surjective map of algebras

\[
\begin{align*}
\mathbb{k}[G^\times m] & \xrightarrow{\sim} (\mathbb{k}G)^\otimes m \\
& \xrightarrow{\sim} \text{End}_\mathbb{k}(S)^\otimes m \\
& \xrightarrow{\sim} \text{End}_\mathbb{k}(S^\otimes m)
\end{align*}
\]

with \( g^i \in G \). It follows that \( S^\otimes m \) is an absolutely irreducible representation of the group \( G^\times m \) over \( \mathbb{k} \). For \( c \in \mathbb{Z} := \mathbb{Z}G \), the operator \( c_S \in \text{End}_\mathbb{k}(S) \) is a scalar; so each \( (c^1, \ldots, c^m) \in \mathbb{Z}^\times m \) acts on \( S^\otimes m \) as the scalar \( c_S^1 \cdots c_S^m = (c^1 \cdots c^m)_S \). Therefore, \( S^\otimes m \) is in fact a representation (absolutely irreducible) of the group \( G^\times m/C \), where we have put \( C := \{(c^1, \ldots, c^m) \in \mathbb{Z}^\times m \mid c^1 \cdots c^m = 1\} \). Now Frobenius’ Divisibility Theorem implies that \( \dim_\mathbb{k} S^\otimes m = (\dim_\mathbb{k} S)^m \) divides \( |G^\times m|/|C| = |G|^m/|\mathbb{Z}^\times m| \). In other words, \( q = \frac{|G| \cdot \mathbb{Z}}{\dim_\mathbb{k} S} \) satisfies \( q^m \in \frac{1}{|G|} \mathbb{Z} \) for all \( m \), and so \( \mathbb{Z}[q] \subseteq \frac{1}{|G|} \mathbb{Z} \). By the facts about integrality stated in §2.2.7, it follows that \( q \in \mathbb{Z} \), proving the proposition. \( \square \)

The culminating point of the developments described in this subsection is the following result due to N. Ito from 1951 [107].

Ito’s Theorem. Let \( G \) be finite, let \( A \) be a normal abelian subgroup of \( G \), and let \( \mathbb{k} \) be a field of characteristic 0 that contains all \( e^{th} \) roots of unity, where \( e \) is the exponent of \( A \). Then the degree of every absolutely irreducible representation of \( G \) over \( \mathbb{k} \) divides \( |G : A| \).

Proof. The proof is by induction on the index \( |G : A| \). If \( G = A \), then the result is clear, because all absolutely irreducible representations of \( \mathbb{k}A \) have degree 1.

The inductive step will be based on Proposition 3.28 and on a special case of Clifford’s Theorem (§3.6.4), which we explain here from scratch. Let \( S \) be an absolutely irreducible representation of \( \mathbb{k}G \). Our hypothesis on \( \mathbb{k} \) implies that \( S \) contains a common eigenvector for all operators \( a_S \) with \( a \in A \); so there is a group homomorphism \( \phi : A \to \mathbb{k}^\times \) with \( S_\phi = \{ s \in S \mid a.s = \phi(a)s \} \neq \emptyset \). Put

\[ H = \{ g \in G \mid \phi(g^{-1}ag) = \phi(a) \text{ for all } a \in A \} \]

and observe that \( H \) is a subgroup of \( G \) such that \( A \subseteq H \). Furthermore, \( h.S_\phi \subseteq S_\phi \) for all \( h \in H \) and the sum \( \sum_{g \in G/H} g.S_\phi \) is direct, because the various \( g.S \) are distinct homogeneous components of \( S|_A \). Since \( S = \mathbb{k}.S_\phi \) by irreducibility, we must have \( S = \bigoplus_{g \in G/H} g.S_\phi \). Thus, \( S_\phi \) is a subrepresentation of \( S|_H \) and the canonical map \( S_\phi |_H \to S \) from (3.8) is an isomorphism. It follows that \( S_\phi \) is an absolutely irreducible \( \mathbb{k}H \)-representation.
First assume that $H \neq G$. Then we know by induction that $\dim_k S_\phi$ divides $[H : A]$, and hence, $\dim_k S = [G : H] \dim_k S_\phi$ divides $[G : A]$. Finally, if $H = G$, then $S_\phi = S$ and so $A$ acts by scalars on $S$. Letting $\overline{\sigma}$ denote the images in $G_{\overline{G}}$, we have $\overline{A} \leq \mathcal{Z}(\overline{G})$. Hence, Proposition 3.28 gives that $\dim_k S$ divides $[\overline{G} : \overline{A}]$ and therefore also $[G : A]$, proving the theorem. □

### 3.6.2. Burnside’s $p^aq^b$-Theorem

The principal results in this section are purely group theoretical. Therefore, we choose to work over the field $\mathbb{C}$ of complex numbers. We will write

$$A_\mathbb{C} \overset{\text{def}}{=} \{ s \in \mathbb{C} \mid s \text{ is integral over } \mathbb{Z} \}.$$ 

As was mentioned in §2.2.7, $A$ is a subring of $\mathbb{C}$ such that $A \cap \mathbb{Q} = \mathbb{Z}$.

Now let $V$ be a finite-dimensional complex representation of a finite group $G$. Then the eigenvalues of all operators $g_V$ $(g \in G)$ are roots of unity, of order dividing the exponent $m = \exp G$, and hence they belong to $\mathbb{Z}[\zeta_m] \subseteq A$, where $\zeta_m := e^{2\pi i/m} \in \mathbb{C}$. Thus all character values $\chi_V(g)$ and all $\mathbb{Z}$-linear combinations of character values are contained in the subring $\mathbb{Z}[\zeta_m] \subseteq \mathbb{C}$. The following lemma contains the technicalities needed for the proof of Burnside’s $p^aq^b$-Theorem. Recall that $G_g$ denotes the $G$-conjugacy class of an element $g \in G$.

**Lemma 3.29.** Let $G$ be finite and let $S \in \text{Irr}(\mathbb{C}G)$ and $g \in G$ be such that $\dim_\mathbb{C} S$ and $|G_g|$ are relatively prime. Then either $\chi_S(g) = 0$ or $g_V$ is a scalar operator.

**Proof.** First, let $V \in \text{Rep}_{\text{fin}} \mathbb{C}G$ and $g \in G$ be arbitrary and put $s := \frac{\chi_V(g)}{\dim_\mathbb{C} V} \in \mathbb{Q}(\zeta_m)$.

**Claim.** $s \in A_\mathbb{C}$ if and only if $\chi_V(g) = 0$ or $g_V \in \text{End}_\mathbb{C}(V)$ is a scalar operator.

Indeed, $\chi_V(g) = 0$ implies $s = 0 \in A_\mathbb{C}$. Also, if $g_V \in \text{End}_\mathbb{C}(V)$ is a scalar operator, necessarily of the form $g_V = \zeta_m^r \text{Id}_V$ for some $r$, then again $s = \zeta_m^r \in A_\mathbb{C}$. Conversely, assume that $s \in A_\mathbb{C}$ and $\chi_V(g) \neq 0$. Then $0 \neq \gamma(s) \in A$ for all $\gamma \in \Gamma := \text{Gal}(\mathbb{Q}(\zeta_m)/\mathbb{Q})$. Thus, $0 \neq s := \prod_{\gamma \in \Gamma} \gamma(s) = \prod_{\gamma \in \Gamma} \frac{\gamma(\chi_V(g))}{\dim_\mathbb{C} V} \in A_\mathbb{C} \cap \mathbb{Q} = \mathbb{Z}$ and so $|s| \geq 1$. On the other hand, since each $\gamma(\chi_V(g))$ is a sum of $\dim_\mathbb{C} V$ many $m$th roots of unity, the triangle inequality implies that $|\gamma(\chi_V(g))| \leq \dim_\mathbb{C} V$ for all $\gamma \in \Gamma$. It follows that $|s| = 1$ and $|\chi_V(g)| = \dim_\mathbb{C} V$, which forces all eigenvalues of $g_V$ to be identical. Therefore, $g_V$ is a scalar operator, proving the Claim.

Now let $V = S \in \text{Irr}(\mathbb{C}G)$ and consider the class sum $\sigma_g = \sum_{x \in G_g} x \in \mathcal{Z}(\mathbb{C}G)$ (Example 3.14). By Schur’s Lemma, the operator $(\sigma_g)_S \in \text{End}_\mathbb{C}(S)$ is a scalar. Hence,

$$(\sigma_g)_S = \frac{\text{trace} (\sigma_g)_S}{\dim_\mathbb{C} S} = \frac{|G_g| \chi_S(g)}{\dim_\mathbb{C} S}.$$ 

Since $\sigma_g \in \mathbb{Z}G$, a subring of $\mathbb{C}G$ that is finitely generated as $\mathbb{Z}$-module, $\sigma_g$ satisfies a monic polynomial over $\mathbb{Z}$, and hence $(\sigma_g)_S$ satisfies the same polynomial. This
Some Classical Theorems shows that $|G| \mid \chi_S(g)$ as above. Finally, if $|G|$ and $\dim C_S$ are relatively prime, then it follows that $s \in A$, because $|G|s \in A$ and $(\dim C_S)s = \chi_S(g) \in A$. We may now invoke the Claim to finish the proof. □

The following result of Burnside originally appeared in the second edition (1911) of his monograph The Theory of Groups of Finite Order [36].

**Burnside’s $p^a q^b$-Theorem.** Every group of order $p^a q^b$, where $p$ and $q$ are primes, is solvable.

Before embarking on the argument, let us make some preliminary observations. By considering composition series of the groups in question, the assertion of the theorem can be reformulated as the statement that every simple group $G$ of order $p^a q^b$ is abelian. Assume that $a > 0$ and let $P$ be a Sylow $p$-subgroup of $G$. Then $Z P$, and, for every $g \in Z P$, the centralizer $C_G(g)$ contains $P$ and, consequently, the size of the conjugacy class of $g$ is a power of the prime $q$. Therefore, Burnside’s $p^a q^b$-Theorem will be a consequence of the following

**Theorem 3.30.** Let $G$ be a finite nonabelian simple group. Then $\{1\}$ is the only conjugacy class of $G$ having prime power size.

**Proof.** Assume, for a contradiction, that there is an element $1 \neq g \in G$ such that $|G|$ is a power of the prime $p$. Representation theory enters the argument via the following

**Claim.** For $1 \neq S \in \text{Irr } C_G$, we have $\chi_S(g) = 0$ or $p \mid \dim C_S$.

To prove this, assume that $p \nmid \dim C_S$. Then Lemma 3.29 tells us that either $\chi_S(g) = 0$ or else $g_S$ is a scalar operator. However, since $S \neq 1$ and $G$ is simple, the representation $g \mapsto g_S$ is an embedding $G \hookrightarrow \text{GL}(S)$. Thus the possibility $g_S \in \mathbb{C}$ would imply that $g \in \mathcal{Z} G$, which in turn would force $G$ to be abelian contrary to our hypothesis. Thus, we are left with the other possibility, $\chi_S(g) = 0$.

We can now complete the proof of the theorem as follows. Since the regular representation of $C G$ has the form $(C G)_{\text{reg}} \cong \bigoplus_{S \in \text{Irr } C_G} S^{\dim C_S}$ by Maschke’s Theorem (§3.4.1), we can write $\chi_{\text{reg}}(g) = 1 + ps$ with $s := \sum_{1 \neq S \in \text{Irr } C_G} \frac{\dim C_S}{p} \chi_S(g)$. Note that $s \in A$ by the claim and our remarks about character values above. On the other hand, since $\chi_{\text{reg}}(g) = 0$ by (3.21), we obtain $s = -\frac{1}{p} \in \mathbb{Q} \setminus \mathbb{Z}$, contradicting the fact that $A \cap Q = \mathbb{Z}$ and finishing the proof. □

**3.6.3. The Brauer-Fowler Theorem**

The Brauer-Fowler Theorem [32] (1955) is another purely group theoretical result. It is of historical significance inasmuch as it led to Brauer’s program of classifying finite simple groups in terms of the centralizers of their involutions. Indeed, as
had been conjectured by Burnside in his aforementioned monograph *The Theory of Groups of Finite Order* ([36, Note M]), all finite non-abelian simple groups are of even order—this was eventually proved by Feit and Thompson in 1963 in their seminal odd-order paper [71]. Thus, any finite non-abelian simple group $G$ must contain an involution, that is, an element $1 \neq u \in G$ such that $u^2 = 1$. The Brauer-Fowler Theorem states that $G$ is “almost” determined by the size of the centralizer $C_G(u) = \{g \in G \mid gu = ug\}$:

**Brauer-Fowler Theorem.** Given $n$, there are at most finitely many finite non-abelian simple groups (up to isomorphism) containing an involution with centralizer of order $n$. In fact, each such group embeds into the alternating group $A_{n^2-1}$.

In light of this result, Brauer proposed a two-step strategy to tackle the problem of classifying all finite simple groups: investigate the possible group theoretical structures of the centralizers of their involutions and then, for each group $C$ in the resulting list, determine the finitely many possible finite simple groups $G$ containing an involution $u$ with $C_G(u) \cong C$. This program was the start of the classification project for finite simple groups. The project was essentially completed, with D. Gorenstein at the helm, in the early 1980s; some gaps had to be filled in later. In the course of these investigations, it turned out that, with a small number of exceptions, $G$ is in fact uniquely determined by the involution centralizer $C$. (Exercise 3.6.2 considers the easiest instance of this.) For an overview of the classification project, its history and the statement of the resulting Classification Theorem, see R. Solomon’s survey article [188] or Aschbacher’s monograph [6].

To explain the representation theoretic tools used in the proof of the Brauer-Fowler Theorem below, let $G$ be any finite group and consider the following function, for any given positive integer $n$,

$$\theta_n : G \rightarrow \mathbb{Z}$$

$$g \mapsto \# \{h \in G \mid h^n = g\}$$

(3.50)

Thus, $\theta_2(1) - 1$ is the number of involutions of $G$. Each $\theta_n$ is clearly a $\mathbb{C}$-valued class function on $G$, and hence we know that $\theta_n$ is a $\mathbb{C}$-linear combination of the irreducible complex characters of $G$ (Corollary 3.21). To wit:

**Lemma 3.31.** Let $G$ be a finite group and let $\theta_n$ be as in (3.50). Then:

$$\theta_n = \sum_{S \in \text{Irr } C_G} \nu_n(S)\chi_S \quad \text{with} \quad \nu_n(S) \overset{\text{def}}{=} \frac{1}{|G|} \sum_{g \in G} \chi_S(g^n)$$

In particular, $\sum_{1 \neq S \in \text{Irr } C_G} \nu_2(S) \dim \mathbb{C} S$ is the number of involutions of $G$. 

3.6. Some Classical Theorems

Proof. Write \( \theta_n = \sum_{S \in \text{Irr} \mathbb{C}G} \lambda_S \chi_S \) with \( \lambda_S \in \mathbb{C} \) and note that \( \theta_n(g^{-1}) = \theta_n(g) \) for \( g \in G \). Now use the orthogonality relations to obtain

\[
\lambda_S = \langle \chi_S, \theta_n \rangle = \frac{1}{|G|} \sum_{g \in G} \chi_S(g)\theta_n(g) = \frac{1}{|G|} \sum_{h \in G} \chi_S(h^n) = \nu_n(S)
\]

The involution count formula just expresses \( \theta_2(1) - 1 \).

The complex numbers \( \nu_n(S) \) are called the Frobenius-Schur indicators of the representation \( S \); they will be considered in more detail and in greater generality in Section 12.5. In particular, we will show there that \( \nu_n(S) \) can only take the values 0 and ±1 for any \( S \in \text{Irr} \mathbb{C}G \) (Theorem 12.27).Granting this fact for now, we can give the

Proof of the Brauer-Fowler Theorem. Let \( G \) be a finite non-abelian simple group containing an involution \( u \in G \) such that \( |C_G(u)| = n \), and let \( t \) denote the number of involutions of \( G \). It will suffice to prove the following

Claim. For some \( 1 \neq g \in G \), the size of the conjugacy class \( G \cdot g \) is at most \( \left( \frac{|G| - 1}{t} \right)^2 \).

To see how the Brauer-Fowler Theorem follows from this, note that \( t \geq \frac{|G|}{n} = |G_u| \), because the conjugacy class \( G_u \) consists of involutions. Thus, \( |G_g| \leq \left( \frac{|G| - 1}{t} \right)^2 \leq \left( \frac{|G| - 1}{|G|} \right)^2 n^2 \) and so the conjugation action \( G \to G_g \) gives rise to a homomorphism \( G \to S_{n^2-1} \). Since \( G \) is simple and not isomorphic to \( C_2 \), this map is injective and has image in \( A_{n^2-1} \).

It remains to prove the Claim, which does in fact hold for any finite group \( G \). Suppose on the contrary that \( |G_g| > \left( \frac{|G| - 1}{t} \right)^2 \) for all \( 1 \neq g \in G \) and let \( k \) denote the number of conjugacy classes of \( G \). Then \( |G| - 1 > (k - 1)(\frac{|G| - 1}{t})^2 \) or, equivalently,

\[
t^2 > (k - 1)(|G| - 1).
\]

In order to prove that this is absurd, we use Lemma 3.31 and the fact that all \( \nu_2(S) \in \{0, \pm1\} \) to obtain the estimate

\[
t = \sum_{1 \neq S \in \text{Irr} \mathbb{C}G} \nu_2(S) \dim_S S \leq \sum_{1 \neq S \in \text{Irr} \mathbb{C}G} \dim_S S.
\]

The inequality\(^{11}\) \( \left( \sum_{i=1}^d x_i \right)^2 \leq d \sum_{i=1}^d x_i^2 \) in conjunction with the equalities \( k = |\text{Irr} \mathbb{C}G| \) and \( \sum_{S \in \text{Irr} \mathbb{C}G} (\dim_S S)^2 = |G| \) (Corollary 3.21) then further implies

\[
t^2 \leq \left( \sum_{1 \neq S \in \text{Irr} \mathbb{C}G} \dim_S S \right)^2 \leq (k - 1) \sum_{1 \neq S \in \text{Irr} \mathbb{C}G} (\dim_S S)^2 = (k - 1)(|G| - 1).
\]

This contradicts our prior inequality for \( t^2 \), thereby proving the Claim and finishing the proof of the Brauer-Fowler Theorem.

\(^{11}\)This inequality follows from the Cauchy-Schwarz inequality \(|x \cdot y| \leq |x| |y| \) in Euclidean space \( \mathbb{R}^d \) by taking \( y = (1, 1, \ldots, 1) \).
3.6.4. Clifford Theory

Originally developed by A. H. Clifford [45] in 1937, Clifford theory studies the interplay between the irreducible representations of a group $G$, not necessarily finite, over an arbitrary field $k$ and those of a normal subgroup $N \trianglelefteq G$ having finite index in $G$. Special cases have already been considered in passing earlier (Lemma 3.26 and the proof of Itô’s Theorem). We will now address the passage between $\text{Rep}_k G$ and $\text{Rep}_k N$ more systematically and in greater generality, the principal tools being restriction $\downarrow_N = \text{Res}_{kG}^{kN}$ and induction $\uparrow_G = \text{Ind}_{kN}^{kG}$. In fact, it turns out that the theory is not specific to group algebras and can be explained at little extra cost in the more general setting of crossed products.

Crossed products. First, let $N$ be an arbitrary normal subgroup of $G$ and put $\Gamma = G/N$. It is an elementary fact that if $\{x \mid x \in \Gamma\} \subseteq G$ is any transversal for the cosets of $N$ in $G$, then $kG = \bigoplus_{x \in \Gamma} kNx$ and $kNx = \overline{x}kN$ for all $x$. Thus, putting $B = kG$ and $B^x = kNx$, the algebra $B$ is $\Gamma$-graded (Exercise 1.1.11) and each homogeneous component $B^x$ contains a unit, for example $x$:

\begin{equation}
B = \bigoplus_{x \in \Gamma} B^x, \quad B^x B^y \subseteq B^{xy} \quad \text{and} \quad B^x \cap B^y \neq \emptyset \quad \text{for all } x \in \Gamma.
\end{equation}

In general, any $\Gamma$-graded algebra $B$ as in (3.51), with $\Gamma$ a group and with identity component $A = B^1$, is called a crossed product of $\Gamma$ over $A$ and denoted by

$$B = A * \Gamma.$$ 

Thus, the group algebra $kG$ is a crossed product, $kG = (kN) * \Gamma$ with $\Gamma = G/N$. It is also clear that, for any crossed product $B = A * \Gamma$ and any submonoid $\Delta \subseteq \Gamma$, the sum $A * \Delta := \sum_{x \in \Delta} B^x$ is a subalgebra of $B$. Furthermore, for any choice of units $\overline{x} \in B^x \cap B^y$, one easily shows (Exercise 3.6.5) that the homogeneous components of $B = A * \Gamma$ are given by

\begin{equation}
B^x = A\overline{x} = \overline{x}A.
\end{equation}

Therefore, the units $\overline{x}$ are determined up to a factor in $A^\times$ and conjugation by $\overline{x}$ gives an automorphism $\overline{x}(,)\overline{x}^{-1} \in \text{Aut}_{\text{Alg}}(A)$, which depends on the choice of $\overline{x}$ only up to an inner automorphism of $A$.

Twisting. Let $B = A * \Gamma$ be an arbitrary crossed product. Then each homogeneous component $B^x$ is an $(A, A)$-bimodule via multiplication in $B$. Thus, for any $W \in \text{Rep} A$, we may define

\begin{equation}
xW \overset{\text{def}}{=} B^x \otimes_A W \in \text{Rep} A \quad \text{and} \quad \Gamma_W \overset{\text{def}}{=} \{x \in \Gamma \mid xW \cong W\}.
\end{equation}

**Lemma 3.32.** With the above notation, $\Gamma_W$ is a subgroup of $\Gamma$ and $xW \cong yW$ if and only if $x\Gamma_W = y\Gamma_W$. 

3.6. Some Classical Theorems

Proof. It follows from (B.5) and (3.52) that multiplication of $B$ gives an isomorphism $B^x \otimes_A B^y \cong B^{xy}$ as $(A, A)$-bimodules for any $x, y \in \Gamma$. By associativity of the tensor product $\otimes_A$ and the canonical isomorphism $A \otimes_A W \cong W$ in $\text{Rep} \ A$, we obtain isomorphisms $x^*(W) \cong xyW$ and $1^W \cong W$ in $\text{Rep} \ A$. Both assertions of the lemma are immediate consequences of these isomorphisms. □

Restriction. The main result of this subsection concerns the behavior of irreducible representations of a crossed product $B = A \ast \Gamma$ with $\Gamma$ finite under restriction to the identity component $A$. This covers of the process restricting irreducible representations of an arbitrary group $G$ to a normal subgroup $N \leq G$ having finite index in $G$. For any $W \in \text{Rep} \ A$, we will consider the subalgebra

$$B_W \overset{\text{def}}{=} B \ast \Gamma_W \subseteq B$$

with $\Gamma_W$ as in (3.53).

Clifford’s Theorem. Let $B = A \ast \Gamma$ be a crossed product with $\Gamma$ a finite group. Then, for any $V \in \text{Irr} \ B$, the restriction $V_{\downarrow A}$ is completely reducible of finite length. More precisely, if $S$ is any irreducible subrepresentation of $V_{\downarrow A}$ and $V(S)$ is the $S$-homogeneous component of $V_{\downarrow A}$, then

$$V_{\downarrow A} \cong \left( \bigoplus_{x \in \Gamma} xS \right) \oplus \text{length} \, V(S).$$

Furthermore, $V(S)$ is a subrepresentation of $V_{\downarrow B_S}$ and $V \cong V(S)|_{B_S}$.

Proof. Since $\Gamma$ is finite, the restriction $V_{\downarrow A}$ is finitely generated. Hence there exists a maximal subrepresentation $M \subseteq V_{\downarrow A}$ (Exercise 1.1.3). All $\overline{x}.M$ with $x \in \Gamma$ are maximal subrepresentations of $V_{\downarrow A}$ as well and (3.52) implies that $\overline{x}(\overline{y}.M) = \overline{x y}.M$ for $x, y \in \Gamma$. Therefore, $\bigcap_{x \in \Gamma} \overline{x}.M$ is a proper subrepresentation of $V$ and, consequently, $\bigcap_{x \in \Gamma} \overline{x}.M = 0$ by irreducibility. This yields an embedding $V_{\downarrow A} \hookrightarrow \bigoplus_{x \in \Gamma} V/\overline{x}.M$, proving that $V_{\downarrow A}$ is completely reducible of finite length.

Fix an irreducible subrepresentation $S \subseteq V_{\downarrow A}$. Then all $\overline{x}.S$ are also irreducible subrepresentations of $V_{\downarrow A}$ and $\overline{x}S \cong \overline{x}.S$ via $\overline{x} \otimes s \mapsto \overline{x}.s$. The sum $\sum_{x \in \Gamma} \overline{x}.S$ is a nonzero subrepresentation of $V$ and so we must have $\sum_{x \in \Gamma} \overline{x}.S = V$. By Corollary 1.29 and Lemma 3.32, the irreducible constituents of $V_{\downarrow A}$ are exactly the various twists $\overline{x}S$ with $x \in \Gamma/\Gamma_S$. Therefore, $V_{\downarrow A} = \bigoplus_{x \in \Gamma/\Gamma_S} V(\overline{x}S)$. Furthermore, for $x, y \in \Gamma$, we certainly have $\overline{x}.V(\overline{y}S) \subseteq V(\overline{xy}S)$, because $\overline{x}S \subseteq \overline{x}S \cong \overline{xy}S$. In particular, $V(S)$ is stable under the action of $B_S$ and $\overline{x}.V(S) = V(\overline{x}S)$ for all $x \in \Gamma$. Consequently, $V_{\downarrow A} = \bigoplus_{x \in \Gamma/\Gamma_S} \overline{x}.V(S) \overset{\text{def}}{=} \left( \bigoplus_{x \in \Gamma/\Gamma_S} xS \right) \oplus \text{length} \, V(S)$. It follows that $V \cong V(S)|_{B_S}$ via the canonical map $V(S)|_{B_S} \rightarrow V$ that corresponds to the inclusion $V(S) \hookrightarrow V_{\downarrow B_S}$ in Proposition 1.9. This completes the proof of Clifford’s Theorem. □
Outlook. More can be said in the special case where \( B = \mathbb{F}G = (\mathbb{F}N) \ast (G/N) \) for a finite group \( G \) and \( N \triangleleft G \). If \( \mathbb{F} \) is a field of characteristic 0 that contains sufficiently many roots of unity, then one can use Schur’s theory of projective representations to show that, in the situation of Clifford’s Theorem, \( \dim V(S) \) divides the order of \( \Gamma_S \) or, equivalently, \( \frac{\dim V}{\dim S} \) divides \( |\Gamma| \). This allows for a generalization of Itô’s Theorem to subnormal abelian subgroups \( A \trianglelefteq G \); see [106, Corollaries 11.29 and 11.30]. The monograph [165] by Passman is devoted to the ring theoretic properties of crossed products \( B = A \ast \Gamma \) with \( \Gamma \) (mostly) infinite. Crossed products also play a crucial role in the investigation of division algebras, Galois cohomology and Brauer groups; see [167, Chapter 14] for an introduction to this topic.

Exercises for Section 3.6

3.6.1 (Absolutely irreducible representations of \( \text{SL}_2(\mathbb{F}_p) \) in characteristic \( p \)). Let \( \mathbb{F} \) be a field with \( \text{char} \mathbb{F} = p > 0 \) and let \( G = \text{SL}_2(\mathbb{F}_p) \). To justify a claim made in Example 3.27, show:

(a) The number of \( p \)-regular conjugacy classes of \( G \) is \( p \).

(b) \( V(m) = \text{Sym}^m(\mathbb{F}_2^2) \in \text{Irr}_\mathbb{F}G \) for \( m = 0, 1, \ldots, p-1 \). Consequently, the \( V(m) \) are a full set of non-equivalent irreducible representations of \( \mathbb{F}G \) (Theorem 3.7).

3.6.2 (The Brauer program for involution centralizer \( C_2 \)). Let \( G \) be a finite group containing an involution \( u \in G \) such that the centralizer \( C_G(u) \) has order 2. Use the column orthogonality relations (Exercise 3.4.7) to show that \( G^{ab} = G/[G, G] \) has order 2. Consequently, if \( G \) is simple, then \( G \cong C_2 \).

3.6.3 (Frobenius-Schur indicators). Let \( G \) be a finite group and assume that \( \text{char} \mathbb{F} \nmid |G| \). Recall that the \( n \)-th Frobenius-Schur indicator of \( V \in \text{Rep}_\mathbb{F}G \) is defined by \( \nu_n(V) = \frac{1}{|G|} \sum_{g \in G} \chi_V(g^n) \) (Lemma 3.31). The goal of this exercise is to show that, if \( n \) is relatively prime to \( |G| \), then \( \nu_n(S) = 0 \) for all \( \mathbb{F} \neq S \in \text{Irr} \mathbb{F}G \). To this end, prove:

(a) The \( n \)-th power map \( G \to G, g \mapsto g^n \), is a bijection.

(b) For any field \( \mathbb{F} \) and any \( \mathbb{F} \neq S \in \text{Irr} \mathbb{F}G \), the equalities \( \sum_{g \in G} g^n_S = \sum_{g \in G} g_S = 0 \) hold in \( \text{End}_{\mathbb{F}}G(S) \). Consequently, \( \sum_{g \in G} X_S(g^n) = 0 \).

3.6.4 (Frobenius’ formula). Let \( G \) be a finite group and let \( C_1, \ldots, C_k \) be arbitrary conjugacy classes of \( G \). Put

\[
N(C_1, \ldots, C_k) = \#\{(g_1, \ldots, g_k) \in C_1 \times \cdots \times C_k \mid g_1g_2\cdots g_k = 1\}.
\]

For any \( S \in \text{Irr} \mathbb{C}G \), let \( \chi_{S,i} \) denote the common value of all \( \chi_S(g) \) with \( g \in C_i \). Prove the equalities (a) and (b) below for the product of the class sums \( \sigma_i = \sum_{g \in C_i} g \in \mathcal{Z}(\mathbb{C}G) \) to obtain the following formula of Frobenius:

\[
N(C_1, \ldots, C_k) = \frac{|C_1| \cdots |C_k|}{|G|} \sum_{S \in \text{Irr} \mathbb{C}G} \frac{\chi_{S,1} \cdots \chi_{S,k}}{(\dim_{\mathbb{C}S})^{k-2}}.
\]
3.7. Characters, Symmetric Polynomials, and Invariant Theory

3.7.1. Symmetric Polynomials

This subsection collects some basic facts concerning symmetric polynomials. Throughout, $x_1, x_2, \ldots, x_d$ denote independent commuting variables over $k$. A polynomial in $x_1, x_2, \ldots, x_d$ is called symmetric if it is unchanged under any permutation of the variables: $f(x_1, \ldots, x_d) = f(x_{s(1)}, \ldots, x_{s(d)})$ for any $s \in S_d$.

Foremost among them are the elementary symmetric polynomials, the $k^{th}$ of which is defined by

$$e_k = e_k(x_1, x_2, \ldots, x_d) \overset{\text{def}}{=} \sum_{1 \leq i_1 < i_2 < \cdots < i_k \leq d} x_{i_1}x_{i_2} \cdots x_{i_k} = \sum_{S \subseteq [d]} \prod_{i \in S} x_i.$$

A brief contemplation will convince the reader that the following identity holds; the sum on the left is called the generating function of the polynomials $e_k$:

$$\sum_{k=0}^{d} e_k t^k = \prod_{i=1}^{d} (1 + x_i t). \tag{3.54}$$

Next, the $k^{th}$ complete symmetric polynomial is defined by

$$h_k = h_k(x_1, x_2, \ldots, x_d) \overset{\text{def}}{=} \sum_{1 \leq i_1 \leq i_2 \leq \cdots \leq i_k \leq d} x_{i_1}x_{i_2} \cdots x_{i_k}$$

$$= \sum_{l_1 + l_2 + \cdots + l_d = k \atop l_i \geq 0} x_1^{l_1}x_2^{l_2} \cdots x_d^{l_d}.$$
It is easy to see that the generating function of the polynomials $h_k$ has the following form:

$$
\sum_{k \geq 0} h_k t^k = \prod_{i=1}^{d} \left(1 + x_i t + x_i^2 t^2 + \ldots \right)
= \prod_{i=1}^{d} \frac{1}{1 - x_i t}.
$$

(3.55)

The symmetric polynomials $e_k$ and $h_k$ are closely related to the **power sums**, the $k^{th}$ of which is defined by

$$
p_k = p_k(x_1, x_2, \ldots, x_d) \overset{\text{def}}{=} \sum_{i=1}^{d} x_i^k \quad (k \geq 1).
$$

The connection is described in the classical **Newton identities**:

$$
nh_n = \sum_{k=1}^{n} p_k h_{n-k}

(3.56)

ne_n = \sum_{k=1}^{n} (-1)^{k-1} p_k e_{n-k}

\text{Proof of the Newton identities.} \quad \text{We may view } h_k, e_k, p_k \in \mathbb{Z}[x_1, \ldots, x_d] \text{ and establish (3.56) working over } \mathbb{k} = \mathbb{Q}; \text{ over an arbitrary field } \mathbb{k}, \text{ the relations (3.56) then follow via the canonical homomorphism } \mathbb{Z}[x_1, \ldots, x_d] \rightarrow \mathbb{k}[x_1, \ldots, x_d]. \text{ Thus, we may use logarithmic differentiation.}

\text{Put } P(t) = \sum_{k \geq 1} p_k t^{k-1} \text{ and } H(t) = \sum_{k \geq 0} h_k t^k; \text{ the latter series is given by (3.55). Then}

$$
P(t) = \sum_{i=1}^{d} \sum_{k \geq 1} x_i^k t^{k-1} = \sum_{i=1}^{d} \frac{x_i}{1 - x_i t} = \sum_{i=1}^{d} \frac{d}{dt} \log \frac{1}{1 - x_i t}
= \frac{d}{dt} \log \prod_{i=1}^{d} \frac{1}{1 - x_i t} \overset{\text{(3.55)}}{=} \frac{d}{dt} \log H(t) = \frac{H'(t)}{H(t)}.
$$

Comparing coefficients at $t^{n-1}$ in $H'(t) = P(t)H(t)$ gives the first identity. The second identity is derived similarly using the generating function $E(t) = \sum_{k \geq 0} e_k t^k$ together with the observation that $P(-t) = \frac{d}{dt} \log E(t)$ by (3.54). \qed
If \( \text{char } k = 0 \), then we may solve the Newton identities recursively to express all \( e_k \) and all \( h_k \) in terms of the power sums \( p_{k'} \) with \( k' \leq k \). For example,

\[
\begin{align*}
e_1 &= p_1 \\
e_2 &= \frac{1}{2}(p_1^2 - p_2) \\
e_3 &= \frac{1}{3!}(p_1^3 - 3p_1p_2 + 2p_3) \\
e_4 &= \frac{1}{4!}(p_1^4 - 4p_1^2p_2 + 3p_2^2 + 8p_1p_3 - 6p_4)
\end{align*}
\]

3.7. Generating Functions for Linear Transformations

Let \( V \) be a finite-dimensional \( k \)-vector space. Any \( f \in \text{End}_k(V) \) gives rise to the endomorphism \( T_f = \bigoplus_k f^{\otimes k} \in \text{End}_k(TV) \) (§1.1.2). Using the formula \( \text{trace}(f^{\otimes k}) = \text{trace}(f)^k \), which follows from (B.26), the generating function of these traces can be expressed as follows:

\[
\sum_{k \geq 0} \text{trace}(f^{\otimes k}) t^k = \frac{1}{1 - \text{trace}(f)t}.
\]

Similarly, we have the series \( \sum_{k \geq 0} \text{trace}(\text{Sym}^k f) t^k \) and \( \sum_{k \geq 0} \text{trace}(\Lambda^k f) t^k \) for the endomorphisms \( \text{Sym} f \) and \( \Lambda f \). The latter series is in fact a polynomial, because \( \Lambda^k V = 0 \) for \( k > \dim_k V \). The following lemma\(^{13}\) gives expressions for these generating functions analogous to (3.58).

**Lemma 3.33.** Let \( V \) be a finite-dimensional \( k \)-vector space and let \( f \in \text{End}_k(V) \). Then

\[
\sum_{k \geq 0} \text{trace}(\text{Sym}^k f) t^k = \frac{1}{\det(\text{Id}_V - ft)}
\]

and

\[
\sum_{k \geq 0} \text{trace}(\Lambda^k f) t^k = \det(\text{Id}_V + ft).
\]

**Proof.** By passing to an algebraic closure of \( k \), we may assume that \( V \) has a \( k \)-basis \( b_1, b_2, \ldots, b_d \) so that the matrix of \( f \) has upper triangular form with the eigenvalues \( \lambda_i \in k \) of \( f \) along the diagonal:

\[
\begin{pmatrix}
\lambda_1 & * \\
\lambda_2 & \\
0 & \ddots \\
& & \lambda_d
\end{pmatrix}
\]

\(^{13}\)The lemma is a one-variable version of the “MacMahon Master Theorem” [142, p. 97-98] from enumerative combinatorics.
By (1.10) a basis of $\text{Sym}^k V$ is given by the standard monomials $b_I = b_{i_1} b_{i_2} \ldots b_{i_k}$ of degree $k$, with $I = (i_1, i_2, \ldots, i_k)$ and $1 \leq i_1 \leq i_2 \leq \cdots \leq i_k \leq d$. We order the basis $(b_I)_I$ by using the lexicographic ordering on the set of all sequences $I$ and put $\lambda_I = \lambda_{i_1} \lambda_{i_2} \ldots \lambda_{i_k}$. Then

\[
(\text{Sym}^k f)(b_I) = f(b_{i_1}) f(b_{i_2}) \ldots f(b_{i_k})
\]

\[
= \prod_{j=1}^k \left( \lambda_{i_j} b_{i_j} + \text{contributions from basis vectors } b_I \text{ with } l < i_j \right)
\]

\[
= \lambda_I b_I + \left( \text{contributions from basis vectors } b_J \text{ with } J < I \right).
\]

Thus, the matrix of $\text{Sym}^k f$ for the basis $(b_I)_I$ is upper triangular with the values $\lambda_I$ on the diagonal. Consequently,

\[
(3.59) \quad \text{trace}(\text{Sym}^k f) = \sum_I \lambda_I = h_k(\lambda_1, \lambda_2, \ldots, \lambda_d),
\]

where $h_k$ is the $k^{th}$ complete symmetric polynomial. The following computation now proves the first formula:

\[
\sum_{k \geq 0} \text{trace}(\text{Sym}^k f)t^k = (3.59) \sum_{k \geq 0} h_k(\lambda_1, \lambda_2, \ldots, \lambda_d)t^k
\]

\[
= (3.55) \prod_{i=1}^d \frac{1}{1 - \lambda_i t} = \frac{1}{\det(\text{Id}_V - ft)}.
\]

The proof of the second formula is similar. A basis of $\wedge^k V$ is given by the elements $\wedge b_I = b_{i_1} \wedge b_{i_2} \wedge \cdots \wedge b_{i_k}$ with $I = (i_1, i_2, \ldots, i_k)$ and $1 \leq i_1 < i_2 < \cdots < i_k \leq d$ by (1.13). Using the lexicographic ordering on the sequences $I$, we again have

\[
(\wedge^k f)(\wedge b_I) = \lambda_I (\wedge b_I) + \text{contributions from basis vectors } \wedge b_J \text{ with } J < I
\]

Therefore,

\[
(3.60) \quad \text{trace}(\wedge^k f) = e_k(\lambda_1, \lambda_2, \ldots, \lambda_d)
\]

where $e_k$ is the $k^{th}$ elementary symmetric polynomial. Finally,

\[
\sum_{k \geq 0} \text{trace}(\wedge^k f)t^k = (3.60) \sum_{k \geq 0} e_k(\lambda_1, \lambda_2, \ldots, \lambda_d)t^k
\]

\[
= (3.54) \prod_{i=1}^d (1 + \lambda_i t) = \det(\text{Id}_V + ft),
\]

which completes the proof of the lemma. \(\square\)
3.7.3. Formal Characters

Now let us consider the situation where $G$ is an arbitrary group and $V \in \text{Rep} \, \mathbb{k} G$ is given. By §3.3.3, we also have the representations $TV, \text{Sym} V$ and $\Lambda V \in \text{Rep} \, \mathbb{k} G$ along with their homogeneous components, $V^{\otimes k}, \text{Sym}^k V$ and $\Lambda^k V \in \text{Rep} \, \mathbb{k} G$. Assuming $V$ to be finite dimensional, we define the formal character of $TV$ by

$$X_{TV}(g, t) \overset{\text{def}}{=} \sum_{k \geq 0} X_{V^{\otimes k}}(g) t^k = \frac{1}{1 - X_V(g) t} \quad (g \in G).$$

(3.61)

Here, the expression on the right follows from (3.58), since $X_{V^{\otimes k}}(g) = \text{trace}(g^{\otimes k})$. Similarly, Lemma 3.33 yields the formal characters

$$X_{\text{Sym} V}(g, t) \overset{\text{def}}{=} \sum_{k \geq 0} X_{\text{Sym}^k V}(g) t^k = \frac{1}{\det(\text{Id}_V - g_V t)},$$

$$X_{\Lambda V}(g, t) \overset{\text{def}}{=} \sum_{k \geq 0} X_{\Lambda^k V}(g) t^k = \det(\text{Id}_V + g_V t).$$

(3.62)

The series $X_{\Lambda V}(g, t)$ is a polynomial having degree $d = \dim \mathbb{k} V$ and leading coefficient $X_{\Lambda^d V}(g) = \det(g_V)$.

**Corollary 3.34.** Let $G$ be an arbitrary group and assume that $\text{char} \, \mathbb{k} = 0$. Then, for any $V \in \text{Rep}_{\text{fin}} \, \mathbb{k} G$, the character $X_V$ determines the characters $X_{\Lambda^k V}$ and $X_{\text{Sym}^k V}$ for all $k \geq 0$ as well as the characteristic polynomial of $g_V$ for all $g \in G$.

**Proof.** By (3.62),

$$X_{\text{Sym} V}(g, t) = \frac{1}{X_{\Lambda V}(g, -t)}.$$

Putting $d = \dim \mathbb{k} V$, (3.62) also gives the following expression for the characteristic polynomial of $g_V$:

$$\det(\text{Id}_V t - g_V) = t^d \det(\text{Id}_V - g_V t^{-1}) = t^d X_{\Lambda V}(g, -t^{-1}).$$

Therefore, it suffices to show that $X_V$ determines all $X_{\Lambda^k V}$. To prove this, recall that $X_{\Lambda^k V}(g) = \text{trace}(\Lambda^k g_V) = e_k(\lambda_1, \lambda_2, \ldots, \lambda_d)$ by (3.60), where $\lambda_1, \ldots, \lambda_d$ are the eigenvalues of $g_V$ in some algebraic closure of $\mathbb{k}$. Moreover, clearly, $X_V(g^k) = \sum_{i=1}^d \lambda_i^k = p_k(\lambda_1, \lambda_2, \ldots, \lambda_d)$. Therefore, (3.57) yields the following expressions for the characters $X_{\Lambda^k V}$ in terms of $X_V$:

$$X_{\Lambda^1 V}(g) = X_V(g)$$

$$X_{\Lambda^2 V}(g) = \frac{1}{2}(X_V(g)^2 - X_V(g^2))$$

$$X_{\Lambda^3 V}(g) = \frac{1}{3!}(X_V(g)^3 - 3X_V(g)X_V(g^2) + 2X_V(g^3))$$

$$\ldots$$

□
3.7.4. An Application to Invariant Theory

Let $G$ be a group, not necessarily finite, and let $V \in \text{Rep}_k G$ be given. Then the tensor algebra $TV$, the symmetric algebra $\text{Sym} V$, and the exterior algebra $\Lambda V$ are graded $G$-algebras, that is, $G$ acts by graded $k$-algebra automorphisms on these algebras (§3.3.3). For any $G$-algebra $A$, the $G$-invariants $A^G = \{ a \in A \mid g.a = a \text{ for all } g \in G \}$ are easily seen to be a subalgebra of $A$, the so-called algebra of $G$-invariants in $A$. Classical invariant theory is mostly concerned with the symmetric algebra, $A = \text{Sym} V$ with $V \in \text{Rep}_k G$, and seeks to determine the structure of the algebra of $G$-invariants, $$(\text{Sym} V)^G = \bigoplus_{k \geq 0} (\text{Sym}^k V)^G.$$

In practice, one often considers the dual representation $V^*$ instead of $V$. From the viewpoint of representation theory, the difference is immaterial. The algebra $O(V) = \text{Sym} V^*$ is called the algebra of polynomial functions on $V$ (Section C.3) and the invariant algebra $O(V)^G$ is called an algebra of polynomial invariants.

In this subsection, we present a classical result from invariant theory, Molien’s Theorem. Under the assumption that $kG$ is semisimple, the theorem gives an expression for the Hilbert series $14$ of the invariant algebra $(\text{Sym} V)^G$ and, more generally, of any homogeneous component $(\text{Sym} V)(S)$ with $S \in \text{Irr}_k G$. Here, the Hilbert series of a $\mathbb{Z}_+$-graded $k$-vector space $W = \bigoplus_{k \geq 0} W^k$ with finite-dimensional homogeneous components $W^k$ is defined by $$H(W, t) \overset{\text{def}}{=} \sum_{k \geq 0} \dim_k W^k t^k \in \mathbb{Z}[t].$$

Molien’s Theorem concerns $W = (\text{Sym} V)(S) = \bigoplus_{k \geq 0} (\text{Sym}^k V)(S)$, the case $S = 1$ being especially relevant. Indeed, a priori knowledge of the Hilbert series $H((\text{Sym} V)^G, t) = \sum_{k \geq 0} \dim_k (\text{Sym}^k V)^G t^k$ is an important tool of computational invariant theory; see the monograph by Derksen and Kemper [55]. Molien’s Theorem will be stated in terms of $k$-valued characters and determinants; so the theorem gives the images of Hilbert series $\in \mathbb{Z}[t]$ in $k[t]$.

**Molien’s Theorem.** Let $G$ be a finite group with $\text{char} k \nmid |G|$ and let $V \in \text{Rep}_k G$. Then, for every $S \in \text{Irr}_k G$, $$H((\text{Sym} V)(S), t) \cdot 1_k = \frac{\dim_{D(S)}}{|G|} \sum_{g \in G} \frac{\chi_S(g^{-1})}{\det(\text{Id}_V - g_V t)}.$$ In particular, the Hilbert series of the invariant algebra $(\text{Sym} V)^G$ is given by $$H((\text{Sym} V)^G, t) \cdot 1_k = \frac{1}{|G|} \sum_{g \in G} \frac{1}{\det(\text{Id}_V - g_V t)}.$$

---

14Hilbert series are also called Poincaré series.
Proof. By Proposition 3.22(c), the formula \( \dim_k M(S) \cdot 1_k = \dim_{D(S)} S \cdot (\chi_S \cdot \chi_M) \) holds for every \( M \in \text{Rep}_{fin} kG \). Therefore,

\[
\sum_{k \geq 0} \dim_k (\text{Sym}^k V)(S) \cdot 1_k t^k = \dim_{D(S)} S \cdot \sum_{k \geq 0} \left( \chi_S \cdot \chi_{\text{Sym}^k V} \right) t^k
= \dim_{D(S)} S \cdot \sum_{k \geq 0} \frac{1}{|G|} \sum_{g \in G} \chi_S(g^{-1}) \chi_{\text{Sym}^k V}(g) t^k
= \dim_{D(S)} S \cdot \frac{1}{|G|} \sum_{g \in G} \chi_S(g^{-1}) \sum_{k \geq 0} \chi_{\text{Sym}^k V}(g) t^k
\]

This proves the formula for \( H((\text{Sym} V)(S), t) \). The Hilbert series for \( (\text{Sym} V)^G \) is the special case \( S = 1 \). \( \square \)

Example 3.35. The following example is borrowed from the same article where Molien’s Theorem was published \([152]\).

Let \( G \) denote the group of rotational symmetries of the regular icosahedron in \( \mathbb{R}^3 \). There are 12 vertices, 30 edges and 20 faces all of which come in opposite pairs. Besides the identity, the group \( G \) contains the rotations by \( \frac{2\pi}{5} k \) with \( 1 \leq k \leq 4 \) about the axis through any pair of opposite vertices. In addition, there are the rotations by \( \pi \) about the axis through the midpoints of any pair of opposite edges as well as the rotations by \( \pm \frac{2\pi}{3} \) about the axis through the midpoints of any pair of opposite faces. Altogether, this accounts for \( 1 + 6 \cdot 4 + 15 + 2 \cdot 10 = 60 \) distinct elements of \( G \). In fact, there are no further elements. For, any rotational symmetry must send a chosen vertex \( v \) to one of the 12 vertices, say \( v' \); a chosen vertex adjacent to \( v \) must be sent to one of the 5 vertices adjacent to \( v' \); and any of these 12 \cdot 5 choices determines the rotation. Thus, \( G \) has order 60. It is not hard to see that \( G \) is isomorphic to the alternating group \( \mathcal{A}_5 \), but we shall not use this fact now.

Placing the center of gravity of the icosahedron at the origin, we obtain an inclusion of \( G \) into the rotation group \( \text{SO}_3(\mathbb{R}) \), which defines a degree-3 representation \( V \) of \( G \) over \( k = \mathbb{R} \). Following Molien, we determine the Hilbert series \( H((\text{Sym} V)^G, t) \). To this end, we need to find \( \det(\text{Id}_V - g_V t) \) for all \( g \in G \). Put \( \zeta_n = e^{2\pi i/n} \in \mathbb{C} \). If \( g \) is one of the six \( \frac{2\pi}{5} k \)-rotations, then the eigenvalues of \( g_V \) are \( 1, \zeta_5^k, \zeta_5^{-k} \), and hence

\[
\det(\text{Id}_V - g_V t) = (1 - t)(1 - \zeta_5^k t)(1 - \zeta_5^{-k} t).
\]
Similarly, since \( \zeta_2 = -1 \), each of the fifteen \( \pi \)-rotations gives
\[
\det(\text{Id}_V - g_V t) = (1 - t)(1 + t)^2 = (1 - t^2)(1 + t)
\]
and each of the twenty \( \pm \frac{2\pi}{3} \)-rotations contributes
\[
\det(\text{Id}_V - g_V t) = (1 - t)(1 - \zeta_3^{\pm 1} t)(1 - \zeta_3^{\mp 1} t) = 1 - t^3.
\]
Therefore, also taking the identity element into account, Molien’s Theorem yields
the following expression for \( H((\text{Sym} V)^G, t) \):
\[
\frac{1}{60} \left( \frac{1}{(1 - t)^3} + \frac{15}{(1 - t^2)(1 + t)} + \frac{20}{1 - t^3} + \frac{6}{1 - t} \sum_{k=1}^{4} \frac{1}{(1 - \zeta_3^k t)(1 - \zeta_3^{\mp k} t)} \right).
\]
This formula suggests—without actually proving it, but see Exercise 3.7.3—that
the invariant algebra has the form
\[
(\text{Sym} V)^G = P \oplus Ps
\]
for algebraically independent homogeneous invariants \( a, b, c \) of respective degrees 2, 6, 10
and \( s \) homogeneous degree 15. This is indeed the structure of the invariant algebra
in this case; see [187, p. 99].

**Exercises for Section 3.7**

**3.7.1 (Newton formulae)**. Let \( \operatorname{char} k = 0 \) and let \( V \in \text{Rep}_{\text{fin}} kG \). Prove the analogs of (3.63) for the characters \( \chi_{\text{Sym}^k V} \):
\[
\chi_{\text{Sym}^k V}(g) = \chi_V(g)
\]
\[
\chi_{\text{Sym}^k V}(g) = \frac{1}{k!} (\chi_V(g)^2 + \chi_V(g^2))
\]
\[
\chi_{\text{Sym}^k V}(g) = \frac{1}{3!} (\chi_V(g)^3 + 3 \chi_V(g) \chi_V(g^2) + 2 \chi_V(g^3))
\]

**3.7.2 (Some representations of \( kS_3 \) in characteristic 0)**. In this problem, \( V = V_2 \)
denotes the standard representation of the symmetric group \( S_3 \) and we assume \( \operatorname{char} k = 0 \) for simplicity. Recall that \( \text{Irr} kS_3 = \{ 1, \text{sgn}, V \} \).

(a) Show that \( \text{Sym}^{k+6} V \cong \text{Sym}^k V \oplus (kS_3)_{\text{reg}} \) in \( \text{Rep} kS_3 \). (Use Lemma 3.33.)

(b) For \( 0 \leq k \leq 5 \), find the decomposition of \( \text{Sym}^k V \) into irreducible representations of \( kS_3 \).

(c) Show that \( \text{Sym}^2(\text{Sym}^3 V) \cong \text{Sym}^3(\text{Sym}^2 V) \) in \( \text{Rep} kS_3 \) and find the decomposition of this representation into irreducible constituents.
3.7.3 (Hilbert series). This exercise concerns graded \( k \)-vector spaces \( V = \bigoplus_{k \in \mathbb{Z}} V^k \) such that all \( V^k \) are finite dimensional and \( V^k = 0 \) for \( k < 0 \). The Hilbert series of \( V \) is the Laurent power series 
\[
H(V, t) = \sum_k \dim_k V^k t^k.
\]
(a) For graded vector spaces \( V \) and \( W \), show that \( H(V \otimes W, t) = H(V, t)H(W, t) \) and 
\[
H(V \oplus W, t) = H(V, t) + H(W, t),
\]
where \( V \otimes W \) and \( V \oplus W \) carry the standard gradings (Exercise 1.1.11).

(b) If \( 0 \to V_1 \to V_2 \to \cdots \to V_r \to 0 \) is an exact sequence of graded vector spaces, with degree-preserving maps, then 
\[
\sum_i (-1)^i H(V_i, t) = 0.
\]
(c) Let \( A = k[x_1, x_2, \ldots, x_n] \) be the polynomial algebra, graded by \( \deg x_i = d_i \) for given positive integers \( d_i \). Show that 
\[
H(A, t) = \prod_i \frac{1}{1 - t^{d_i}}.
\]
(d) Let \( A = \bigoplus_k A^k \) be a graded \( k \)-algebra and assume that \( V \) is a graded free left \( A \)-module having a finite basis \( b_1, \ldots, b_m \) consisting of homogeneous elements, say \( \deg b_j = m_j \). Then 
\[
H(V, t) = H(A, t) \cdot \sum_j t^{m_j}.
\]
In particular, with \( A \) as in (c), we have 
\[
H(V, t) = \frac{f(t)}{\prod_i 1 - t^{d_i}}
\]
for some Laurent polynomial \( f(t) \in \mathbb{Z}[t^{\pm 1}] \).

3.7.4 (Hilbert-Serre Theorem). Let \( A = \bigoplus_k A^k \) be an affine commutative graded \( k \)-algebra, with homogeneous algebra generators \( x_1, x_2, \ldots, x_n \) of degrees \( \deg x_i = d_i \) for given positive integers \( d_i \). Then, for any finitely generated graded left \( A \)-module \( V \), we have 
\[
H(V, t) = \frac{f(t)}{\prod_i 1 - t^{d_i}}
\]
for some Laurent polynomial \( f(t) \in \mathbb{Z}[t^{\pm 1}] \).

3.7.5 (Molien’s Theorem). (a) Let \( G = \langle g \rangle \cong C_4 \) act on the polynomial algebra \( k[x, y] \) by \( g.x = -y \) and \( g.y = x \) and assume that \( \text{char} k = 0 \). Use Molien’s Theorem to show that 
\[
H(k[x, y]^G, t) = \frac{1}{4} \left( \frac{1}{(1 - t^3)} + \frac{2}{1 + t^2} + \frac{1}{1 - t^4} \right) = \frac{1 + t^4}{(1 - t^3)(1 - t^4)}.
\]
Use Exercise 3.7.3 to conclude that the invariant algebra has the form 
\[
k[x, y]^G = k[x^2 + y^2, x^2 y^2] \oplus k[x^2 + y^2, x^2 y^2](x^3 y - y x^3).
\]
(b) Let \( G = \langle g \rangle \cong C_3 \) act on the polynomial algebra \( \mathbb{C}[x, y] \) by \( g.x = \zeta_3 x \) and \( g.y = \zeta_3^2 y \), where \( \zeta_3 = e^{2\pi i/3} \in \mathbb{C} \). Use Molien’s Theorem to show that 
\[
H(k[x, y]^G, t) = \frac{1}{3} \left( \frac{1}{(1 - t^3)} + \frac{2}{(1 - \zeta_3 t)(1 - \zeta_3^2 t)} \right) = \frac{1 - t + t^2}{(1 - t)(1 - t^3)}.
\]
Show that this expression can be rewritten in the following alternative forms:
\[
H(k[x, y]^G, t) = \frac{1 + t^2 + t^4}{(1 - t^3)^2} = \frac{1 + t^3}{(1 - t^3)(1 - t^4)}.
\]
Conclude that 
\[
k[x, y]^G = k[x^3, y^3] \oplus k[x^3, y^3]xy \oplus k[x^3, y^3]x^2 y^2
\]
\[
= k[xy, x^3 + y^3] \oplus k[xy, x^3 + y^3]x^2 y^2.
\]
3.8. Decomposing Tensor Powers

We conclude this chapter by having an initial look at the “place permutation” action of the symmetric group $S_n$ on the $n^{th}$ tensor power of a given representation $V \in \text{Rep}_k G$ for some group $G$. The $S_n$-action and the resulting decomposition of $V^\otimes n$ will be further explored in Sections 4.7 and 8.8.

Throughout this section, $G$ is an arbitrary group and $V \in \text{Rep}_k G$.

3.8.1. Symmetric and Antisymmetric Tensors

Our ultimate goal is to give an explicit decomposition of the $n^{th}$ tensor power $V^\otimes n \in \text{Rep}_k G$ into certain subrepresentations. To this end, we take advantage of the fact that, in addition to the diagonal action $G V^\otimes n$, we also have an action $S_n V^\otimes n$ by place permutations:

\begin{equation}
\sigma(v_1 \otimes v_2 \otimes \cdots \otimes v_n) = v_{\sigma^{-1}1} \otimes v_{\sigma^{-1}2} \otimes \cdots \otimes v_{\sigma^{-1}n} \quad (\sigma \in S_n).
\end{equation}

This action manifestly commutes with the diagonal action of $G$ and it makes $V^\otimes n$ a $kS_n$-representation. This allows us to regard $V^\otimes n \in \text{Rep}_k [G \times S_n]$.

**Decomposition.** For the remainder of this subsection, assume that $\text{char } k \nmid n!$. Then the group algebra $kS_n$ is semisimple by Maschke’s Theorem and so $V^\otimes n \in \text{Rep}_k kS_n$ is completely reducible and can be decomposed into homogenous components. Furthermore, all these homogenous components are also stable under the action $G \subset V^\otimes n$ by Proposition 1.31(b). Thus, we obtain the following decomposition in $\text{Rep}_k [G \times S_n]$:

\begin{equation}
V^\otimes n = \bigoplus_{S \in \text{Irr } kS_n} (V^\otimes n)(S)
\end{equation}

We will describe the components $(V^\otimes n)(S)$ in detail later; see (4.54).

**Symmetrization.** For now, let us focus on the two degree-1 representations, $S = 1$ and $S = \text{sgn}$. The weight spaces $(V^\otimes n)(1) = (V^\otimes n)^S_n$ and $(V^\otimes n)(\text{sgn})$ are commonly referred to as the spaces of symmetric and antisymmetric $n$-tensors, respectively. We will use the following notation:

\[
\begin{align*}
\text{ST}^n V & \overset{\text{def}}{=} (V^\otimes n)(1) = (V^\otimes n)^{S_n} \\
\text{AT}^n V & \overset{\text{def}}{=} (V^\otimes n)(\text{sgn})
\end{align*}
\]
The projections (3.44) onto the corresponding components of (3.65) are the following maps in \( \text{Rep} \, \mathbb{k}G \):

\[
\mathcal{J} : V^\otimes n \rightarrow ST^n V \quad \text{and} \quad \mathcal{A} : V^\otimes n \rightarrow \Lambda^n V
\]

\[
(3.66)
\]

\[
x \mapsto \frac{1}{n!} \sum_{s \in S_n} s.x \quad \text{and} \quad x \mapsto \frac{1}{n!} \sum_{s \in S_n} \text{sgn}(s)s.x
\]

The projections \( \mathcal{J} \) and \( \mathcal{A} \) are called the symmetrization and antisymmetrization maps, respectively; the corresponding primitive central idempotents \( e(\mathcal{J}) = \frac{1}{n!} \sum_{s \in S_n} s \) and \( e(\mathcal{A}) = \frac{1}{n!} \sum_{s \in S_n} \text{sgn}(s)s \) of \( \mathbb{k}S_n \) are called symmetrizer and antisymmetrizer.

**Isomorphism with Symmetric and Exterior Powers.** Recall from §3.3.3 that we have canonical epimorphisms in \( \text{Rep} \, \mathbb{k}G \),

\[
V^\otimes n \rightarrow \text{Sym}^n V, \quad v_1 \otimes v_2 \otimes \cdots \otimes v_n \mapsto v_1 v_2 \cdots v_n
\]

and

\[
V^\otimes n \rightarrow \Lambda^n V, \quad v_1 \otimes v_2 \otimes \cdots \otimes v_n \mapsto v_1 \wedge v_2 \wedge \cdots \wedge v_n.
\]

The foregoing allows us to identify \( \text{Sym}^n V \) and \( \Lambda^n V \) with the spaces of symmetric and antisymmetric \( n \)-tensors, respectively:

**Lemma 3.36.** Assume that \( \text{char} \, \mathbb{k} = 0 \) or \( \text{char} \, \mathbb{k} > n \). Then:

(a) The map \( V^\otimes n \rightarrow \text{Sym}^n V \) factors through \( \mathcal{J} \) and its restriction to symmetric \( n \)-tensors is an isomorphism \( ST^n V \rightarrow \text{Sym}^n V \) in \( \text{Rep} \, \mathbb{k}G \).

(b) The map \( V^\otimes n \rightarrow \Lambda^n V \) factors through \( \mathcal{A} \) and its restriction to antisymmetric \( n \)-tensors is an isomorphism \( \text{AT}^n V \rightarrow \Lambda^n V \) in \( \text{Rep} \, \mathbb{k}G \).

**Proof.** (a) The kernel of \( V^\otimes n \rightarrow \text{Sym}^n V \) is \( I \cap V^\otimes n \), where \( I \) is the ideal of the tensor algebra \( TV \) that is generated by the elements \( v \otimes v' \cdot v' \otimes v \) with \( v, v' \in V \) (§1.1.2). Since \( V^\otimes n = ST^n V \oplus \text{Ker} \, \mathcal{J} \), it suffices to show that \( I \cap V^\otimes n = \text{Ker} \, \mathcal{J} \). If \( x \in \text{Ker} \, \mathcal{J} \), then \( x = x - \mathcal{J}x = \frac{1}{n!} \sum_{s \in S_n} (x - s.x) \). Since \( \text{Sym} V \) is commutative, \( x \) and each \( s.x \) have the same image in \( \text{Sym}^n V \). It follows that \( x \) maps to \( 0 \in \text{Sym}^n V \), whence \( I \cap V^\otimes n \supseteq \text{Ker} \, \mathcal{J} \). For the reverse inclusion, it suffices to show that \( \mathcal{J}x = 0 \) for \( x = y \otimes (v \otimes v' - v' \otimes v) \otimes z \) with \( y, v, v', v' \in V^\otimes r \) and \( z \in V^\otimes (n-r-2) \). Note that the transposition \( t = (r + 1, r + 2) \in S_n \) satisfies \( t.x = -x \). Since \( S_n = \mathcal{A}_n \sqcup \mathcal{A}_n t \), we obtain \( \mathcal{J}x = \frac{1}{n!} \sum_{s \in \mathcal{A}_n} (s.x + st.x) = \frac{1}{n!} \sum_{s \in \mathcal{A}_n} (s.x - s.x) = 0 \), as desired.

(b) The argument for \( \Lambda^n V \) is analogous, using the ideal \( J = \langle v \otimes v | v \in V \rangle \) of \( TV \) in place of \( I \) and the antisymmetrizer \( \mathcal{A} \) instead of \( \mathcal{J} \). The inclusion \( \text{Ker} \, \mathcal{A} \subseteq J \cap V^\otimes n \) holds because the images of \( x \) and \( \text{sgn}(s)x \) in \( \Lambda^n V \) are identical for all \( x \in V^\otimes n \), \( s \in S_n \) by (1.12). For \( J \cap V^\otimes n \subseteq \text{Ker} \, \mathcal{A} \), it suffices to show that \( \mathcal{A}x = 0 \) for \( x = y \otimes v \otimes z \) with \( y \in V^\otimes r \) and \( z \in V^\otimes (n-r-2) \). Now the transposition
\[ t = (r + 1, r + 2) \in S_n \] satisfies \( t \cdot x = x \) and so \( \text{sgn}(s)s \cdot x + \text{sgn}(st)st \cdot x = \text{sgn}(s)s \cdot x - \text{sgn}(s)s \cdot x = 0 \) for all \( s \). Therefore, \( \mathcal{A}x = \frac{1}{n} \sum_{s \in \mathcal{A}_n} (\text{sgn}(s)s \cdot x + \text{sgn}(st)st \cdot x) = 0 \), which completes the proof. \( \square \)

Exercise 3.8.1 interprets the isomorphism \( \text{ST}^n V \cong \text{Sym}^n V \) as an isomorphism between \( S_n \)-invariants and \( S_n \)-coinvariants.

The Case \( n = 2 \). If \( n = 2 \) and \( \text{char} \, \mathbb{k} \neq 2 \), then the representations \( \mathbb{1} \) and \( \text{sgn} \) comprise all of \( \text{Irr} \, \mathbb{k}S_n \). Thus, in view of Lemma 3.36, the decomposition (3.65) takes the following form, which in particular justifies a claim made in §3.5.2:

\[
V^{\otimes 2} = \text{ST}^2 V \oplus \text{ST}^2 V \cong \text{Sym}^2 V \oplus \Lambda^2 V.
\]

3.8.2. Young Modules and Young Subgroups

We now focus on the case where \( d = \dim_{\mathbb{k}} V < \infty \). The base field \( \mathbb{k} \) can be arbitrary for now. Fix a \( \mathbb{k} \)-basis \( (x_j)^d_1 \) of \( V \). Then the monomials

\[
x_I = x_{i_1} \otimes x_{i_2} \otimes \cdots \otimes x_{i_n}
\]

with \( I = (i_1, \ldots, i_n) \in X := [d]^n \) form a \( \mathbb{k} \)-basis of \( V^{\otimes n} \) that is permuted by the \( S_n \)-action: \( s \cdot x_I = x_{s \cdot I} \) with \( s \cdot I = (i_{s^{-1}1}, \ldots, i_{s^{-1}n}) \). Therefore, \( V^{\otimes n} \cong \mathbb{k}X \) is a permutation representation of \( S_n \). A transversal for the orbit set \( S_n \backslash X \) is given by the sequences

\[
I_m = (1, 2, \ldots, 2, \ldots, 2, \ldots, 2)\left(\begin{array}{llll}
m_1 & m_2 & \cdots & m_d
\end{array}\right)
\]

with \( m = (m_1, m_2, \ldots, m_d) \in \mathbb{Z}_+^d \) and \( |m| := \sum m_j = n \). The isotropy group of \( I_m \) is the subgroup of \( S_n \) consisting of those permutation of \([n]\) that stabilize all subsets \( \{1, \ldots, m_1\}, \{m_1 + 1, \ldots, m_1 + m_2\}, \ldots, \{m_1 + \cdots + m_{d-1} + 1, \ldots, n\} \). Thus, denoting this subgroup by \( S_m \), we obtain the following decomposition of \( V^{\otimes n} \) in \( \text{Rep} \, \mathbb{k}S_n \), with one summand for each \( S_n \)-orbit:

\[
V^{\otimes n} \cong \bigoplus_{m \in \mathbb{Z}_+^d : |m| = n} \mathbb{k}[S_n/S_m]
\]

The summands \( \mathbb{k}[S_n/S_m] \) are referred to as Young modules and the groups \( S_m \) as Young subgroups of \( S_n \). In combinatorics, a \( d \)-tuple \( m = (m_1, m_2, \ldots, m_d) \in \mathbb{Z}_+^d \) such that \( |m| = n \) is called a weak composition of \( n \) into \( d \) parts; if all \( m_i \) are nonzero, then \( m \) is called a composition. Omitting all zero-parts from \( m \), we obtain a composition \( m' \) of \( n \) with possibly fewer than \( d \) parts, but the Young subgroup remains unchanged: \( S_{m'} = S_m \). More generally, if \( M = M_1 \sqcup M_2 \sqcup \cdots \sqcup M_l \) for a set \( M \) and subsets \( M_i \subseteq M \), then we obtain a Young subgroup \( Y \leq S_M \), consisting of
those permutations of $M$ that stabilize all $M_i$. Clearly, we may remove all $M_i = \emptyset$ without changing $Y$ and

$$Y \cong S_{M_1} \times S_{M_2} \times \cdots \times S_{M_l}.$$  

### 3.8.3. Generating Symmetric Tensors

The $n^{\text{th}}$ power $v^{\otimes n} = v \otimes v \otimes \cdots \otimes v \in V^{\otimes n}$ of any $v \in V$ is evidently symmetric. Assuming the base field $\mathbb{k}$ to be infinite\(^{15}\) and $\dim_{\mathbb{k}} V < \infty$, we will show that the space $\text{ST}^n V$ of symmetric $n$-tensors is generated by powers. The proof will make use of some basic facts concerning polynomial functions and the Zariski topology on $V$; see Section C.3. As usual, $O(V) = \text{Sym} V^*$ denotes the algebra of polynomial functions on $V$. Since $\mathbb{k}$ is infinite, we have the embedding $\Phi: O(V) \hookrightarrow \mathbb{k}^V = \{\text{functions } V \to \mathbb{k}\}$ as in (C.1). Thus, for any $0 \neq f \in O(V)$, the principal open subset $V_f = \{v \in V \mid (\Phi f)(v) \neq 0\}$ is nonempty.

**Proposition 3.37.** Let $V$ be finite dimensional and assume that $\mathbb{k}$ is infinite. Then $\text{ST}^n V = \langle v^{\otimes n} \mid v \in V \rangle_{\mathbb{k}}$; in fact, it suffices to let $v$ range over a Zariski-dense subset of $V$.

**Proof.** Let $(x_i^d)^d$ be a $\mathbb{k}$-basis of $V$ and let $D \subseteq V$ be a Zariski-dense subset. The basis (3.68) of $V^{\otimes n}$ and the description (3.69) of the $S_n$-orbits on this basis yields a $\mathbb{k}$-basis of $\text{ST}^n V$: by (3.24), a basis is given by the orbit sums

$$\sigma_m = \sum_{s \in S_n/S_m} x_s I_m$$

with $m = (m_1, m_2, \ldots, m_d) \in \mathbb{Z}_+^d$ and $|m| = n$. Our goal is to show that the vector space $\text{ST}^n V$ is generated by the elements $v^{\otimes n}$ with $v \in D$. This in turn amounts to showing that, for any linear form $0 \neq l \in (\text{ST}^n V)^*$, there is some $v \in D$ such that $\langle l, v^{\otimes n} \rangle \neq 0$. So let us fix $0 \neq l \in (\text{ST}^n V)^*$ and let $(x_i^d)^d$ be the dual basis of $V^*$ for the given basis of $V$. Writing $v = \sum_i \lambda_i x_i = \sum_i \langle x_i, v \rangle x_i$, the development of $v^{\otimes n}$ in terms of the above basis $(\sigma_m)$ is as follows:

$$v^{\otimes n} = \sum_{m \in \mathbb{Z}_+^d : |m| = n} \lambda^m \sigma_m \quad \text{with} \quad \lambda^m = \lambda_1^{m_1} \lambda_2^{m_2} \cdots \lambda_d^{m_d}.$$ 

Put $h = h(l) := \sum_{|m| = n} x^m \langle l, \sigma_m \rangle \in O^n(V)$, where $x^m = (x_1)^{m_1} \cdots (x_d)^{m_d}$. Then $h \neq 0$, because $\langle l, \sigma_m \rangle \neq 0$ for some $m$ and the standard monomials $x^m$ are $\mathbb{k}$-independent. Therefore, the principal open subset $V_h$ of $V$ is nonempty, and so we also have $D \cap V_h \neq \emptyset$. The function $\Phi h \in \mathbb{k}^V$ is given by

$$(\Phi h)(v) = \sum_{|m| = n} x^m \langle l, \sigma_m \rangle = \langle l, v^{\otimes n} \rangle.$$ 

Therefore, $\langle l, v^{\otimes n} \rangle \neq 0$ for some $v \in D$, as desired.  

\(^{15}\)See Exercise 3.8.3 for finite fields.
3.8.4. Polarization

Finally, we briefly discuss another invariant theoretic theme. We continue to assume that $V$ is finite dimensional.

**Forms and Multilinear Maps.** The homogeneous component $O^n(V) = \text{Sym}^n V^*$ is called the space of *forms of degree* $n$ on $V$ in invariant theory. We wish to relate degree-$n$ forms to multilinear maps $l: V^n = V \times V \times \ldots \times V \rightarrow \mathbb{k}$ (§B.1.3).

Consider the action $G \subseteq \text{MultLin}(V^n, \mathbb{k})$ that is defined by

$$(g.l)(v_1, \ldots, v_n) = l(g^{-1}.v_1, \ldots, g^{-1}.v_n).$$

The following map is an isomorphism in $\text{Rep}_kG$:

$$(V^*)^\otimes n \xrightarrow{\sim} (V^\otimes n)^* \xrightarrow{\sim} \text{MultLin}(V^n, \mathbb{k})$$

(3.71)

$$f_1 \otimes \cdots \otimes f_n \longmapsto \left((v_1, \ldots, v_n) \mapsto \prod_i \langle f_i, v_i \rangle \right)$$

**Symmetric Multilinear Maps.** Similarly, the place permutation action $S_n \subseteq V^n$, given by $s.(v_1, v_2, \ldots, v_n) = (v_{s^{-1}1}, v_{s^{-1}2}, \ldots, v_{s^{-1}n})$ for $s \in S_n$, gives rise to the following action $S_n \subseteq \text{MultLin}(V^n, \mathbb{k})$:

$$(s.l)(v_1, \ldots, v_n) = l(v_{s1}, v_{s2}, \ldots, v_{sn}).$$

The isomorphism (3.71) is also equivariant for this action and the place permutation action (3.64) of $S_n$ on $(V^*)^\otimes n$. Indeed, denoting the image of $f \in (V^*)^\otimes n$ under the map (3.71) by $I_f$, we calculate

$$(s.I_f)(v_1, \ldots, v_n) = I_f(v_{s1}, v_{s2}, \ldots, v_{sn}) = \prod_i \langle f_i, v_{si} \rangle = \prod_i \langle f_{s^{-1}i}, v_i \rangle.$$

A multilinear map $l: V^n \rightarrow \mathbb{k}$ is said to be symmetric if $l$ is $S_n$-invariant, that is, $l$ is constant on $S_n$-orbits in $V^n$. Thus, in (3.71), the $S_n$-invariants $\text{ST}^n(V^*) = ((V^*)^\otimes n)^{S_n}$ correspond to the symmetric multilinear maps $V^n \rightarrow \mathbb{k}$.

**Polarization and Restitution.** Now let us assume that $\text{char} \mathbb{k} = 0$ or $\text{char} \mathbb{k} > n$. Then we have the following isomorphisms in $\text{Rep}_kG$:

$$(3.72)$$

$$\mathcal{P}: O^n(V) \xrightarrow{\text{Lemma 3.36}} \text{ST}^n(V^*) \xrightarrow{\sim} \left\{ \text{symmetric multilinear maps } V^n \rightarrow \mathbb{k} \right\}$$

In the invariant theory literature, this isomorphism is called *polarization* and its inverse, *restitution*; see Weyl [205, p. 5ff] or Procesi [168, p. 40ff]. To find the polarization $\mathcal{P}f$ of a given $f \in O^n(V)$, choose any preimage $t = t_f \in (V^*)^\otimes n$ for $f$.
and symmetrize it. Letting $O_t = S_n.t$ denote the $S_n$-orbit of $t$, the symmetrization (3.66) may be written as

$$
\mathcal{S}_t = \frac{1}{|O_t|} \sum_{t' \in O_t} t' \in \text{ST}^n(V^*).
$$

Then $\mathcal{P} f$ is the image of $\mathcal{S} t$ in $\text{MultLin}(V^n, \mathbb{k})$ under the map (3.71).

**Examples 3.38** (Some polarizations). Fix a basis $x_1, \ldots, x_d$ of $V^*$ and identify $V$ with $\mathbb{k}^d$ via $v \leftrightarrow ((x^i, v))_i$. For simplicity, let us write $x = x^1$ and $y = x^2$. The element $x^2 \in \mathcal{O}^2(V)$ comes from $x \otimes x \in (V^*)^\otimes 2$, which is already symmetric. The tensor $x \otimes x$ corresponds in (3.71) to the symmetric bilinear map $V^2 \to \mathbb{k}$ that is given by $((\xi_1, \ldots), (\xi_2, \ldots)) \mapsto \xi_1 \xi_2$. Denoting this bilinear map by $x_1 x_2$, we obtain

$$
\mathcal{P}(x^2) = x_1 x_2.
$$

For $xy \in \mathcal{O}^2(V)$, we need to symmetrize, $\mathcal{P}(x \otimes y) = \frac{1}{2} (x \otimes y + y \otimes x) \in \text{ST}^2(V^*)$, which corresponds to the bilinear map $((\xi_1, \eta_1, \ldots), (\xi_2, \eta_2, \ldots)) \mapsto \frac{1}{2} (\xi_1 \eta_2 + \eta_1 \xi_2)$. Using the above notational convention, this yields

$$
\mathcal{P}(xy) = \frac{1}{2} (x_1 y_2 + x_2 y_1).
$$

Similarly, $\mathcal{P}(x^2 y)$ comes from $\mathcal{P}(x \otimes x \otimes y) = \frac{1}{2} (x \otimes x \otimes y + x \otimes y \otimes x + y \otimes x \otimes x)$, whence

$$
\mathcal{P}(x^2 y) = \frac{1}{3} (x_1 x_2 y_3 + x_1 y_2 x_3 + y_1 x_2 x_3).
$$

For future reference, we give an explicit description of the restitution map. By our hypothesis on $\mathbb{k}$, the canonical map $O^n(V) \to \mathbb{k}^V$ is an embedding (Exercise C.3.2); so we may describe elements of $O^n(V)$ as functions on $V$.

**Lemma 3.39.** The preimage of a symmetric multilinear map $l: V^n \to \mathbb{k}$ under the isomorphism (3.72) is the degree-$n$ form $(v \mapsto l(v, v, \ldots)) \in O^n(V)$.

**Proof.** Let $(x_i)_i$ be a basis of $V$ and let $(x^i)_i$ be the dual basis of $V^*$. For $I = (i_1, \ldots, i_n) \in [d]^n$, put $x^I = x^{i_1} \otimes \cdots \otimes x^{i_n} \in (V^*)^\otimes n$ and let $x^{i_1} \cdots x^{i_n}$ denote the multilinear map $V^n \to \mathbb{k}$ that corresponds to $x^I$ under the isomorphism (3.71). Then $l = \sum_I \lambda_I x^I$ with $\lambda_I = l(x_{i_1}, \ldots, x_{i_n})$. The preimage of $l$ in (3.71) is $\sum_I \lambda_I x^I \in \text{ST}^n(V^*)$ and the image of $\sum_I \lambda_I x^I$ under the isomorphism of Lemma 3.36 is the degree-$n$ form $f_I := \sum_I \lambda_I x^{i_1} \cdots x^{i_n} \in O^n(V)$. Thus, $\mathcal{P} f_I = l$. 

The following computation shows that \( l(v, \ldots, v) = f_l(v) \), proving the lemma:

\[
l(v, \ldots, v) = l(\sum_i \langle x^i, v \rangle x_i, \ldots, \sum_i \langle x^i, v \rangle x_i)
= \sum_i \langle x^i, v \rangle \cdots \langle x^i, v \rangle l(x_{i_1}, \ldots, x_{i_n})
= \sum_i x^{i_1} \cdots x^{i_n}(v) \Lambda_f. \quad \square
\]

**Exercises for Section 3.8**

3.8.1 (\( S_n \)-coinvariants). (a) Consider the coinvariants \( (V^\otimes n)_{S_n} \) (Exercise 3.3.3) for the place permutation action \( S_n \subset V^\otimes n \). Show the canonical map \( V^\otimes n \twoheadrightarrow (V^\otimes n)_{S_n} \) is a morphism in \( \text{Rep}_k G \), which has the same kernel as the canonical map \( V^\otimes n \twoheadrightarrow \text{Sym}^n V \). Thus, \( \text{Sym}^n V \cong (V^\otimes n)_{S_n} \) in \( \text{Rep}_k G \).

(b) For \( V \in \text{Rep}_{\text{fin}} kG \), show that \( (\text{Sym}^n V)^* \cong \text{Sym}^n (V^*) \) in \( \text{Rep}_k G \) provided \( \text{char} k \nmid n! \) and always \( (\wedge^n V)^* \cong \wedge^n (V^*) \) in \( \text{Rep}_k G \).

3.8.2 (Groups of odd order). Let \( G \) be finite of odd order and let \( \text{char} k = 0 \). For any \( 1 \neq S \in \text{Irr}_k G \), show:

(a) \( (1, \chi_{\text{Sym}^2 S}) = (1, \chi_{\Lambda^2 S}). \) (Use (3.63) and Exercises 3.7.1 and 3.6.3.)

(b) \( S \neq S^* \). (Use (3.67) and part (a).)

(c) The number \( c \) of conjugacy classes of \( G \) satisfies \(|G| \equiv c \mod 16\).

3.8.3 (Generating tensor powers). Prove the following statements by modifying the proof of Proposition 3.37 and referring to Exercise C.3.2:

(a) \( (V^\otimes n)_{S_n} = \langle v^\otimes n \mid v \in V \rangle_k \) provided \(|k| \geq n\).

(b) Let \( 0 \neq f \in O^n(V) \). If \(|k| \geq n + m\), then \( (V^\otimes n)^{S_n} = \langle v^\otimes n \mid v \in V_f \rangle_k \).
Chapter 4

Symmetric Groups

This chapter presents the representation theory of the symmetric groups $S_n$ over an algebraically closed base field of characteristic 0. The main focus will be on the sets Irr $S_n$ of all irreducible representations (up to isomorphism) of the various $S_n$. We will follow Okounkov and Vershik [161], [200] rather than taking the classical approach invented by the originators of the theory, Frobenius, Schur and especially Young [207], about a century earlier. A remarkable feature of the Okounkov-Vershik method is that the entire chain of groups

(4.1) \[ 1 = S_1 \leq S_2 \leq \ldots \leq S_{n-1} \leq S_n \leq \ldots \]

is considered simultaneously and relations between the irreducible representations of successive groups in this chain are systematically exploited from the outset. Here, $S_{n-1}$ is identified with the subgroup of $S_n$ consisting of all permutations of $[n] = \{1, 2, \ldots, n\}$ that fix $n$.

The starting point for our investigation of the sets Irr $S_n$ is the Multiplicity-Freeness Theorem (Section 4.2), which states that the restriction of any $V \in \text{Irr } S_n$ to $S_{n-1}$ decomposes into a direct sum of pairwise non-equivalent irreducible representations. This fact naturally leads to the definition of the so-called branching graph $\mathbb{B}$: the set of vertices is the disjoint union of all Irr $S_n$ and we draw an arrow from $W \in \text{Irr } S_{n-1}$ to $V \in \text{Irr } S_n$ if $W$ is an irreducible constituent of $V \downarrow_{S_{n-1}}$. Another fundamental result in the representation theory of the symmetric groups, the Graph Isomorphism Theorem, establishes an isomorphism between $\mathbb{B}$ and a more elementary graph, the so-called Young graph $\mathbb{Y}$. The vertex set of $\mathbb{Y}$ is the disjoint union of the sets $\mathcal{P}_n$ consisting of all partitions of $n$, each represented by its Young diagram. An arrow from $\mu \in \mathcal{P}_{n-1}$ to $\lambda \in \mathcal{P}_n$ signifies that the diagram of $\lambda$ is obtained by adding a box to the diagram of $\mu$. 

185
The Graph Isomorphism Theorem makes combinatorial tools available for the investigation of $\text{Irr} \, S_n$. In particular, we will present the elegant “hook-walk” proof, due to Greene, Nijenhuis and Wilf [91], for the hook-length formula giving the degrees of the irreducible representations of $S_n$. The proof of the isomorphism $\mathcal{B} \rightarrow \mathcal{Y}$ uses an analysis of the spectrum of the so-called Gelfand-Zetlin algebra $GZ_n$, a commutative subalgebra of $kS_n$ whose definition is based directly on the chain (4.1). It turns out that each $V \in \text{Irr} \, S_n$ has a basis consisting of eigenvectors for $GZ_n$ and this basis is unique up to rescaling. For a suitable choice of scalars, the matrices of all operators $s_V$ ($s \in S_n$) will be shown to have rational entries; another normalization leads to orthogonal matrices (Young’s orthogonal form). Finally, we will present an efficient method for computing the irreducible characters of $S_n$, the Murnaghan-Nakayama Rule.

The general strategy employed in this chapter emulates traditional methods from the representation theory of semisimple Lie algebras, which will be the subject of Chapters 6 - 8. The role of the GZ-algebra $GZ_n$ is analogous to the one played by the Cartan subalgebra of a semisimple Lie algebra $\mathfrak{g}$ and analogs of $\mathfrak{sl}_2$-triples of $\mathfrak{g}$ will also occur in the present chapter. More extensive accounts of the representation theory of the symmetric groups along the lines followed here can be found in the original papers of Vershik and Okounkov [161], [200], [199], [201] and in the monographs [124] and [41].

*If not explicitly stated otherwise, the base field $k$ is understood to be algebraically closed of characteristic 0 throughout this chapter. Therefore, the group algebra $kS_n$ is split semisimple by Maschke’s Theorem (§3.4.1). We will often suppress $k$ in our notation, writing $\text{Irr} \, S_n = \text{Irr} \, kS_n$ as above and also $\text{Hom}_S = \text{Hom}_kS_n$, $\text{dim} = \text{dim}_k$ etc.*
4.1. Gelfand-Zetlin Algebras

The chain (4.1) gives rise to a corresponding chain of group algebras,
\begin{equation}
\mathbb{k} = \mathbb{k}S_1 \subseteq \mathbb{k}S_2 \subseteq \ldots \subseteq \mathbb{k}S_{n-1} \subseteq \mathbb{k}S_n \subseteq \ldots
\end{equation}

The **Gelfand-Zetlin (GZ)** algebra \(\mathcal{GZ}_n\) [82], [83] is defined to be the subalgebra of \(\mathbb{k}S_n\) that is generated by the centers \(\mathcal{Z}_k := \mathcal{Z}(\mathbb{k}S_k)\) for \(k \leq n\):
\[
\mathcal{GZ}_n \overset{\text{def}}{=} \mathbb{k}[\mathcal{Z}_1, \mathcal{Z}_2, \ldots, \mathcal{Z}_n] \subseteq \mathbb{k}S_n
\]

Note that all \(\mathcal{Z}_k\) commute elementwise with each other: if \(\alpha \in \mathcal{Z}_k\) and \(\beta \in \mathcal{Z}_l\) with \(k \leq l\), say, then \(\alpha \in \mathbb{k}S_k \subseteq \mathbb{k}S_l\) and \(\beta \in \mathcal{Z}(\mathbb{k}S_l)\) and hence \(\alpha \beta = \beta \alpha\). Therefore, \(\mathcal{GZ}_n\) is certainly commutative; in fact, the same argument will work for any chain of algebras in place of (4.2). In order to derive more interesting facts about \(\mathcal{GZ}_n\), we will need to use additional properties of (4.2). For example, we will show that \(\mathcal{GZ}_n\) is a maximal commutative subalgebra of \(\mathbb{k}S_n\) and that \(\mathcal{GZ}_n\) is semisimple; see Theorem 4.4 below.

4.1.1. Centralizer Subalgebras

Our first goal is to exhibit a more economical set of generators for the algebra \(\mathcal{GZ}_n\). This will be provided by the so-called **Jucys-Murphy (JM) elements**, which will play an important role throughout this chapter. The \(n^{\text{th}}\) JM-element, denoted by \(X_n\), is defined as the orbit sum of the transposition \((1, n) \in S_n\) under the conjugation action \(S_{n-1} \subseteq \mathbb{k}S_n\):
\[
X_n \overset{\text{def}}{=} \sum_{i=1}^{n-1} (i, n) \in (\mathbb{k}S_n)^{S_{n-1}}
\]

Here, \((\mathbb{k}S_n)^{S_{n-1}}\) denotes the subalgebra of \(\mathbb{k}S_n\) consisting of all \(S_{n-1}\)-invariants. Evidently, \((\mathbb{k}S_n)^{S_{n-1}}\) is contained in the invariant algebra \((\mathbb{k}S_n)^{S_k}\) of the conjugation action \(S_k \subseteq \mathbb{k}S_n\) for all \(k < n\), and \((\mathbb{k}S_n)^{S_k}\) can also be described as the centralizer of the subalgebra \(\mathbb{k}S_k\) in \(\mathbb{k}S_n\):
\[
(\mathbb{k}S_n)^{S_k} = \{a \in \mathbb{k}S_n \mid ab = ba \text{ for all } b \in \mathbb{k}S_k\}.
\]

By the foregoing, the JM-elements \(X_{k+1}, \ldots, X_n\) all belong to \((\mathbb{k}S_n)^{S_k}\), and this algebra clearly also contains the center \(\mathcal{Z}_k = \mathcal{Z}(\mathbb{k}S_k)\) as well as the subgroup \(S'_{n-k} \leq S_n\) consisting of the permutations of \([n] = \{1, 2, \ldots, n\}\) that fix all elements of \([k]\). The following theorem is due to Olshanskiĭ [162].

**Theorem 4.1.** The \(\mathbb{k}\)-algebra \((\mathbb{k}S_n)^{S_k}\) is generated by the center \(\mathcal{Z}_k = \mathcal{Z}(\mathbb{k}S_k)\), the subgroup \(S'_{n-k} \leq S_n\), and the JM-elements \(X_{k+1}, \ldots, X_n\).
Note that, for \( m \geq k + 1 \),
\[
(4.3) \quad X_m - (k + 1, m) - \cdots - (m - 1, m) = (k + 1, m)X_{k+1}(k + 1, m).
\]
Since \((i, m) \in S'_{n-k} \) for \( k < i < m \), all but one of the JM-elements could be deleted from the list of generators in the theorem. However, our focus later on will be on the JM-elements rather than the other generators.

Before diving into the proof of Olshanski˘ı’s Theorem, we remark that the \( S_k \)-conjugacy class of any \( s \in S_n \) can be thought of in terms of “marked cycle shapes.” In detail, if \( s \) is given as a product of disjoint cycles, possibly including 1-cycles, then we can represent \( S_k \) by the shape that is obtained by keeping each of \( k + 1, \ldots, n \) in its position in the given product while placing the symbol \( * \) in all other positions. For example, the marked cycle shape
\[
(*, *, *)(*, *)(12, *, 15)(13, *, *)(14)
\]
represents the \( S_{11} \)-conjugacy class consisting of all permutations of \([15]\) that are obtained by filling the positions marked \( * \) by the elements of \([11] \) in some order.

**Proof of Theorem 4.1.** We already know that \( \mathcal{A} := \mathbb{k}[\mathcal{Z}_k, S'_{n-k}, X_{k+1}, \ldots, X_n] \subseteq \mathcal{B} := (\mathbb{k}S_n)^{S_k} \). In order to prove the inclusion \( \mathcal{B} \subseteq \mathcal{A} \), observe that \( \mathbb{k}S_n \) is a permutation representation of \( S_k \). Hence a \( \mathbb{k} \)-basis of \( \mathcal{B} \) is given by the \( S_k \)-orbit sums (§3.3.1),
\[
\sigma_s := \sum_{t \in S_k} t \quad (s \in S_n),
\]
where \( S_k \) denotes the \( S_k \)-conjugacy class of \( s \). Our goal is to show that \( \sigma_s \in \mathcal{A} \) for all \( s \in S_n \).

To this end, we use a temporary notion of length\(^1\) for elements \( s \in S_n \), defining \( l(s) \) to be the number of points from \([n]\) that are moved by \( s \) or, equivalently, the number of symbols occurring in the disjoint cycle decomposition of \( s \) with all 1-cycles omitted. Clearly, \( l(\_\_\_\_) \) is constant on conjugacy classes of \( S_n \). Moreover, \( l(ss') \leq l(s) + l(s') \) for \( s, s' \in S_n \) and equality holds exactly if \( s \) and \( s' \) do not move a common point. Letting \( \mathcal{F}_l \subseteq \mathbb{k}S_n \) denote the \( \mathbb{k} \)-linear span of all \( s \in S_n \) with \( l(s) \leq l \), we have \( \mathcal{F}_0 = \mathbb{k} \subseteq \cdots \subseteq \mathcal{F}_{l-1} \subseteq \mathcal{F}_l \subseteq \cdots \subseteq \mathcal{F}_n = \mathbb{k}S_n \) and \( \mathcal{F}_l \mathcal{F}_l' \subseteq \mathcal{F}_{l+l'} \). Moreover, all subspaces \( \mathcal{F}_l \) are \( S_k \)-stable. Put
\[
\mathcal{B}_l = \mathcal{B} \cap \mathcal{F}_l = \mathcal{F}_l^{S_k}.
\]
A basis of \( \mathcal{B}_l \) is given by the orbit sums \( \sigma_s \) with \( l(s) \leq l \). We will show by induction on \( l \) that \( \mathcal{B}_l \subseteq \mathcal{A} \) or, equivalently, \( \sigma_s \in \mathcal{A} \) for all \( s \in S_n \) with \( l(s) = l \).

To start, if \( l = l(s) \leq 1 \) then \( \sigma_s = s = 1 \in \mathcal{A} \). More generally, if \( s = rt \) with \( r \in S_k \) and \( t \in S'_{n-k} \), then \( \sigma_s = \sigma_r t \) and \( \sigma_r \in \mathcal{F}_k \). Hence, \( \sigma_s \in \mathcal{A} \) again. Thus, we may assume that \( s \notin S_k \times S'_{n-k} \); so the disjoint cycle decomposition of \( s \) involves a cycle of the form \((\_\_\_\_, i, m)\) with \( i \leq k < m \). If \( l = 2 \) then \( s = (i, m) \) and \( \sigma_s \) is

---

\(^1\)This is not to be confused with another notion of “length,” considered in Example 7.10.
4.1. Gelfand-Zetlin Algebras

identical to the left-hand side of (4.3), which belongs to \( \mathcal{A} \). Therefore, we may assume that \( l > 2 \) and that \( B_{l-1} \subseteq \mathcal{A} \).

Next, assume that \( s = rt \) with \( r, t \neq 1 \) and \( l(r) + l(t) = l \). Then \( \sigma_r, \sigma_t \in \mathcal{A} \) by induction, and hence \( \mathcal{A} \ni \sigma_r \sigma_t = \sum r', i'' \ r' t'' \), where \( r' \) and \( t'' \) run over the \( S_k \)-conjugacy classes of \( r \) and \( t \), respectively. If \( r' \) and \( t'' \) move a common point from \([n]\), then \( l(r' t'') < l(r') + l(t'') = l \). The sum of all these products \( r' t'' \) is an \( S_k \)-invariant belonging to \( B_{l-1} \subseteq \mathcal{A} \). Therefore, it suffices to consider the sum of the non-overlapping products \( r' t'' \). Each of these products has the same marked cycle shape as \( s \), and hence it belongs to the \( S_k \)-conjugacy class of \( s \). By \( S_k \)-invariance, the sum of the non-overlapping products \( r' t'' \) is a positive integer multiple of \( \sigma_s \). This shows that \( \sigma_r \sigma_t \equiv z \sigma_s \mod \mathcal{A} \) for some \( z \in \mathbb{N} \), and we conclude that \( \sigma_s \in \mathcal{A} \).

It remains to treat the case where \( s \) is a cycle, say \( s = (j_1, \ldots, j_{l-2}, i, m) \) with \( i \leq k < m \). Write \( s = rt \) with \( r = (i, m) \) and \( t = (j_1, \ldots, j_{l-2}, m) \). Since \( \sigma_r, \sigma_t \in \mathcal{A} \) by induction, we once again have \( \mathcal{A} \ni \sigma_r \sigma_t = \sum i', i'' \ (i', m)' \) with \( i' \leq k \) and \( i'' \) running over the \( S_k \)-conjugacy class of \( t \). As above, the sum of all these products \( (i', m)' \) having length less than \( l \) belongs to \( \mathcal{A} \) by induction. The products of length equal to \( l \) all have the form \( (i', m)' = (i', m)(j_1', \ldots, j_{l-2}', m) = (j_1', \ldots, j_{l-2}', i', m) \), and these products form the \( S_k \)-conjugacy class of \( s \). Therefore, we again have \( \sigma_r \sigma_t \equiv z \sigma_s \mod \mathcal{A} \) for some \( z \in \mathbb{N} \), which finishes the proof. \( \Box \)

4.1.2. Generators of the Gelfand-Zetlin Algebra

As a consequence of Theorem 4.1, we obtain the promised generating set for the Gelfand-Zetlin algebra: \( \mathcal{GZ}_n \) is generated by the JM-elements \( X_k \) with \( k \leq n \). Even though \( X_1 = 0 \) is of course not needed as a generator, it will be convenient to keep this element in the list.

**Corollary 4.2.** \( \mathcal{GZ}_n = \mathbb{k}[X_1, X_2, \ldots, X_n] \).

**Proof.** First note that
\[
X_k = \sum \{ \text{all transpositions of } S_k \} - \sum \{ \text{all transpositions of } S_{k-1} \}.
\]
Since the first sum belongs to \( \mathcal{X}_k \) and the second to \( \mathcal{X}_{k-1} \), it follows that \( X_k \in \mathcal{GZ}_n = \mathbb{k}[\mathcal{X}_1, \ldots, \mathcal{X}_n] \). For the inclusion \( \mathcal{GZ}_n \subseteq \mathbb{k}[X_1, \ldots, X_n] \), we proceed by induction on \( n \). The case of \( \mathcal{GZ}_1 = \mathbb{k} \) being clear, assume that \( n > 1 \) and that \( \mathcal{GZ}_{n-1} \subseteq \mathbb{k}[X_1, \ldots, X_{n-1}] \). Since \( \mathcal{GZ}_n = \mathbb{k}[\mathcal{GZ}_{n-1}, \mathcal{X}_n] \) by definition, it suffices to show that \( \mathcal{X}_n \subseteq \mathbb{k}[\mathcal{GZ}_{n-1}, X_n] \). But
\[
\mathcal{X}_n = (\mathbb{k}S_n)^S_n \subseteq (\mathbb{k}S_n)^{S_{n-1}} = \mathbb{k}[\mathcal{X}_{n-1}, X_n] \subseteq \mathbb{k}[\mathcal{GZ}_{n-1}, X_n]
\]
as desired. \( \Box \)
Exercises for Section 4.1

4.1.1 (Relations between JM-elements and Coxeter generators). The transpositions $s_i = (i, i + 1)$ ($i = 1, \ldots, n - 1$) are called the Coxeter generators of $S_n$. Show that the following relations hold for the Coxeter generators and the JM-elements:

$s_i X_{i+1} = X_{i+1} s_i$ and $s_i X_j = X_j s_i$ if $j \neq i, i + 1$

4.1.2 (Product of the JM-elements). Show that $X_2 X_3 \ldots X_n$ is the sum of all $n$-cycles in $S_n$.

4.1.3 (Semisimplicity of some subalgebras of $kS_n$). Show that any subalgebra of $kS_n$ that is generated by some of the JM-elements $X_i$ ($i \leq n$) and some subgroups of $S_n$ and the centers of some subgroup algebras of $kS_n$ is semisimple. In particular, the centralizer algebras $(kS_n)^{S_k}$ and $GZ_n$ are semisimple. (Use Exercise 3.4.2 and the fact that all these subalgebras are stable under the standard involution of $kS_n$ and defined over $\mathbb{Q}$.)

4.2. The Branching Graph

In this section, we define the first of two directed graphs that will play a major role in this chapter: the branching graph $B$. This graph efficiently encodes a great deal of information concerning the irreducible representations of the various symmetric groups $S_n$.

4.2.1. Restricting Irreducible Representations

The developments in this section hinge on the following observation.

Multiplicity-Freeness Theorem. For each $V \in \text{Irr } S_n$, the restriction $V \downarrow_{S_{n-1}}$ is a direct sum of non-isomorphic irreducible representations of $S_{n-1}$.

Proof. Since $kS_{n-1}$ is split semisimple, we have $V \downarrow_{S_{n-1}} \cong \bigoplus_{W \in \text{Irr } S_{n-1}} W^{\oplus m(W)}$ with $m(W) \in \mathbb{Z}_+$ and so $\text{End}_{S_{n-1}}(V) \cong \prod_W \text{Mat}_{m(W)}(k)$ (Proposition 1.33). The theorem states that $m(W) \leq 1$ for all $W$, which is equivalent to the assertion that the algebra $\text{End}_{S_{n-1}}(V)$ is commutative. Similarly, since $kS_n$ is split semisimple, we have the standard isomorphism (1.46) of $k$-algebras,

$$
\begin{align*}
\begin{array}{ccc}
kS_n & \sim & \prod_{\text{Irr } S_n} \text{End}_k(V) \\
\text{w} & \downarrow_{\text{Irr } S_n} & \w
\end{array}
\end{align*}
$$

Under this isomorphism, the conjugation action $S_n \subseteq kS_n$ translates into the standard $S_n$-action on each component $\text{End}_k(V)$: $(^s a)_V = s_V \circ a_V \circ s_V^{-1} = s a_V$.
Therefore, the isomorphism (4.4) restricts to an isomorphism of algebras of $S_{n-1}$-invariants,

$$(kS_n)^{S_{n-1}} \cong \prod_{V \in \text{Irr} S_n} \text{End}_k(V)^{S_{n-1}} \cong \prod_{V \in \text{Irr} S_n} \text{End}_{S_{n-1}}(V).$$

By Theorem 4.1, $(kS_n)^{S_{n-1}} = k[\mathbb{F}_{n-1}, X_n]$ is a commutative algebra. Consequently, all $\text{End}_{S_{n-1}}(V)$ are commutative as well, as desired. □

4.2.2. The Graph $\mathbb{B}$

Consider the following graph $\mathbb{B}$, called the \textit{branching graph} of the chain (4.2). The set of vertices of $\mathbb{B}$ is defined by

$$\text{vert} \mathbb{B} = \bigsqcup_{n \geq 1} \text{Irr } S_n.$$

For given vertices $W \in \text{Irr } S_{n-1}$ and $V \in \text{Irr } S_n$, the graph $\mathbb{B}$ has a directed edge $W \to V$ if and only if there is a nonzero map $W \to V\downarrow_{S_{n-1}}$ in $\text{Rep } S_{n-1}$, that is, $W$ is an irreducible constituent of $V\downarrow_{S_{n-1}}$. Thus, the vertices of $\mathbb{B}$ are organized into levels, with $\text{Irr } S_n$ being the set of level-$n$ vertices, and all arrows in $\mathbb{B}$ are directed toward the next higher level. Figure 4.1 shows the first five levels $\mathbb{B}$ (Exercise 4.2.1).

![Branching Graph](image)

\textbf{Figure 4.1.} Bottom of the branching graph $\mathbb{B}$ (notation of §3.5.2)

The Multiplicity-Freeness Theorem can now be stated as follows:

$$(4.5) \quad V\downarrow_{S_{n-1}} \cong \bigoplus_{W \to V \in \mathbb{B}} W$$
Note that the decomposition (4.5) is canonical: the image of $W$ in $V_{\downarrow S_{n-1}}$ is uniquely determined as the $W$-homogeneous component of $V_{\downarrow S_{n-1}}$ and the map $W \to V_{\downarrow S_{n-1}}$ is a monomorphism in $\text{Rep}_{S_{n-1}}$ that is uniquely determined up to a scalar multiple by Schur’s Lemma.

4.2.3. Gelfand-Zetlin Bases

Let $V \in \text{Irr}_{S_n}$ be given. The following procedure yields a canonical decomposition of $V$ into 1-dimensional subspaces. Start by decomposing $V_{\downarrow S_{n-1}}$ into irreducible constituents as in (4.5); then, for each arrow $W \to V$ in $B$, decompose $W_{\downarrow S_{n-2}}$ into irreducibles. Proceeding in this manner all the way down to $S_1 = 1$, we obtain the desired decomposition of the vector space $V_{\downarrow S_1}$ into 1-dimensional subspaces, one for each path $T: \mathbb{I}_{S_1} \to \cdots \to V$ in $B$. The resulting decomposition of $V$ is uniquely determined, because the various decompositions at each step are unique. Choosing $0 \neq v_T$ in the subspace of $V$ corresponding to the path $T$, we obtain a basis $(v_T)$ of $V$ and each $v_T$ is determined up to a scalar multiple. This basis is called “the” Gelfand-Zetlin (GZ) basis of $V$; of course, any rescaling of this basis would also be a GZ-basis. To summarize,

\begin{equation}
V_{\downarrow S_1} = \bigoplus_{T: \mathbb{I}_{S_1} \to \cdots \to V \text{ in } B} \mathbb{K} v_T
\end{equation}

It follows in particular that

\begin{equation}
\dim V = \#\{\text{paths } \mathbb{I}_{S_1} \to \cdots \to V \text{ in } B\}.
\end{equation}

The reader is invited to check that there are five such paths in Figure 4.1 for $V = W$, and six paths for $V = \wedge^2 V_4$. This does of course agree with what we know already about the dimensions of these representations (§3.5.2). For a generalization of (4.7), see Exercise 4.2.3.

**Example 4.3** (GZ-basis of the standard representation $V_{n-1}$). Consider the chain $M_1 \subseteq M_2 \subseteq \cdots \subseteq M_n \subseteq \cdots$, where $M_n = \bigoplus_{i=1}^{n} \mathbb{K} b_i$ is the standard permutation representation of $S_n$ (§3.2.4). Working inside $\bigcup_{n \geq 1} M_n = \bigoplus_{i \geq 1} \mathbb{K} b_i$, we have $V_{n-1} = \{\sum_i \lambda_i b_i \mid \sum_i \lambda_i = 0 \text{ and } \lambda_i = 0 \text{ for } i > n\}$ and so $\cdots \subseteq V_{n-2} \subseteq V_{n-1} \subseteq \cdots$. Thus, $V_{n-2}$ provides us with an irreducible component of $V_{n-1\downarrow S_{n-1}}$. The vector

$$v_{n-1} = \sum_{i=1}^{n-1} (b_i - b_n) = \sum_{i=1}^{n-1} b_i - (n-1)b_n \in V_{n-1}$$

is a nonzero $S_{n-1}$-invariant that does not belong to $V_{n-2}$. For dimension reasons, we conclude that $V_{n-1\downarrow S_{n-1}} = \mathbb{K} v_{n-1} \oplus V_{n-2} \cong \mathbb{I}_{S_{n-1}} \oplus V_{n-2}$ is the decomposition of $V_{n-1\downarrow S_{n-1}}$ into irreducible constituents. Inductively we further deduce that $V_{n-1} = \cdots$
The Branching Graph

193

\[ \bigoplus_{j=1}^{n-1} k v_j \] and that \((v_1, \ldots, v_{n-1})\) is the GZ-basis of \(V_{n-1}\). It is straightforward to check that the Coxeter generator \(s_i = (i, i+1) \in S_n\) acts on this basis as follows:

\[
s_i . v_j = \begin{cases} v_j & \text{for } j \neq i-1, i \\ \frac{1}{i} v_{i-1} + (1 - \frac{1}{i}) v_i & \text{for } j = i-1 \\ (1 + \frac{1}{i}) v_{i-1} - \frac{1}{i} v_i & \text{for } j = i \end{cases}
\]

(4.8)

These equations determine the vectors \(v_j \in V_{n-1}\) up to a common scalar factor: if (4.8) also holds with \(w_j\) in place of \(v_j\), then \(v_j \mapsto w_j\) is an \(S_n\)-equivariant endomorphism of \(V_{n-1}\) and hence an element of \(D(V_{n-1}) = k\). We shall discuss some rescalings of the GZ-basis \((v_j)\) in Examples 4.17 and 4.19.

4.2.4. Properties of \(GZ_n\)

We have seen that the Gelfand-Zetlin algebra \(GZ_n\) is commutative and generated by the JM-elements \(X_1, \ldots, X_n\). Now we derive further information about \(GZ_n\) from the foregoing.

Theorem 4.4. (a) \(GZ_n\) is the set of all \(a \in k S_n\) such that the GZ-basis of each \(V \in \text{Irr } S_n\) consists of eigenvectors for \(a V\).

(b) \(GZ_n\) is a maximal commutative subalgebra of \(k S_n\).

(c) \(GZ_n\) is semisimple: \(GZ_n \cong k^{d_n}\) with \(d_n = \sum_{V \in \text{Irr } S_n} \text{dim } V\).

Proof. For each \(V \in \text{Irr } k S_n\), identify \(\text{End}_k(V)\) with the matrix algebra \(\text{Mat}_{\text{dim } V}(k)\) via the GZ-basis of \(V\). Then the isomorphism (4.4) identifies the group algebra \(k S_n\) with the direct product of these matrix algebras. Let \(D\) denote the subalgebra of \(k S_n\) that corresponds to the direct product of the algebras of diagonal matrices in each component. Part (a) asserts that \(D = GZ_n\):

\[
\begin{align*}
\begin{array}{c}
k S_n \cong \prod_{V \in \text{Irr } S_n} \text{Mat}_{\text{dim } V}(k) \\
\text{via GZ-bases}
\end{array}
\end{align*}
\]

(4.9)

The isomorphism \(GZ_n \cong k^{d_n}\) in (c) is then clear and so is the maximality assertion in (b). Indeed, the subalgebra of diagonal matrices in any matrix algebra \(\text{Mat}_d(k)\) is self-centralizing: the only matrices that commute with all diagonal matrices are themselves diagonal. Therefore, \(D\) is a self-centralizing subalgebra of \(k S_n\), and hence \(D\) is a maximal commutative subalgebra. In particular, in order to prove the equality \(D = GZ_n\), it suffices to show that \(D \subseteq GZ_n\), because we already know that \(GZ_n\) is commutative.

To prove the inclusion \(D \subseteq GZ_n\), let \(e(V) \in Z_n\) denote the primitive central idempotent of \(k S_n\) corresponding to \(V \in \text{Irr } S_n\) (§1.4.4). Recall that,
for any $W \in \text{Rep} S_n$, the operator $e(V)_W$ projects $W$ onto the $V$-homogeneous component $W(V)$, annihilating all other homogeneous components of $W$. Therefore, for any path $T: 1_{S_i} = W_1 \rightarrow W_2 \rightarrow \cdots \rightarrow W_n = V$ in $\mathbb{B}$, the element $e(T) := e(W_1)e(W_2)\cdots e(W_n) \in \mathbb{Z}[X_1, X_2, \ldots, X_n] = \mathcal{G} Z_n$ acts on $V$ as the projection $\pi_T: V \rightarrow V_T = \mathbb{k}v_T$ in (4.6) and $e(T)v' = 0_{v'}$ for all $V \neq V' \in \text{Irr} S_n$. Thus, in (4.9), we have

$$\text{k} S_n \sim \prod_{V \in \text{Irr} S_n} \text{Mat}_{\dim V}(\mathbb{k})$$

This shows that the idempotents $e(T)$ form the standard basis of the diagonal algebra $\mathcal{D}$, consisting of the diagonal matrices with one entry equal to 1 and all others 0, which proves the desired inclusion $\mathcal{D} \subseteq \mathcal{G} Z_n$.  

\[ \square \]

4.2.5. The Spectrum of $\mathcal{G} Z_n$

We now give a description of $\text{Spec} \mathcal{G} Z_n = \text{MaxSpec} \mathcal{G} Z_n \cong \text{Hom}_{\mathbb{Z}[x]}(\mathcal{G} Z_n, \mathbb{k})$ (§1.3.2) that will play an important role in Section 4.4. For this, we elaborate on some of the properties of $\mathcal{G} Z_n$ stated in Theorem 4.4. First, the fact that the $GZ$-basis $(v_T)$ of any $V \in \text{Irr} S_n$ consists of eigenvectors for $\mathcal{G} Z_n$ says that each $v_T$ is a weight vector for a suitable weight $\phi_T \in \text{Hom}_{\mathbb{Z}[x]}(\mathcal{G} Z_n, \mathbb{k})$:

\[ a.v_T = \phi_T(a)v_T \quad (a \in \mathcal{G} Z_n). \]

Moreover, in view of (4.7), the dimension $d_n = \dim \mathcal{G} Z_n$ is equal to the total number of paths in $\mathbb{B}$ from $1_{S_i}$ to some vertex $\in \text{Irr} S_n$. Finally, the isomorphism $\mathcal{G} Z_n \cong \mathbb{k}^{d_n}$ in (4.9) is given by $a \mapsto (\phi_T(a))_T$. Therefore, each $\phi_T$ is a weight of a unique $V \in \text{Irr} S_n$, the endpoint of the path $T: 1_{S_i} \rightarrow \cdots \rightarrow V$ in (4.6), and

\[ \text{Hom}_{\mathbb{Z}[x]}(\mathcal{G} Z_n, \mathbb{k}) = \left\{ \phi_T \mid T \text{ a path $1_{S_i} \rightarrow \ldots \rightarrow$ in } \mathbb{B} \right\} \]

Since the algebra $\mathcal{G} Z_n$ is generated by the JM-elements $X_1, \ldots, X_n$ (Corollary 4.2), each weight $\phi_T$ is determined by the $n$-tuple $(\phi_T(X_i))^n_1 \in \mathbb{k}^n$. Therefore, the spectrum $\text{Hom}_{\mathbb{Z}[x]}(\mathcal{G} Z_n, \mathbb{k})$ of $\mathcal{G} Z_n$ is in one-to-one correspondence with the set

\[ \text{Spec}(n) \overset{\text{def}}{=} \left\{ (\phi(X_1), \phi(X_2), \ldots, \phi(X_n)) \in \mathbb{k}^n \mid \phi \in \text{Hom}_{\mathbb{Z}[x]}(\mathcal{G} Z_n, \mathbb{k}) \right\} \]

To summarize, we have bijections

\[ \text{Spec}(n) \overset{\sim}{\leftrightarrow} \text{Hom}_{\mathbb{Z}[x]}(\mathcal{G} Z_n, \mathbb{k}) \overset{\sim}{\leftrightarrow} \left\{ \text{paths } 1_{S_i} \rightarrow \ldots \rightarrow \text{ in } \mathbb{B} \right\} \]

\[ \phi_T(X_i)^n \leftrightarrow \phi_T \leftrightarrow T \]
4.3. The Young Graph

We now start afresh, working in purely combinatorial rather than representation theoretic territory.

4.3.1. Partitions and Young Diagrams

The main player in this section is the following set of non-negative integer sequences:

\[ \mathcal{P} \overset{\text{def}}{=} \left\{ (\lambda_1, \lambda_2, \ldots) \in \mathbb{Z}_+^\infty \mid \lambda_1 \geq \lambda_2 \geq \ldots \text{ and } \sum \lambda_i < \infty \right\} \]

We will denote the sequence \((\lambda_1, \lambda_2, \ldots) \in \mathcal{P}\) by \(\lambda\) and write \(|\lambda| = \sum \lambda_i\). Thus,

\[ \mathcal{P} = \bigsqcup_n \mathcal{P}_n \quad \text{with} \quad \mathcal{P}_n \overset{\text{def}}{=} \left\{ \lambda \in \mathcal{P} \mid |\lambda| = n \right\}. \]

The members of \(\mathcal{P}_n\) are called \textit{partitions} of \(n\). For \(\lambda \in \mathcal{P}_n\), we will also use the standard notation \(\lambda + n\). Partitions will be visualized by \textit{Young diagrams}: the Young diagram of \(\lambda = (\lambda_1, \lambda_2, \ldots)\) consists of rows of boxes that are aligned on the left, with \(\lambda_1\) boxes in the first row, \(\lambda_2\) in the second, etc. The unique partition with \(|\lambda| = 0\) thus has an empty Young diagram. We will generally only consider partitions with \(|\lambda| \geq 1\) (except in Exercise 4.3.1) and we will typically

\[ \text{For more information on this sequence, see the On-Line Encyclopedia of Integer Sequences [186, Sequence A000085].} \]
write partitions as finite sequences, omitting a tail of 0-components. Here, for example, is the Young diagram of the partition \((7, 5, 4, 4, 2)\):

\[
\begin{array}{c c c c c c}
\cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \\
\cdot & \cdot & \cdot & \\
\cdot & \cdot & \\
\end{array}
\]

Young diagrams are also called Ferrers diagrams, particularly when represented using dots instead of boxes, and sometimes the rows are arranged in the reverse order, from the bottom up (French notation), but we will use the above convention.

Reflecting a given partition \(\lambda\) across the line \(y = x\) yields the so-called \textbf{conjugate partition}; it will be denoted by \(\lambda^c\). For example, \((7, 5, 4, 4, 2)^c = (5, 5, 4, 4, 2, 1, 1)\).

**4.3.2. The graph \(\mathcal{Y}\) and the Graph Isomorphism Theorem**

The \textbf{Young graph} \(\mathcal{Y}\) has vertex set

\[
\text{vert } \mathcal{Y} = \mathcal{P},
\]

with each \(\lambda \in \mathcal{P}\) represented by its Young diagram. An arrow \(\mu \rightarrow \lambda\) in \(\mathcal{Y}\) means that the Young diagram of \(\mu\) is obtained from the one of \(\lambda\) by removing one box, necessarily a “southeast” corner box. Note that the number of removable boxes of \(\lambda = (\lambda_i)_{i \geq 1}\) is equal to the number of distinct values among the \(\lambda_i\). We define a partial order \(\leq\) on \(\mathcal{P}\) by declaring \(\mu \leq \lambda\) if the Young diagram of \(\mu\) fits into the diagram of \(\lambda\) or, more formally, \(\mu_i \leq \lambda_i\) for all \(i\). Thus, there is an arrow \(\mu \rightarrow \lambda\) in \(\mathcal{Y}\) iff \(\mu < \lambda\) but no \(\nu \in \mathcal{P}\) satisfies \(\mu < \nu < \lambda\).

The vertices of \(\mathcal{Y}\) are divided into levels, with \(\mathcal{P}_n\) at level \(n\). The first five levels of \(\mathcal{Y}\), as displayed in Figure 4.2, show a striking similarity to the corresponding levels of the branching graph \(\mathcal{B}\) (Figure 4.1). In fact, we have the following fundamental

**Graph Isomorphism Theorem.** The graphs \(\mathcal{Y}\) and \(\mathcal{B}\) are isomorphic.

Explicitly, the theorem asserts the existence of a bijection \(\phi: \text{vert } \mathcal{Y} \xrightarrow{\sim} \text{vert } \mathcal{B}\) such that there is an arrow \(\mu \rightarrow \lambda\) in \(\mathcal{Y}\) if and only if there is an arrow \(\phi(\mu) \rightarrow \phi(\lambda)\) in \(\mathcal{B}\). We will then write \(\phi: \mathcal{Y} \xrightarrow{\sim} \mathcal{B}\). We may of course also speak of automorphisms \(\mathcal{B} \xrightarrow{\sim} \mathcal{B}\) and likewise for \(\mathcal{Y}\). In fact, it is not hard to see that conjugation, \(\lambda \mapsto \lambda^c\), is
the only non-identity automorphism of $\mathcal{Y}$ (Exercise 4.3.2). Thus, there are at most two possible graph isomorphisms $\mathcal{Y} \xrightarrow{\sim} \mathcal{B}$.

The proof of the Graph Isomorphism Theorem will be given in Section 4.4. In the remainder of this section, we will discuss some consequences. Throughout, let us write the bijection on vertices given by the graph isomorphism $\mathcal{Y} \xrightarrow{\sim} \mathcal{B}$ as

$$\lambda \mapsto V^\lambda \quad (\lambda \in \mathcal{P}).$$

For example, we clearly must have $V^\square = \mathbb{1}_{S_1}$, because $\square$ and $\mathbb{1}_{S_1}$ are the sole vertices of $\mathcal{Y}$ and $\mathcal{B}$ with no incoming arrows. More generally, any isomorphism $\mathcal{Y} \xrightarrow{\sim} \mathcal{B}$ must bijectively map the $n$-vertex paths $\square \rightarrow \mu_2 \rightarrow \cdots \rightarrow \mu_n$ in $\mathcal{Y}$ to the corresponding paths in $\mathcal{B}$, and hence it will give match the level-$n$ vertices of $\mathcal{Y}$ with the level-$n$ vertices of $\mathcal{B}$, giving bijections $\mathcal{P}_n \xrightarrow{\sim} \text{Irr}\, S_n$ for all $n$. Of course, we already know that such bijections exists: there are as many irreducible representations of $S_n$ as there are conjugacy classes of $S_n$ (Corollary 3.21) and the conjugacy classes in turn are in bijection with the partitions of $n$ (§3.5.2). However, the full form of the the Graph Isomorphism Theorem gives much more detailed information and its proof will require more work.

### 4.3.3. Consequences of the Graph Isomorphism Theorem

The Graph Isomorphism Theorem allows us to derive information about the irreducible representations of all $S_n$ from combinatorial features of $\mathcal{Y}$. Here are some examples.
The Branching Rule. Since $\mu \to \lambda$ in $\mathcal{Y}$ is equivalent to $V^\mu \to V^\lambda$ in $\mathcal{B}$, we may rewrite (4.5) as

$$V^\lambda \downarrow_{S_n} \equiv \bigoplus_{\mu \to \lambda \text{ in } \mathcal{Y}} V^\mu \tag{4.13}$$

This formula is referred to as a branching rule. The number of arrows $\mu \to \lambda$ in (4.13) is equal to the number of removable boxes of $\lambda$, which in turn is equal to the number of distinct values (row lengths in the Young diagram) of $\lambda$. Therefore,

$$\text{length } V^\lambda \downarrow_{S_n} = \# \{ \text{distinct values of } \lambda \}.$$

In particular, $V^\lambda \downarrow_{S_n}$ is irreducible if and only if the Young diagram of $\lambda$ is a rectangle. The branching rule (4.13) has the following reformulation:

$$V^\mu \uparrow_{S_n} \equiv \bigoplus_{\mu \to \lambda \text{ in } \mathcal{Y}} V^\lambda \tag{4.14}$$

Indeed, (4.13) says that the multiplicity of $V^\mu$ in $V^\lambda \downarrow_{S_n}$ is equal to 1 if there is an arrow $\mu \to \lambda$ in $\mathcal{Y}$ and equal to 0 otherwise; (4.14) makes the same statement about the multiplicity of $V^\lambda$ in $V^\mu \uparrow_{S_n}$. The equivalence of (4.13) and (4.14) thus follows from Frobenius reciprocity (Corollary 1.37): $m(V^\mu, V^\lambda \downarrow_{S_n}) = m(V^\lambda, V^\mu \uparrow_{S_n})$.

Dimension. Any graph isomorphism $\mathcal{Y} \xrightarrow{\simeq} \mathcal{B}$ will induce bijections between the set of all paths $\square \to \cdots \to \lambda$ in $\mathcal{Y}$ and the set of all paths $\mathbb{1}_{S_n} \to \cdots \to V^\lambda$ in $\mathcal{B}$. Since the size of the latter set of paths equals $\dim V^\lambda$ by (4.7), we obtain

$$\dim V^\lambda = f^\lambda \overset{\text{def}}{=} \# \{ \text{paths } \square \to \cdots \to \lambda \text{ in } \mathcal{Y} \} \tag{4.15}$$

The number $f^\lambda$ will be determined in §4.3.5 below. Note that the dimension $d_n = \dim G\mathcal{Z}_n$ (Theorem 4.4) can now also be written as $d_n = \sum_{\lambda \vdash n} f^\lambda$.

4.3.4. Paths in $\mathcal{Y}$ and Standard Young Tableaux

In this subsection, we will describe the number $f^\lambda$ defined in (4.15) in terms of standard Young tableaux. By definition, a standard Young tableau of shape $\lambda \vdash n$, or $\lambda$-tableau for short, is obtained by filling the numbers $1, 2, \ldots, n$ into the boxes of the Young diagram of $\lambda$ in such a way that the numbers increase along rows (left to right) and along columns (top to bottom). Clearly, the $(1, 1)$-box must contain the number 1 and $n$ must occur in some removable corner box. Removing this box, we obtain the Young diagram of a partition $\mu$ with $\mu \to \lambda$. Continuing in this manner,
successively removing the boxes containing the highest number, we eventually end up with the tableau \( \square \). This process is easily seen to yield a bijection
\[
\{ \text{paths } \square \to \cdots \to \lambda \text{ in } \mathcal{Y} \} \leftrightarrow \{ \lambda \text{-tableaux} \}.
\]

Just as we have identified partitions with their Young diagrams in the description of the Young graph \( \mathcal{Y} \), we will also oftentimes not distinguish between paths in \( \mathcal{Y} \) and standard Young tableaux. In particular, we may rewrite the definition of \( f^\lambda \) in (4.15) as follows:
\[
\text{(4.17) } f^\lambda = \# \{ \lambda \text{-tableaux} \}
\]

**Example 4.5.** Here are all standard Young tableaux of shape \( \lambda = (2, 2, 1) \) along with the corresponding paths in \( \mathcal{Y} \):

\[
\begin{align*}
1 & \Rightarrow \begin{array}{cccc}
1 & 2 & 3 & 4 \\
5 & 6 & 7 & 8
\end{array} \\
1 & 3 & 2 & 4 \\
5 & 6 & 7 & 8 \\
1 & 3 & 2 & 4 \\
5 & 6 & 7 & 8 \\
1 & 3 & 2 & 4 \\
5 & 6 & 7 & 8
\end{align*}
\]

### 4.3.5. The Hook-Length Formula

Identifying a given \( \lambda \vdash n \) with its Young diagram as usual, we consider the hook at a given box \( x \) of \( \lambda \) as in the picture on the right. The **hook length** at \( x \) is defined by
\[
h(x) \overset{\text{def}}{=} \# \{ \text{boxes in the hook at } x \}.
\]

Corner boxes or, equivalently, removable boxes are exactly those boxes, \( x \), with \( h(x) = 1 \). The following formula is due to Frame, Robinson and Thrall [73]; the probabilistic proof that we shall present below is due to Greene, Nijenhuis and Wilf [91].

**Hook-Length Formula.** For \( \lambda \vdash n \), we have
\[
f^\lambda = \frac{n!}{\prod_{x \in \lambda} h(x)}.
\]

For example, consider the partition \( \lambda = (3, 2, 2, 1) \). Filling each box in the Young diagram of \( \lambda \) with the length of the hook at this box, we obtain the scheme on the left. Hence, \( f^\lambda = \frac{8!}{6 \cdot 4 \cdot 3 \cdot 2 \cdot 1^3} = 70 \).
Here is some brief background on probability and the description of the particular experiment that is used in the proof of the hook-length formula.

The **Hook-Walk Experiment**. Suppose that a certain experiment has a finite set \( \Omega \) of possible outcomes; the set \( \Omega \) is called the **sample space** and subsets \( E \subseteq \Omega \) are referred to as **events**. Assume further that, for each \( \omega \in \Omega \), there is a probability value \( p(\omega) \in \mathbb{R}_{\geq 0} \) such that \( \sum_{\omega \in \Omega} p(\omega) = 1 \). Then the probability of a given event \( E \subseteq \Omega \) is defined by

\[
P(E) \overset{\text{def}}{=} \sum_{\omega \in E} p(\omega) .
\]

For example, if all outcomes \( \omega \in \Omega \) have the same probability, then

\[
P(E) = \frac{|E|}{|\Omega|}.
\]

The particular experiment that we will consider below is, in probability parlance, a memoryless random walk or Markov chain: each step in the walk depends only on the current position and not on the sequence of steps that preceded it. Specifically, consider a partition \( \lambda \vdash n \), identified with its Young diagram. To start with, choose a box, \( x = x_0 \), among the \( n \) boxes of \( \lambda \) with uniform probability \( \frac{1}{n} \). If \( x \) is a corner box, then stop; otherwise, choose a different box, \( x_1 \), in the hook at \( x \) with uniform probability \( q(x) := \frac{1}{h(x) - 1} \). If \( x_1 \) is a corner box, then stop; otherwise, choose a different box, \( x_2 \), in the hook at \( x_1 \) with uniform probability \( q(x_1) \) etc. Each step moves either down or right, and the walk will terminate at some corner box \( x_t = c \in \lambda \) after finitely many steps:

\[
\omega: x = x_0 \to x_1 \to \cdots \to x_{t-1} \to x_t = c .
\]

We will refer to \( \omega \) as a **hook walk** in \( \lambda \). Our sample set \( \Omega \) consists of all such hook walks; this is clearly a finite set. The probability of the hook walk \( \omega \) is given by

\[
p(\omega) = \frac{1}{n} q(\omega) \quad \text{with} \quad q(\omega) := q(x_0)q(x_1)\cdots q(x_{t-1}) .
\]

It is not hard to see that \( \sum_{\omega \in \Omega} p(\omega) = 1 \). For each corner box \( c \), consider the event \( E_c = \{ \text{all hook walks in } \lambda \text{ that end at } c \} \). These events form a partition of the sample set \( \Omega \). Therefore,

\[
1 = \sum_{\omega \in \Omega} p(\omega) = \sum_c P(E_c) = \sum_c \frac{1}{n} \sum_{\omega \in E_c} q(\omega) ,
\]

where \( c \) runs over the corner boxes of \( \lambda \).

**Proof of the Hook-Length Formula.** Our goal is to prove the formula \( f^A = \frac{n!}{H(\lambda)} \), where we have put

\[
H(\lambda) := \prod_{x \in A} h(x)
\]

We proceed by induction on \( n \). The formula is trivially true for \( n = 1 \). To deal with \( n > 1 \), we use the following recursion, which is evident from the definition of \( f^A \)
as the number of paths $\square \to \cdots \to \lambda$ in $\mathcal{Y}$:

$$f^\lambda = \sum_{\mu \to \lambda} f^\mu.$$  

By induction, we know that $f^\mu = (n-1)! H(\mu)$ for all $\mu$ in the above sum. Thus, we need to show that $\frac{n!}{H(\lambda)} = \sum_{\mu \to \lambda} \frac{(n-1)!}{H(\mu)}$. Recall that $\mu \to \lambda$ means that $\mu$ arises from $\lambda$ by removing one corner box, say $c$. Denoting the resulting $\mu$ by $\lambda \setminus c$, our goal is to show that

$$1 = \sum_{c} \frac{1}{n} \frac{H(\lambda)}{H(\lambda \setminus c)}.$$  

where $c$ runs over the corner boxes of $\lambda$. Comparison with (4.19) shows that it suffices to prove the following equality, for each corner box $c$:

(4.20)  

$$\sum_{\omega \in E_c} q(\omega) = \frac{H(\lambda)}{H(\lambda \setminus c)}.$$  

First, let us consider the right-hand side of (4.20). Note that $\lambda \setminus c$ has the same hooks as $\lambda$, except that the hook at $c$, of length 1, is missing and the hooks at all boxes in the gray region marked $B = B(c)$ on the right have lengths shorter by 1 than the corresponding hooks of $\lambda$. Therefore, the right hand side of (4.20) can be written as follows:

$$H(\lambda) H(\lambda \setminus c) = \prod_{b \in B} \frac{h(b)}{h(b) - 1}.$$  

Using the notation $q(b) = \frac{1}{h(b)}$ from the hook-walk experiment, we can write the product on the right as $\prod_{b \in B} (1 + q(b)) = \sum_{S \subseteq B} \prod_{b \in S} q(b)$. Hence,

(4.21)  

$$\frac{H(\lambda)}{H(\lambda \setminus c)} = \sum_{S \subseteq B} \prod_{b \in S} q(b).$$  

Now for the left-hand side of (4.20). For each hook walk $\omega \in E_c$, let $S_\omega \subseteq B$ denote the set of boxes that arise as the horizontal and vertical projections of the boxes of $\omega$ into $B$; see the green boxes on the left. Note that, while $S_\omega$ generally does not determine the entire walk $\omega$, the starting point, $x$, is certainly determined, and if $x \in B$, then $\omega$ is determined. We claim that, for each subset $S \subseteq B$,
This will give the following expression for the left-hand side of (4.20):

\[ \sum_{\omega \in E_c} q(\omega) = \sum_{S \subseteq B} \sum_{\omega \in E_c} q(\omega) = \sum_{S \subseteq B} \prod_{b \in S} q(b). \]

By (4.21) this is identical to the right-hand side of (4.20), thereby proving (4.20).

We still need to justify the claimed equality (4.22). For this, we argue by induction on \(|S|\). The only hook walk \(\omega \in E_c\) with \(S_\omega = \emptyset\) is the walk starting and ending at \(c\) without ever moving; so the claim is trivially true for \(|S| = 0\). The claim is also clear if the starting point, \(x\), of \(\omega\) belongs to \(B\), because then the sum on the left has only one term, which is equal to the right-hand side of (4.22). So assume that \(x \not\in B \cup \{c\}\). Then there are two kinds of possible hook walks \(\omega \in E_c\) with \(S_\omega = S\): those that start with a move to the right and those that start with a move down. Letting \(\eta\) denote the remainder of the walk and letting \(x', x'' \in B\) be the vertical and horizontal projections of \(x\) into \(B\), we have \(S_\eta = S \setminus \{x'\}\) in the former case and \(S_\eta = S \setminus \{x''\}\) in the latter. Therefore,

\[
\sum_{\omega \in E_c, S_\omega = S} q(\omega) = q(x) \left( \sum_{\eta \in E_c, S_\eta = S \setminus \{x'\}} q(\eta) + \sum_{\eta \in E_c, S_\eta = S \setminus \{x''\}} q(\eta) \right)
= q(x) \left( \prod_{b \in S \setminus \{x'\}} q(b) + \prod_{b \in S \setminus \{x''\}} q(b) \right)
= q(x) \left( \frac{1}{q(x')} + \frac{1}{q(x'')} \right) \prod_{b \in S} q(b).
\]

To complete the proof of (4.22), it remains to observe that \(q(x) \left( \frac{1}{q(x')} + \frac{1}{q(x'')} \right) = 1\) or, equivalently, \(h(x) + 1 = h(x') + h(x'')\), which is indeed the case.

This finishes the proof of the hook-length formula. \(\Box\)
Exercises for Section 4.3

4.3.1 (Up and down operators). Let $\mathbb{Z}\mathcal{P} = \bigoplus_n \mathbb{Z}\mathcal{P}_n$ denote the $\mathbb{Z}$-module of all formal $\mathbb{Z}$-linear combinations of partitions $\lambda$. Here, $\mathbb{Z}\mathcal{P}_n = 0$ for $n < 0$, because $\mathcal{P}_n$ is empty in this case, and $\mathbb{Z}\mathcal{P}_0 \cong \mathbb{Z}$, because $\mathcal{P}_0$ contains only $(0, 0, \ldots)$, with Young diagram $\varnothing$. Thus, we have added $\varnothing$ as a root vertex to $\mathcal{Y}$ and a unique arrow $\varnothing \rightarrow \Box$. Consider the operators $U, D \in \text{End}_\mathbb{Z}(\mathbb{Z}\mathcal{P})$ that are defined by $U(\lambda) = \sum_{\lambda \rightarrow \mu} \mu$ and $D(\lambda) = \sum_{\mu \rightarrow \lambda} \mu$. Show that these operators satisfy the Weyl algebra relation $DU = UD + 1$.

4.3.2 (Automorphisms of $\mathcal{Y}$). (a) Show that each $\lambda \in \mathcal{P}_n$ ($n > 2$) is determined by the set $S(\lambda) := \{\mu \in \mathcal{P}_{n-1} \mid \mu \rightarrow \lambda \text{ in } \mathcal{Y}\}$.

(b) Conclude by induction on $n$ that the graph $\mathcal{Y}$ has only two automorphisms: the identity and conjugation.

4.3.3 (Rectangle partitions). Show:

(a) $f^{(n,n)} = \frac{1}{n+1} \binom{2n}{n}$, the $n$th Catalan number.

(b) $f^A > rc$ for $\lambda = (c^r) := (c, \ldots, c)$ with $c, r \geq 2$ and $rc \geq 8$.

4.3.4 (Dimensions of the irreducible representations of $S_6$). Extend Figure 4.2 up to layer $\mathcal{P}_6$ and find the degrees of all irreducible representations of $S_6$ (assuming the Graph Isomorphism Theorem).

4.3.5 (Hook partitions and exterior powers of the standard representation). Let $\lambda \mapsto V^A$ be the bijection on vertices that is given by a graph isomorphism $\mathcal{Y} \cong \mathcal{B}$ as in §4.3.2. Assume that the isomorphism has been chosen so that $\Box \mapsto 1_{S_6}$. Show that $V^{(n-k,1^k)} = \wedge^k V_{n-1}$ holds for all $n$ and $k = 0, \ldots, n-1$, where $(n-k, 1^k)$ is the “hook partition” $(n-k, 1, \ldots, 1)$.

4.3.6 (Irreducible representations of dimension < $n$). Let $n \geq 7$. Assuming the Graph Isomorphism Theorem, show that $1$, sgn, the standard representation $V_{n-1}$, and its sign twist $V_{n-1}^\pm$ are the only irreducible representations of degree < $n$ of $S_n$. (Use Exercise 4.3.3(b).)

4.4. Proof of the Graph Isomorphism Theorem

The main goal of this section is to provide the proof of the Graph Isomorphism Theorem. This will be accomplished in Corollary 4.13 after some technical tools have been deployed earlier in this section. In short, the strategy is to set up a bijection between the collection of all paths $\Box \rightarrow \ldots$ in $\mathcal{Y}$ with endpoint in $\mathcal{P}_n$ and the seemingly rather more complex collection of paths $1_{S_n} \rightarrow \ldots$ in $\mathcal{B}$ with endpoint in $\text{Irr } S_n$. We have already constructed a bijection between the latter paths
and a certain subset $\text{Spec}(n) \subseteq \mathbb{X}^n$ in (4.12) and we have also identified paths in $\mathbb{Y}$ with standard Young tableaux in (4.16). We will show in §4.4.1 that standard Young tableaux with $n$ boxes in turn are in one-to-one correspondence with a certain set of $n$-tuples $\text{Cont}(n) \subseteq \mathbb{Z}^n$. The main point of the proof is to show that $\text{Cont}(n) = \text{Spec}(n)$; this will yield the desired bijection

\[
\begin{array}{c}
\{ \text{paths } \square \to \ldots \text{ in } \mathbb{Y} \} \xrightarrow{\sim} \text{Cont}(n) = \text{Spec}(n) \xleftarrow{\sim} \{ \text{paths } 1_S \to \ldots \text{ in } \mathbb{B} \}.
\end{array}
\]

4.4.1. Contents

Let $T$ be a standard Young tableau with $n$ boxes. We define the **content** of $T$ to be the $n$-tuple $c_T = (c_{T,1}, c_{T,2}, \ldots, c_{T,n})$, where $c_{T,i} = a$ means that the box of $T$ containing the number $i$ lies on the line $y = x + a$; we will call this line the $a$-diagonal. Thus,

\[
c_{T,i} = (\text{column number}) - (\text{row number}) \text{ where } i \text{ occurs in } T
\]

Since 1 must occupy the $(1,1)$-box of any standard Young tableau, we always have $c_{T,1} = 0$.

Any standard Young tableau $T$ is determined by its content. To see this, consider the fibers

\[
c_T^{-1}(a) = \{ i \mid c_{T,i} = a \} \quad (a \in \mathbb{Z}).
\]

The size of these fibers tell us the shape of $T$ or, equivalently, the underlying Young diagram: there must be $|c_T^{-1}(a)|$ boxes on the $a$-diagonal. The elements of $c_T^{-1}(a)$ give the content of these boxes: we must fill them into the boxes on the $a$-diagonal in increasing order from top to bottom. In this way, we reconstruct $T$ from $c_T$.

**Example 4.6.** Suppose we are given $c_T = (0, 1, -1, 2, 3, -2, 0, -1, 1)$. The nonempty fibers are $c_T^{-1}(-2) = \{6\}$, $c_T^{-1}(-1) = \{3, 8\}$, $c_T^{-1}(0) = \{1, 7\}$, $c_T^{-1}(1) = \{2, 9\}$, $c_T^{-1}(2) = \{4\}$ and $c_T^{-1}(3) = \{5\}$. Thus, the resulting standard Young tableau $T$ is

Of course, not every $n$-tuple of integers is the content of a standard Young tableau with $n$ boxes, since there are only finitely many such Young tableaux. We also know already that the first component of any content vector must always be 0. Our next goal will be to give a complete description of the following set:
Cont\((n)\) \text{ def } = \{ T \text{ is a standard Young tableau with } n \text{ boxes} \} \subseteq \mathbb{Z}^n

\textbf{Example 4.7.} We list all standard Young tableaux with 4 boxes and their content vectors; these form the set Cont\((4)\). Note that, quite generally, reflecting a given standard Young tableau \(T\) across the diagonal \(y = x\) results in another Young tableau, \(T^c\), satisfying \(c_T^c = -c_T\).

\begin{align*}
\begin{array}{cccc}
1 & 2 & 3 & 4 \\
0 & 1 & 2 & 3 \\
\end{array} & \begin{array}{cccc}
3 & 4 & 1 & 2 \\
0 & -1 & 1 & 2 \\
\end{array} & \begin{array}{cccc}
1 & 2 & 3 & 4 \\
0 & 1 & 2 & -1 \\
\end{array} & \begin{array}{cccc}
3 & 4 & 1 & 2 \\
0 & -1 & 1 & 0 \\
\end{array} \\
\begin{array}{cccc}
1 & 2 & 3 & 4 \\
0 & -1 & -2 & -3 \\
\end{array} & \begin{array}{cccc}
3 & 4 & 1 & 2 \\
0 & 1 & -1 & -2 \\
\end{array} & \begin{array}{cccc}
1 & 2 & 3 & 4 \\
0 & -1 & 1 & -2 \\
\end{array} & \begin{array}{cccc}
1 & 2 & 3 & 4 \\
0 & 1 & -1 & 0 \\
\end{array}
\end{align*}

\textbf{An Equivalence Relation.} As we have remarked above, the shape of a standard Young tableau \(T\) is determined by the multiplicities \(\lfloor c_T^{-1}(a) \rfloor\) for \(a \in \mathbb{Z}\). Thus, two standard Young tableaux \(T\) and \(T'\) have the same shape if and only if \(c_T^c = s \cdot c_T\) for some \(s \in S_n\), where \(S_n\) acts on \(\mathbb{Z}^n\) by place permutations:

\[ s.(c_1, \ldots, c_n) := (c_{s^{-1}1}, \ldots, c_{s^{-1}n}). \]

Alternatively, \(T\) and \(T'\) have the same shape if and only if \(T' = sT\) for some \(s \in S_n\), where \(sT\) denotes the tableau that is obtained from \(T\) by replacing the entries in all boxes by their images under the permutation \(s\). The reader will readily verify (Exercise 4.4.2) that

\[ (4.23) \quad c_{sT} = s \cdot c_T. \]

Define an equivalence relation \(\sim\) on Cont\((n)\) by

\[ c_T \sim c_{T'} \iff c_{T'} = s \cdot c_T \text{ for some } s \in S_n \]
\[ \iff T \text{ and } T' \text{ have the same shape} \]
\[ \iff T \text{ and } T' \text{ describe paths } \square \rightarrow \ldots \text{ in } \mathbb{Y} \]
\[ \text{having the same endpoint.} \]

We summarize the correspondences discussed in the foregoing:

\[ (4.24) \quad \text{Cont}(n) \overset{\sim}{\longleftrightarrow} \left\{ \text{standard Young tableaux with } n \text{ boxes} \right\} \overset{\sim}{\longleftrightarrow} \left\{ \text{paths } \square \rightarrow \ldots \text{ in } \mathbb{Y} \right\} \]
\[ \text{forget places \quad empty boxes} \]
\[ \overset{\sim}{\longleftrightarrow} \left\{ \text{Young diagrams with } n \text{ boxes} \right\} \overset{\sim}{\longleftrightarrow} \mathcal{P}_n \]
Admissible Transpositions. Not all place permutations of the content vector \( c_T \) of a given standard Young tableau \( T \) necessarily produce another content vector of a standard Young tableau. For one, the first component must always be 0. The case of the transpositions \( s_i = (i, i + 1) \in S_n \) will be of particular interest; \( s_i \) swaps the boxes of \( T \) containing \( i \) and \( i + 1 \) while leaving all other boxes untouched and \( s_i(c_1, \ldots, c_n) = (c_1, \ldots, c_{i-1}, c_{i+1}, c_i, c_{i+2}, \ldots, c_n) \). It is easy to see (Exercise 4.4.2) that \( s_i T \) is another standard Young tableau if and only if \( i \) and \( i + 1 \) occur in different rows and columns of \( T \). In this case, \( s_i \) will be called an admissible transposition for \( T \). Note that the boxes containing \( i \) and \( i + 1 \) belong to different rows and columns of \( T \) if and only if these boxes are not direct neighbors in \( T \), sharing a common boundary line, and this is also equivalent to the condition that the boxes of \( T \) containing \( i \) and \( i + 1 \) do not lie on adjacent diagonals. To summarize,

\[
\begin{align*}
\text{\( s_i \) is admissible for \( T \)} & \iff \text{\( i \) and \( i + 1 \) belong to different rows and columns of \( T \)} \\
& \iff \text{\( c_{T, i+1} \neq c_{T, i} \pm 1 \).}
\end{align*}
\]

Alternatively, \( s_i \) is admissible for \( c = (c_1, \ldots, c_n) \in \text{Cont}(n) \) iff \( c_{i+1} \neq c_i \pm 1 \).

For a given partition \( \lambda \vdash n \), consider the particular \( \lambda \)-tableau \( T(\lambda) \) that is obtained from the Young diagram of \( \lambda \) by filling 1, 2, \ldots, \( n \) into the boxes as in the picture on the left. Clearly, for any \( \lambda \)-tableau \( T \), there is a unique \( s \in S_n \) with \( T = sT(\lambda) \). It is a standard fact that the transpositions \( s_1, \ldots, s_{n-1} \) generate \( S_n \) and that the minimal length of a product representing a given permutation \( s \in S_n \) in terms of these generators is equal to the number of inversions of \( s \), that is, the number of pairs \( (i, j) \in [n] \times [n] \) with \( i < j \) but \( si > sj \); see Exercise 4.4.3 or Example 7.10. This number, called the length of \( s \), will be denoted by \( \ell(s) \).

**Lemma 4.8.**

(a) Let \( \lambda \vdash n \) and let \( T \) be a \( \lambda \)-tableau. Then there exists a sequence \( s_{i_1}, \ldots, s_{i_l} \) of admissible transpositions in \( S_n \) such that \( s_{i_1} \ldots s_{i_l} T = T(\lambda) \) and \( l = \ell(s_{i_1} \ldots s_{i_l}) \).

(b) Let \( c, c' \in \text{Cont}(n) \). Then \( c \sim c' \) if and only if there exists a finite sequence of admissible transpositions that transforms \( c \) into \( c' \).

**Proof.** (a) Let \( n_T \) be the number in the box at the end of the last row of \( T \). We argue by induction on \( n \) and \( n - n_T \). The case \( n = 1 \) being trivial, assume that \( n > 1 \).

If \( n_T = n \) then remove the last box from \( T \) and let \( T' \) denote the resulting standard Young tableau, of shape \( \lambda' \vdash n - 1 \). By induction, we can transform \( T' \) into \( T(\lambda') \) by a sequence of admissible transpositions given by Coxeter generators \( s \in S_{n-1} \) and the sequence may be chosen to have the desired length. The same sequence will move \( T \) to \( T(\lambda) \).
Now assume that \( n_T < n \). Since the box of \( T \) containing \( n_T + 1 \) cannot occur in the same row or column as the last box, containing \( n_T \), the transposition \( s_{n_T} \) is admissible for \( T \) by (4.25). The \( \lambda \)-tableau \( T' = s_{n_T}T \) satisfies \( n_T = n_T + 1 \). By induction, there is a finite sequence \( s_{i_1}, \ldots, s_{i_l} \) of admissible transpositions such that \( s = s_{i_1} \cdots s_{i_l} \) satisfies \( sT' = T(\lambda) \) and \( l = \ell(s) \). It follows that \( ss_{n_T}T = T(\lambda) \) and \( l + 1 = \ell(ss_{n_T}) \), where the latter equality holds because \( s(n_T) < n = s(n_T + 1) \) (Exercise 4.4.3).

(b) Clearly, the existence of a sequence \( s_{i_1}, \ldots, s_{i_l} \) such that \( s_{i_1} \cdots s_{i_l} . c = c' \) implies that \( c \sim c' \). Conversely, if \( c = c_T \sim c' = c_{T'} \), then \( T \) and \( T' \) are \( \lambda \)-tableaux for the same \( \lambda \). It follows from (a) that there is a sequence \( s_{i_1}, \ldots, s_{i_l} \) of admissible transpositions such that \( s_{i_1} \cdots s_{i_l} . T = T' \). Hence, \( s_{i_1} \cdots s_{i_l} . c = c' \) by (4.23). \( \square \)

**Description of Cont\((n)\).** The following proposition gives the desired description of the set Cont\((n)\). Since our ultimate goal is to show that Cont\((n)\) is identical to the subset Spec\((n) \subseteq \mathbb{K}^n \) as defined in §4.2.5, we view Cont\((n) \subseteq \mathbb{K}^n \). Note, however, that conditions (i) and (ii) below imply that Cont\((n) \subseteq \mathbb{Z}^n \).

**Proposition 4.9.** Cont\((n)\) is precisely the set of all \( c = (c_1, c_2, \ldots, c_n) \in \mathbb{K}^n \) satisfying the following conditions:

(i) \( c_1 = 0; \)

(ii) \( c_i - 1 \text{ or } c_i + 1 \in \{c_1, c_2, \ldots, c_{i-1}\} \) for all \( i \geq 2; \)

(iii) if \( c_i = c_j = a \) for \( i < j \) then \( \{a + 1, a - 1\} \subseteq \{c_{i+1}, \ldots, c_{j-1}\} \).

**Proof.** Let \( C(n) \) denote the set of \( n \)-tuples \( c = (c_1, c_2, \ldots, c_n) \in \mathbb{K}^n \) satisfying conditions (i) – (iii). We need to show that Cont\((n) = C(n) \).

We first check that Cont\((n) \subseteq C(n) \). As we have observed earlier, (i) certainly holds if \( c = c_T \) for some standard Young tableau \( T \), because the number 1 must be in the \((1,1)\)-box. Any \( i \geq 2 \) must occupy a box of \( T \) in position \((x, y)\) with \( x > 1 \) or \( y > 1 \). In the former case, let \( j \) be the entry in the \((x - 1, y)\)-box. Then \( j < i \) and \( c_j = y - (x - 1) = c_i + 1, \) whence \( c_i + 1 \in \{c_1, c_2, \ldots, c_{i-1}\} \). In an analogous fashion, one shows that \( c_i - 1 \in \{c_1, c_2, \ldots, c_{i-1}\} \) if \( y > 1 \), proving (ii).

Now suppose that \( c_i = c_j = a \) for \( i < j \). Then the entries \( i \) and \( j \) both lie on the \( a \)-diagonal; say \( i \) occupies the \((x, x + a)\)-box and \( j \) the \((x', x' + a)\)-box, with \( x < x' \). Let \( k \) and \( l \) denote the entries in the boxes at positions \((x + 1, x + a)\) and \((x' - 1, x' + a)\), respectively. Then \( k, l \in \{i+1, \ldots, j-1\} \) and \( c_k = (x + a) - (x + 1) = a - 1, c_l = (x' + a) - (x' - 1) = a + 1 \). This proves (iii), thereby completing the proof of the inclusion Cont\((n) \subseteq C(n) \).
For the reverse inclusion, Cont\(n\) \(\supseteq\) C\(n\), we proceed by induction on \(n\).

The case \(n = 1\) being clear, with C\(1\) = \{(0)\} = Cont\(1\), assume that \(n > 1\)
and that C\(n-1\) \(\subseteq\) Cont\(n-1\). Let \(c = (c_1, c_2, \ldots, c_n) \in\) C\(n\) be given. Clearly, the truncated \(c' = (c_1, c_2, \ldots, c_{n-1})\) also satisfies conditions (i) – (iii); so \(c' \in\) C\(n-1\) \(\subseteq\) Cont\(n-1\). Therefore, there exists a (unique) standard Young tableau \(T'\) with \(c_T = c'\). We wish to add a box containing the number \(n\) to \(T'\) so as to obtain a standard Young tableau, \(T\), with \(c_T = c\). Thus, the new box must be placed on the \(c_n\)-diagonal, \(y = x + c_n\), at the first slot not occupied by any boxes of \(T'\). We need to check that the resulting \(T\) has the requisite “flag shape” of a partition; the monotonicity requirement for standard Young tableaux is then automatic, because the new box contains the largest number.

First assume that \(c_n \notin \{c_1, c_2, \ldots, c_{n-1}\}\); so \(T'\) has no boxes on the \(c_n\)-diagonal. Since there are no gaps between the diagonals of \(T'\), the values \(c_1, \ldots, c_{n-1}\), with repetitions omitted, form an interval in \(\mathbb{Z}\) containing \(0\). Therefore, if \(c_n > 0\) then \(c_n > c_i\) for all \(i < n\), while condition (ii) tells us that \(c_n - 1 \in \{c_1, c_2, \ldots, c_{n-1}\}\). Thus, \(c_n = \max\{c_i \mid 1 \leq i \leq n-1\} + 1\) and the new box labeled \(n\) is added at the end of the first row of \(T'\), at position \((1, 1 + c_n)\). Similarly, if \(c_n < 0\), then the new box is added at the bottom of the first column of \(T'\). In either case, the resulting \(T\) has flag shape.

Finally assume that \(c_n \in \{c_1, c_2, \ldots, c_{n-1}\}\) and choose \(i < n\) maximal with \(c_i = c_n =: a\). Then the box labeled \(i\) is the last box on the \(a\)-diagonal of \(T'\).

We also know from condition (iii) that there exist \(r, s \in \{i+1, \ldots, n-1\}\) with \(c_r = a-1\) and \(c_s = a+1\). Both \(r\) and \(s\) are unique. Indeed, if \(i < r < r' < n\) and \(c_r = a-1 = c_{r'}\), then \(a \in \{c_{r+1}, \ldots, c_{r'-1}\}\) by (iii), contradicting maximality of \(i\). This shows uniqueness of \(r\); the argument for \(s\) is analogous. Therefore, \(T'\) has unique boxes on the \((a-1)\)- and \((a+1)\)-diagonals with entries > \(i\). Necessarily these boxes are the last ones on their respective diagonals and they must be neighbors of \(i\), as in the picture above. Therefore, the new box labeled \(n\) is slotted in to the corner formed by the boxes with \(i, r\) and \(s\), again resulting in the desired flag shape. \(\square\)

### 4.4.2. Weights

Returning to the representation theoretic side of matters, let us begin with a few reminders from §4.2.5. With \(X_1, \ldots, X_n \in \mathcal{G}_n\) denoting the JM-elements, the spectrum of \(\mathcal{G}_n\) is in one-to-one correspondence with the set

\[
\text{Spec}(n) = \{(\phi(X_1), \phi(X_2), \ldots, \phi(X_n)) \in \mathbb{K}^n \mid \phi \in \text{Hom}_{\text{Alg}}(\mathcal{G}_n, \mathbb{K})\}.
\]
By (4.12) we have the following bijections:

\[ \Spec(n) \xrightarrow{\sim} \Hom_{\Alg_k}(G\mathbb{Z}_n, k) \xrightarrow{\sim} \{ \text{paths } \mathbb{I}_{S_1} \to \ldots \text{ in } \mathbb{B} \} \]

(4.26)

\[ \alpha_T := (\phi_T(X))_1^n \longleftrightarrow \phi_T \longleftrightarrow T \]

The first component of each \( \alpha_T \) is 0, because \( X_1 = 0 \). Recall from (4.10) that each \( \phi_T \) is a weight of a unique \( V \in \Irr S_n \) and that the weight space is spanned by the GZ-basis vector \( v_T \in V \) that is given by the path \( T: \mathbb{I}_{S_1} \to \cdots \to V \) in \( \mathbb{B} \):

(4.27)

\[ a.v_T = \phi_T(a)v_T \quad (a \in \mathbb{GZ}_n). \]

We will also write elements of \( \Spec(n) \) simply as \( n \)-tuples,

\[ \alpha = (a_1, a_2, \ldots a_n) \in k^n, \]

with \( a_1 = 0 \). Let \( V(\alpha) \in \Irr S_n \) denote the irreducible representation having weight \( \alpha \) and let \( v_\alpha \in V(\alpha) \) be the corresponding GZ-basis vector of \( V(\alpha) \). Then (4.27) becomes

(4.28)

\[ X_k.v_\alpha = a_k v_\alpha \quad \text{for } k = 1, \ldots , n. \]

The vector \( v_\alpha \in V(\alpha) \) is determined by these equations up to a scalar multiple. We will scale the weight vectors \( v_\alpha \) in a consistent way in §4.5.1, but this will not be necessary for now.

**Another Equivalence Relation.** Define \( \approx \) on \( \Spec(n) \) by

\[ \alpha_T \approx \alpha_T' \iff \phi_T \text{ and } \phi_{T'} \text{ are weights of the same representation } \in \Irr S_n \]

\[ \iff T \text{ and } T' \text{ are paths } \mathbb{I}_{S_1} \to \cdots \text{ in } \mathbb{B} \]

with the same endpoint \( \in \Irr S_n \).

Alternatively, \( \alpha \approx \alpha' \) if and only if \( V(\alpha) = V(\alpha') \). From (4.26) we obtain the following bijections:

\[ \Spec(n) \xrightarrow{\sim} \{ \text{paths } \mathbb{I}_{S_1} \to \cdots \text{ in } \mathbb{B} \} \]

(4.29)

\[ \Spec(n)/\approx \longleftrightarrow \{ \text{paths } \mathbb{I}_{S_1} \to \cdots \text{ in } \mathbb{B} \} \]

\[ \Spec(n)/\approx \xrightarrow{\sim} \Irr S_n \]

**Example 4.10** (\( \Spec(4) \)). We need to find the GZ-basis and weights of each \( V \in \Irr S_4 \). For the representation \( \mathbb{I} \), this is trivial: for any \( n \), the unique weight of \( \mathbb{I}_{S_n} \) is \( (0,1,2,\ldots, n-1) \), because \( X_k.1 = (k-1) \). Example 4.3 provides us with the GZ-bases of \( V_3 \) and \( \tilde{V}_2 \). In the case of \( \tilde{V}_2 \), note that \( X_4 \) acts via the canonical map.
\( \mathbb{S}_q \rightarrow \mathbb{S}_3 \), which sends \( X_q \mapsto (2, 3) + (1, 3) + (1, 2) \). Next, for any \( n \) and any \( V \in \text{Irr} \mathbb{S}_n \), the sign twist \( V^\pm = \text{sgn} \otimes V \) has the “same” GZ-basis as \( V \) but with weights multiplied by \(-1\):

\[
X_k \cdot (1 \otimes v_\alpha) = \sum_{i < k} (i, k) \cdot (1 \otimes v_\alpha) = -(1 \otimes X_k \cdot v_\alpha) \quad (4.28)
\]

Leaving the detailed verifications to the reader (Exercise 4.4.1), we list the elements of \( \text{Spec}(4) \), grouped according to \( \approx \)-equivalence:

\[
\begin{align*}
1 : & \quad (0, 1, 2, 3) \\
V_1 : & \quad (0, -1, 1, 2) \\
& \quad (0, 1, -1, 2) \\
& \quad (0, 1, 2, -1) \\
\text{sgn} : & \quad (0, -1, -2, -3) \\
V_2^\pm : & \quad (0, 1, -1, -2) \\
& \quad (0, -1, 1, -2) \\
& \quad (0, -1, -2, 1)
\end{align*}
\]

Observe that \( \approx \)-equivalence amounts to equality up to a place permutation in Example 4.10. Comparison with Example 4.7 further shows that \( \text{Spec}(4) = \text{Cont}(4) \). Theorem 4.12 will give the same conclusions for all \( n \).

**Description of Spec(\( n \)).** For the proof of Theorem 4.12, we will need the following technical proposition, which will also see some use in Section 4.5.

**Proposition 4.11.** Let \( \alpha = (a_1, a_2, \ldots, a_n) \in \text{Spec}(n) \) and let \( v_\alpha \in V(\alpha) \) be a weight vector of weight \( \alpha \) as in (4.28). Then:

(a) \( a_i \neq a_{i+1} \) for all \( i \); so we may put \( d_i := (a_{i+1} - a_i)^{-1} \).

(b) \( d_i = \pm 1 \) if and only if \( v_\alpha \) is an eigenvector for \( s_i \), in which case \( s_i \cdot v_\alpha = d_i v_\alpha \).

(c) If \( d_i \neq \pm 1 \), then \( s_i \cdot \alpha \in \text{Spec}(n) \), \( s_i \cdot \alpha \approx \alpha \) and \( v_{s_i \cdot \alpha} = s_i \cdot v_\alpha - d_i v_\alpha \) (up to a scalar factor).

**Proof.** We will employ the subalgebra \( \mathcal{B}_l = \mathbb{k}[X_l, X_{l+1}, s_l] \subseteq \mathbb{k}\mathbb{S}_n \).³ Exercise 4.1.3 tells us that \( \mathcal{B}_l \) is semisimple and, by Exercise 4.1.1, we have the relations

\[
X_{i+1} s_l = s_l X_i + 1 \quad \text{and} \quad X_l s_i = s_i X_{i+1}.
\]

The subspace \( V_l = \mathbb{k} v_\alpha + \mathbb{k} s_i \cdot v_\alpha \subseteq V(\alpha) \) is in fact a \( \mathcal{B}_l \)-module. Indeed, \( V_l \) is clearly stable under the action of \( s_l \) and it follows from (4.28) and (4.30) that \( V_l \) is stable under \( X_l \) and \( X_{l+1} \) as well:

\[
X_{i+1} s_l \cdot v_\alpha = a_i s_l \cdot v_\alpha + v_\alpha \quad \text{and} \quad X_l s_i \cdot v_\alpha = a_{i+1} s_i \cdot v_\alpha - v_\alpha
\]

³The role of \( \mathcal{B}_l \) is analogous to the one played by \( \mathfrak{sl}_2 \)-triples in the representation theory of semisimple Lie algebras; see Chapter 6.
First assume that \( \dim V_i = 1 \) or, equivalently, \( v_\alpha \) is an eigenvector for \( s_i \). Then we must have \( s_i.v_\alpha = \pm v_\alpha \) and
\[
a_{i+1}v_\alpha = X_{i+1}v_\alpha = \pm X_{i+1}s_i.v_\alpha = \pm (s_iX_i + 1).v_\alpha = (a_i \pm 1)v_\alpha
\]
Therefore, \( a_{i+1} = a_i \pm 1 \) holds in this case or, equivalently, \( d_i = \pm 1 \).

Now let \( \dim V_i = 2 \). Then it follows from (4.28) and (4.31) that the matrices of \( X_i, X_{i+1} \) and \( s_i \) for the basis \((v_\alpha, s_i.v_\alpha)\) of \( V_i \) are
\[
\begin{pmatrix} a_i & -1 \\ 0 & a_{i+1} \end{pmatrix}, \quad \begin{pmatrix} a_{i+1} & 1 \\ 0 & a_i \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.
\]
Since the actions of \( X_i \) and \( X_{i+1} \) on \( V(\alpha) \in \text{Irr} \mathcal{S}_n \) are diagonalizable, they are so on the subspace \( V_i \) as well. Hence the first two matrices are diagonalizable, which forces \( a_i \neq a_{i+1} \). Together with the preceding paragraph, this proves (a). Consider the new basis \((v_\alpha, w_\alpha = s_i.v_\alpha - d_i v_\alpha)\) of \( V_i \). The matrices of \( X_i, X_{i+1} \) and \( s_i \) now take the form
\[
\begin{pmatrix} a_i & 0 \\ 0 & a_{i+1} \end{pmatrix}, \quad \begin{pmatrix} a_{i+1} & 0 \\ 0 & a_i \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & 0 \\ 0 & -d_i \end{pmatrix}.
\]
If \( a_{i+1} = a_i \pm 1 \) then the last matrix becomes \( \begin{pmatrix} \pm 1 & 0 \\ 0 & \pm 1 \end{pmatrix} \). Thus \( k w_\alpha \) is the unique simple \( \mathcal{B}_i \)-submodule of \( V_i \), contradicting semisimplicity of \( \mathcal{B}_i \). Therefore, we must have \( a_{i+1} \neq a_i \pm 1 \), completing the proof of (b). The first two matrices above show that \( X_j.w_\alpha = a_{i+1}w_\alpha \) and \( X_{i+1}.w_\alpha = a_jw_\alpha \). For \( j \neq \{i, i+1\} \), we have \( X_j.w_\alpha = a_jw_\alpha \) by (4.28), because \( s_iX_j = X_js_i \) (Exercise 4.1.1). This shows that \( w_\alpha \) is a weight vector of weight \( s_i.\alpha \) for \( \mathcal{G} \mathbb{Z}_n \), and hence \( s_i.\alpha \) is a weight of \( V(\alpha) \). We have thus shown that \( s_i.\alpha \in \text{Spec}(n) \) and \( s_i.\alpha \approx \alpha \). Moreover, \( v_{s_i.\alpha} = w_\alpha \) up to a scalar factor. This proves (c), finishing the proof of the proposition. \( \square \)

### 4.4.3. Identification of Contents and Weights

In this subsection, we will to merge the following two commutative diagrams, previously stated as (4.24) and (4.29):

\[
\begin{array}{ccc}
\{ \text{paths} \square \rightarrow \ldots \text{in} \mathcal{P}_n \} & \sim & \text{Cont}(n) \\
\{ \text{with endpoint} \in \mathcal{P}_n \} & \sim & \text{Spec}(n) \end{array}
\]

\[
\begin{array}{ccc}
\text{Paths} & \sim & \text{Spec}(n) \\
\text{Irr} \mathcal{S}_n \end{array}
\]

**Theorem 4.12.** \( \text{Cont}(n) = \text{Spec}(n) \) and \( \sim = \approx \).

**Proof.** We will prove the following two claims:

1. \( \text{Spec}(n) \subseteq \text{Cont}(n) \).
2. If \( \alpha \in \text{Spec}(n) \), \( c \in \text{Cont}(n) \) and \( \alpha \sim c \), then \( c \in \text{Spec}(n) \) and \( \alpha \approx c \).
Assuming (1) and (2), we deduce the theorem as follows. For any \( c \in \text{Cont}(n) \), let \([c]_\sim \) denote the \( \sim \)-equivalence class; likewise for \( \approx \). It follows from (1) and (2) that either \([c]_\sim \cap \text{Spec}(n) = \emptyset \) or else \([c]_\sim \subseteq [c]_\approx \subseteq \text{Spec}(n) \). Consequently, 
\[
|\text{Cont}(n)/\sim| \geq |\text{Spec}(n)/\approx|
\]
and equality holds if and only if \( \text{Cont}(n) = \text{Spec}(n) \) and \( \sim = \approx \). But we already know that \( |\mathcal{P}_n^\mathcal{I}| = |\text{Irr} \mathcal{S}_n| \) (Corollary 3.21 and §3.5.2). Hence, the two diagrams above give \( |\text{Cont}(n)/\sim| = |\mathcal{P}_n^\mathcal{I}| = |\text{Irr} \mathcal{S}_n| = |\text{Spec}(n)/\approx| \), as desired. It remains to prove (1) and (2).

**Proof of (1).** We will verify that any \( \alpha = (a_1, \ldots, a_n) \in \text{Spec}(n) \) satisfies conditions (i), (ii) and (iii) of Proposition 4.9. As we have observed earlier, we certainly have \( a_1 = 0 \); so (i) holds.

For (ii), we need to show that \( \{a_i - 1, a_i + 1\} \cap \{a_1, \ldots, a_{i-1}\} \neq \emptyset \) for all \( 1 < i \leq n \). Suppose this fails for \( i \). Then \( a_i \neq a_{i-1} \pm 1 \) and so Proposition 4.11(c) tells us that \( s_{i-1} \cdot \alpha = (a_1, \ldots, a_{i-2}, a_i, a_{i-1}, \ldots, a_n) \in \text{Spec}(n) \). Also, \( a_i \neq a_{i-2} \pm 1 \) and hence we may use Proposition 4.11(c) again to switch \( a_i \) and \( a_{i-2} \). Continuing in this fashion, we may move \( a_i \) all the way to the first position to obtain \( (a_i, a_1, \ldots) \in \text{Spec}(n) \). Therefore, \( a_i = 0 = a_1 \) by (i), contradicting Proposition 4.11(a). This proves (ii).

Finally, suppose that \( \alpha \) violates (iii); so \( a_i = a_j =: a \) for some \( i < j \) but \( \{a-1, a+1\} \nsubseteq \{a_i+1, \ldots, a_{j-1}\} \). Assume that \( \alpha \), \( i \) and \( j \) have been chosen so that \( j - i \) is minimal. Then we must have \( a_{i+1} = a \pm 1 \) and \( a_{j-1} = a \pm 1 \). For, otherwise Proposition 4.11(c) would allow us to replace \( \alpha \) by switching \( a = a_i \) with its neighbor to the right or \( a = a_j \) with its neighbor to the left, thereby decreasing the value of \( j - i \). Thus, \( \alpha \) has the form \( \alpha = (\ldots, a, a \pm 1, \ldots, a \pm 1, a, \ldots) \) with \( a \) in positions \( i \) and \( j \). Furthermore, inasmuch as \( \alpha \) violates (iii), we must in fact have \( a_{i+1} = a_{j-1} = a \pm 1 \). Now minimality of \( j - i \) forces \( i + 1 = j - 1 \); so \( \alpha = (\ldots, a, a \pm 1, a, \ldots) \). Proposition 4.11(b) now yields \( s_i \cdot v_\alpha = \pm v_\alpha \) and \( s_{i+1} \cdot v_\alpha = \mp v_\alpha \). It follows that

\[
 s_i s_{i+1} s_i v_\alpha = \pm v_\alpha \neq \mp v_\alpha = s_i s_{i+1} s_i v_\alpha,
\]

contradicting the braid relation \( s_i s_{i+1} s_i = s_{i+1} s_i s_{i+1} \). This proves condition (iii), and hence (1) is proved.

**Proof of (2).** Let \( \alpha \in \text{Spec}(n) \) and \( c \in \text{Cont}(n) \) be given such that \( \alpha \sim c \); so the \( n \)-tuples \( \alpha \) and \( c \) differ only by a place permutation. Then we know from Lemma 4.8 that there is a sequence \( s_{i_1}, \ldots, s_{i_t} \) of admissible transpositions such that \( c = s_{i_t} \ldots s_{i_1} \cdot \alpha \). Recall from (4.25) that the transposition \( s_i \) is admissible

\[
(\xi = (x_1, \ldots, x_n)) \text{ if and only if } x_{i+1} \neq x_i \pm 1 \text{.}
\]

Since any such transposition \( s_i \) transforms an element of \( \text{Spec}(n) \) into a \( \approx \)-equivalent element of \( \text{Spec}(n) \) by Proposition 4.11(c), it follows that \( c \in \text{Spec}(n) \) and \( c \approx \alpha \). This completes the proof of the theorem.

Now, we are finally ready to prove the Graph Isomorphism Theorem (§4.3.2); it is an immediate consequence of Theorem 4.12.
**Corollary 4.13.** \( \mathbb{Y} \cong \mathbb{B} \).

**Proof.** Theorem 4.12 allows us to merge the two diagrams preceding the statement of the theorem into the following commutative diagram:

\[
\begin{array}{c}
\{ \text{paths} \square \rightarrow \cdots \text{in } \mathbb{Y} \} \xrightarrow{\sim} \text{Cont}(n) = \text{Spec}(n) \xleftarrow{\sim} \{ \text{paths} \mathbb{I}_S \rightarrow \cdots \text{in } \mathbb{B} \} \\
\downarrow \text{remember} \downarrow \text{forget} \downarrow \text{remember} \\
\mathcal{P}_n \xleftarrow{\sim} \text{Cont}(n)/\sim = \text{Spec}(n)/\approx \xrightarrow{\sim} \text{Irr } S_n
\end{array}
\]

The bottom rows of all these diagrams, one for each \( n \), give the desired bijection between the vertices of \( \mathbb{Y} \) and \( \mathbb{B} \). Furthermore, we have an arrow \( \mu \rightarrow \lambda \) in \( \mathbb{Y} \) if and only if there is a path \( \square \rightarrow \cdots \mu \rightarrow \lambda \) in \( \mathbb{Y} \). If \( \mu \in \mathcal{P}_n \) and \( \lambda \in \mathcal{P}_{n+1} \), say, then this means that the content vector \( \in \text{Cont}(n) \) corresponding to the initial part \( \square \rightarrow \cdots \mu \) of this path gives the first \( n \) components of content vector \( \in \text{Cont}(n+1) \) corresponding to the full path \( \square \rightarrow \cdots \mu \rightarrow \lambda \). Similar remarks apply to arrows and paths in \( \mathbb{B} \) of the correspondences above. The Graph Isomorphism Theorem follows from this. \( \square \)

**Some Examples.** As in §4.3.2 let \( \lambda \mapsto V^\lambda \) be the bijection on vertices given by the above graph isomorphism \( \mathbb{Y} \rightarrow \mathbb{B} \). In order to determine the partition \( \lambda \in \mathcal{P}_n \) corresponding to a given \( V \in \text{Irr } S_n \), it suffices to find a weight \( \alpha \) of \( V \), coming from some GZ-basis vector of \( V \), then view \( \alpha \) as the content vector of a (unique) standard Young tableau, and finally take \( \lambda \) to be the partition given by the corresponding Young diagram.

**Examples 4.14.** (a) As we have observed in Example 4.10, the only weight of the trivial representation \( \mathbb{I}_S \) is \((0, 1, 2, \ldots, n - 1)\). This \( n \)-tuple, viewed as a content vector, describes the standard Young tableau \( \begin{array}{c} 1 \ 2 \ 3 \ \cdots \ n \end{array} \), which in turn gives the partition \( \lambda = (n) \). Thus \( \mathbb{I}_S = V^{(n)} \).

(b) The vector \( v_1 = b_1 - b_2 \) occurs in the GZ-basis of the standard representation \( V_{n-1} \) of \( S_n \) (Example 4.3). One checks that \( X_2.v_1 = -v_1 \) and \( X_k.v_1 = (k-2)v_1 \) for \( k \geq 3 \); so the weight of \( v_1 \) is \( (0, -1, 1, 2, \ldots, n-2) \). This weight is the content of the standard Young tableau above, and hence we obtain the partition \( \lambda = (n-1, 1) \). Therefore, \( V_{n-1} = V^{(n-1, 1)} \).

In fact, it follows from (a) that \( \lambda^k V_{n-1} = V^{(n-k, 1^k)} \) for all \( k \) (Exercise 4.3.5).

(c) We had seen in Examples 4.10 and 4.7 that if \( \alpha \) is a weight of \( V \in \text{Irr } S_n \), then \( -\alpha \) is a weight of the sign twist \( V^\pm \) and if the content vector \( \alpha \) yields the partition \( \lambda \), then \( -\alpha \) will yield the conjugate partition \( \lambda^c \). Thus, \( (V^\lambda)^\pm = V^{\lambda^c} \).
Exercises for Section 4.4

4.4.1 (Details for Spec(4)). Verify the weights in Spec(4) as stated in Example 4.10. (Exercise 4.2.5 is useful.)

4.4.2 (Details for place permutations). Let T be a standard Young tableau with n boxes and let $s_i = (i, i+1) \in S_n$. Prove the following assertions made in the text.

(a) $c_{sT} = s.c_T$ for all $s \in S_n$.

(b) $s_i T$ is a standard Young tableau if and only if the boxes of $T$ containing $i$ and $i+1$ occur in different rows and columns of $T$.

(c) The boxes containing $i$ and $i+1$ belong to different rows and columns if and only if these boxes are not direct neighbors in $T$, and this happens if and only if the boxes containing $i$ and $i+1$ do not lie on adjacent diagonals of $T$.

4.4.3 (Length of permutations). Let $l(s)$ denote the number of inversions of $s \in S_n$. Show that

\[ l(sT) = \begin{cases} 
  l(s) + 1 & \text{if } s(i) < s(i+1) \\
  l(s) - 1 & \text{if } s(i) > s(i+1) 
\end{cases} \]

Deduce that $l(s) = \ell(s)$.

4.4.4 (Some decompositions into irreducibles). Let $V_{n-1} = V^{(n-1,1)}$ be the standard representation of $S_n$ and $M_n$ the standard permutation representation. Prove:

(a) $V_{n-1} \otimes M_n = I_{S_{n-2}}^{S_{n-1}} \oplus V^{(n-2,1,1)} \oplus V^{(n-2,2,1)}$.

(b) $V_{n-1} \otimes V_{n-1} = I \oplus V_{n-1} \oplus V^{(n-2,1,1)} \oplus V^{(n-2,2)}$.

4.5. The Irreducible Representations

The purpose of this section is to derive some explicit formulae for the action of $S_n$ on the irreducible representations $V^A$. In particular, we shall see that the GZ-basis of each $V^A$ can be scaled so that the matrices of all $s \in S_n$ have entries in $\mathbb{Q}$; for a different choice of normalization of the GZ-basis, the matrices will be orthogonal.

4.5.1. Realization over $\mathbb{Q}$

Let $\lambda \vdash n$ and let $V^A$ be the corresponding irreducible representation of $S_n$ as per the Graph Isomorphism Theorem. Since paths $\square_{S_i} \to \cdots \to V^A$ in $\mathbb{G}$ are in bijection with paths $\square \to \cdots \to \lambda$ in $\mathbb{Y}$ or, equivalently, $\lambda$-tableaux, we can rewrite (4.6) in the following form, with uniquely determined 1-dimensional subspaces $V^A_T$:

\[ V^A = \bigoplus_{T \text{ a } \lambda\text{-tableau}} V^A_T. \]

Specifically, $V^A_T$ is the $\mathbb{GZ}_n$-weight space of $V^A$ for the weight $c_T$, the content of $T$. We will now select a nonzero vector from each $V^A_T$ in a coherent manner. To this
end, let \( \pi_T : V^A \to V^A_T \) denote the projection along the sum of the weight spaces \( V^A_T \) with \( T' \neq T \) and fix
\[
0 \neq v(\lambda) \in V^A_{T(\lambda)}.
\]
Here \( T(\lambda) \) is the special \( \lambda \)-tableau considered in Lemma 4.8. Each \( \lambda \)-tableau \( T \) has the form \( T = s_T T(\lambda) \) for a unique \( s_T \in S_n \). Put
\[
(4.32) \quad v_T := \pi_T(s_T \cdot v(\lambda)).
\]
In the theorem below, we will check that all \( v_T \) are nonzero and, most importantly, that the action of \( S_n \) on the resulting GZ-basis of \( V^A \) is defined over \( \mathbb{Q} \). Adopting the notation of Proposition 4.11, we will write
\[
(4.33) \quad d_{T,i} := (c_{T,i+1} - c_{T,i})^{-1},
\]
where \( c_T = (c_{T,1}, \ldots, c_{T,n}) \) is the the content of \( T \). Thus, \( d_{T,i} \) is a nonzero rational number. Recall also from (4.25) that \( d_{T,i} \neq \pm 1 \) if and only if the Coxeter generator \( s_i \in S_n \) is admissible for \( T \), that is, \( s_T T \) is a \( \lambda \)-tableau.

**Theorem 4.15.** Let \( \lambda + n \). For each \( \lambda \)-tableau \( T \), let \( s_T, v_T \) and \( d_{T,i} \) be as in (4.32), (4.33). Then \( (v_T) \) is a GZ-basis of \( V^A \) and the action of \( s_i = (i, i+1) \in S_n \) on this basis is as follows:

(i) If \( d_{T,i} = \pm 1 \), then \( s_i v_T = d_{T,i} v_T \).

(ii) If \( d_{T,i} \neq \pm 1 \), then
\[
s_i v_T = \begin{cases} 
  d_{T,i} v_T + v_{s_T} & \text{if } s_T^{-1}(i) < s_T^{-1}(i+1); \\
  d_{T,i} v_T + (1 - d_{T,i}^2) v_{s_T} & \text{if } s_T^{-1}(i) > s_T^{-1}(i+1).
\end{cases}
\]

**Proof.** Proposition 4.11 in conjunction with (4.23), (4.25) implies that \( s_i V^A_T \subseteq V^A_T \) if \( s_i \) is not an admissible transposition for \( T \) and \( s_i V^A_T \subseteq V_{s_T T}^A + V^A_T \) if \( s_i \) is an admissible transposition for \( T \). Consequently, putting \( \ell(T) = \ell(s_T) \) and \( V^A_{<\ell} = \bigoplus_{T: \ell(T) < \ell} V^A_T \), we have \( s_i V^A_{<\ell} \subseteq V^A_{<\ell (i+1)} \) for all \( \ell \) and \( i \). We make the following

**Claim.** \( s_T v(\lambda) \equiv v_T \mod V^A_{<\ell(T)} \) and \( v_T \neq 0 \).

This is trivial for \( \ell(T) = 0 \), because \( s_T = 1, T = T(\lambda) \) and \( v_T = v(\lambda) \) in this case. So let \( \ell(T) > 0 \) and assume that the claim holds for all \( \lambda \)-tableaux \( T' \) with \( \ell(T') < \ell(T) \). By Lemma 4.8(a), \( s_T = s_{i_r} \ldots s_{i_1} \) with admissible transpositions \( s_{i_j} \) and \( \ell = \ell(T) \). Write \( s_T = s_{i_j} s' \) and consider the \( \lambda \)-tableau \( T' = s'T(\lambda) \). Thus, \( T = s_{i_j} T' \) and \( \ell(T') = \ell - 1 \). By induction, \( v_{T'} \in V^A_{T'} \) is a nonzero and \( v_{T'} = s' \cdot v(\lambda) + x \) with \( x \in V^A_{<\ell-1} \). Since \( s_{i_j} \) is admissible, it follows from Proposition 4.11(c) that
\[
0 \neq w_T := s_{i_j} v_{T'} - d_{T',i_j} v_{T'} = s_T v(\lambda) + y \in V^A_T
\]
with \( y = s_{i_j} x - d_{T',i_j} v_{T'} \in V^A_{<\ell} \). Therefore, \( \pi_T(y) = 0 \) and hence \( w_T = \pi_T(w_T) = \pi_T(s_T v(\lambda)) = v_T \), proving the claim.
The claim in particular yields the desired $GZ$-basis $(v_T)$ for $V^A$. Since $v_T \in V^A_T$, the $c_T$-weight space of $V^A_T$, Proposition 4.11 gives that $s_i.v_T = d_{T,i}v_T$ if $d_{T,i} = \pm 1$, and otherwise $s_i.v_T - d_{T,i}v_T$ equals $v_{s_iT}$ up to a scalar multiple, because both are nonzero vectors in the 1-dimensional weight space $V^A_{s_iT}$. First assume that $s^{-1}_T(i) < s^{-1}_T(i + 1)$ or, equivalently, $\ell(s_iT) > \ell(T)$ (Exercise 4.4.3). Then the claim implies that $s_i,v_T - d_i v_T = v_{s_iT}$; in fact, both vectors are congruent to $s_i s_T . v(\lambda)$ modulo $V^A_{s_T}$. Therefore, $s_i,v_T = d_{T,i} v_T + v_{s_iT}$ holds in this case. If $\ell(s_iT) < \ell(T)$, then apply the foregoing to $S = s_iT$, noting that $\ell(s_iS) = \ell(T) > \ell(S)$. We obtain $s_i,v_S = d_{S,i} v_S + v_{s_iS}$ with $d_{S,i} = (c_i - c_{i+1})^{-1} = -d_{T,i}$, because $c_S = s_i c_T = (c_1, \ldots, c_{i-1}, c_i c_{i+1}, c_{i-2}, \ldots, c_n)$. Therefore, $s_i,v_{s_iT} = -d_{T,i} v_{s_iT} + v_T$ and hence

$$v_{s_iT} = -d_{T,i} s_i v_{s_iT} + s_i v_T$$
$$= -d_{T,i} ( -d_{T,i} v_{s_iT} + v_T) + s_i v_T$$
$$= d_{T,i}^2 v_{s_iT} - d_{T,i} v_T + s_i v_T .$$

Equivalently, $s_i.v_T = d_{T,i} v_T + (1 - d_{T,i}^2) v_{s_iT}$, which finishes the proof. \[\square\]

Since the scalars $d_{T,i}$ in Theorem 4.15 are rational numbers, the matrices of the Coxeter generators $s_i$ for the specified basis $(v_T)$ of the irreducible representation $V = V^A$ have entries in $\mathbb{Q}$. Hence the same holds for all $s \in S_n$. Letting $V_\mathbb{Q}$ denote the $\mathbb{Q}$-subspace of $V$ that is spanned by $(v_T)$, we obtain a representation of the rational group algebra $\mathbb{Q}S_n$ such that $V = \mathbb{k} \otimes \mathbb{Q}V_\mathbb{Q}$. In other words, using Exercise 1.4.9 and the terminology employed there, all irreducible representations of $\mathbb{k}S_n$ are defined over $\mathbb{Q}$ and $\mathbb{Q}$ is a splitting field for $S_n$. This fact has been alluded to several times earlier. So, for the record:

**Corollary 4.16.** All irreducible representation of $S_n$ in characteristic 0 are defined over $\mathbb{Q}$. In particular, $\mathbb{Q}$ is a splitting field for $S_n$.

In this connection, it is worth noting that any finite-dimensional representation of a finite group that is defined over $\mathbb{Q}$ can actually be realized with matrices having entries in $\mathbb{Z}$ by a suitable choice of basis (Exercise 4.5.1). The next example spells out Theorem 4.15 for $V_{n-1}$. We have already seen in Example 4.14(b) that $V_{n-1} = V^A$ with $\lambda = (n - 1, 1)$.

**Example 4.17** (The standard representation $V_{n-1}$). For the partition $\lambda = (n - 1, 1)$, we have the $\lambda$-tableaux $T_2, T_3, \ldots, T_n$ with

$$T_j = \begin{array}{cccc}
1 & 2 & \cdots & j-1/p+1 \\
\hline
j & & & \\
\end{array} \begin{array}{cccc}
\cdots & \cdots & \cdots & n \\
\end{array}.$$
Thus, \( T(\lambda) = T_n, s_{T_j} = (j, j + 1, \ldots, n) \) and \( c_{T_j} = (0, 1, \ldots, j - 2, -1, j - 1, \ldots, n - 2) \). Therefore,

\[
d_{T_j,i} = \begin{cases} 
1 & \text{for } j \neq i, i + 1 \\
\frac{1}{\sqrt{2}} & \text{for } j = i \\
-\frac{1}{\sqrt{2}} & \text{for } j = i + 1 
\end{cases}
\]

Writing \( v(j) = v_{T_j} \) and noting that \( s_i \) interchanges \( T_j \) and \( T_{j+1} \), the action formulae in Theorem 4.15 become

\[
s_i v(j) = \begin{cases} 
v(j) & \text{for } j \neq i, i + 1 \\
\frac{1}{\sqrt{2}} v(i) + (1 - \frac{1}{t^2}) v(i + 1) & \text{for } j = i \\
v(i) - \frac{1}{\sqrt{2}} v(i + 1) & \text{for } j = i + 1 
\end{cases}
\]

### 4.5.2. Young’s Orthogonal Form

We now discuss a different choice of normalization for the GZ-basis of an irreducible representation \( V \in \text{Irr} S_n \) that results in orthogonal matrices for the operators \( s_V \) \((s \in S_n)\). Here, an **orthogonal matrix** is an invertible square matrix with real entries whose inverse is equal to its transpose. Again, the action formulae will involve the scalars \( d_{T,i} \), for a standard Young tableau \( T \); see (4.33). Note that \( 1 - d_{T,i}^2 \in \mathbb{Q}_+ \subseteq \mathbb{K} \), since \( d_{T,i} \) is a rational number of absolute value \( \leq 1 \). We fix an embedding of the field \( \mathbb{Q} = \{ c \in \mathbb{C} \mid c \text{ is algebraic over } \mathbb{Q} \} \) into \( \mathbb{K} \) and, for any \( q \in \mathbb{Q}_+ \), we let \( \sqrt{q} = q^{\frac{1}{2}} \in \mathbb{K} \) denote the image of the positive real square root of \( q \). Recall also that \( s_i T \) is a \( \lambda \)-tableau if and only if \( d_{T,i} \neq \pm 1 \); see (4.25).

**Theorem 4.18.** Let \( \lambda \vdash n \). There exists a GZ-basis \( (w_T)_T \) of a \( \lambda \)-tableau of \( V^\lambda \) such that the action of the Coxeter generator \( s_i = (i, i + 1) \in S_n \) is given by

\[
s_i w_T = d_{T,i} w_T + \sqrt{1 - d_{T,i}^2} w_{s_i T}.
\]

(If \( s_i T \) is not a \( \lambda \)-tableau or, equivalently, \( d_{T,i} = \pm 1 \), then the last term is missing.)

**Proof.** Put \( V = V^\lambda \), let \( (v_T) \) be the GZ-basis of \( V \) determined in Theorem 4.15, and let \( V_\mathbb{Q} \) denote the \( \mathbb{Q} \)-subspace of \( V \) that is spanned by \( \{ v_T \} \). Recall that \( V_\mathbb{Q} \) is stable under \( S_n \). By averaging an inner product of \( V_\mathbb{Q} \), we obtain a symmetric bilinear form \((\cdot, \cdot): V_\mathbb{Q} \times V_\mathbb{Q} \to \mathbb{Q} \) satisfying \((v, v) \neq 0 \) for \( v \neq 0 \) and \((s v, s v') = (v, v') \) for all \( v, v' \in V_\mathbb{Q} \) and \( s \in S_n \). Extend \((\cdot, \cdot)\) to a \( \mathbb{K} \)-bilinear form of \( V \) and define

\[
w_T := (v_T, v_T)^{-\frac{1}{2}} v_T \in V
\]

Thus, \( (w_T, w_T) = 1 \) and we also have \((w_T, w_T') = 0 \) if \( T \neq T' \); see Exercise 4.2.4.

Now let us describe the action of \( s_i \) on the GZ-basis \( (w_T) \). If \( d_{T,i} = \pm 1 \), then the formula \( s_i v_T = d_{T,i} v_T \) from Theorem 4.15 clearly also holds with \( w_T \) in place of \( v_T \), proving the asserted formula for \( s_i w_T \) in this case. Thus, we may assume
that \( d_{T,i} \neq \pm 1 \). If \( s_T^{-1}(i) < s_T^{-1}(i + 1) \), then we know from Theorem 4.15 that \( s_i.v_T = d_{T,i}v_T + v_{s_i,T} \). Therefore,

\[
(v_T, v_T) = (s_i.v_T, s_i.v_T)
= d_{T,i}^2(v_T, v_T) + 2d_{T,i}(v_T, v_{s_i,T}) + (v_{s_i,T}, v_{s_i,T})
= d_{T,i}^2(v_T, v_T) + (v_{s_i,T}, v_{s_i,T}),
\]

and so \((v_{s_i,T}, v_{s_i,T}) = (1 - d_{T,i}^2)(v_T, v_T)\). Now \( s_i.v_T = d_{T,i}v_T + v_{s_i,T} \) gives

\[
s_i.w_T = d_{T,i}w_T + (v_T, v_T)^{-\frac{1}{2}}v_{s_i,T}
= d_{T,i}w_T + (v_T, v_T)^{-\frac{1}{2}}(v_{s_i,T}, v_{s_i,T})^{\frac{1}{2}}w_{s_i,T}
= d_{T,i}w_T + (1 - d_{T,i}^2)^{\frac{1}{2}}w_{s_i,T},
\]
as desired. The verification in the case where \( s_T^{-1}(i) > s_T^{-1}(i + 1) \) is analogous using the formula \( s_i.v_T = d_{T,i}v_T + (1 - d_{T,i}^2)v_{s_i,T} \) from Theorem 4.15 instead. This proves the theorem.

The choice of basis in the theorem above is know as **Young’s orthogonal form**. Note that, if \( d_{T,i} = \pm 1 \), then \( w_T \) is a \( s_i \)-eigenvector with eigenvalue \( d_{T,i} \); if \( d_{T,i} \neq \pm 1 \), then \((w_T, w_{s_i,T})\) is the basis of a 2-dimensional \( s_i \)-stable subspace of \( V^\perp \) and the matrix of \( s_i \) with respect to this basis is orthogonal:

\[
\begin{pmatrix}
  d_{T,i} & \sqrt{1 - d_{T,i}^2} \\
  \sqrt{1 - d_{T,i}^2} & -d_{T,i}
\end{pmatrix}
\]

Thus, the matrix of \( s_i \) with respect to the basis of Theorem 4.18 is orthogonal, and hence so are the matrices of all \( s \in S_n \).

**Example 4.19** (The standard representation in orthogonal form). Continuing with the notation of Example 4.17, let us now write \( w(j) = w_T^j \). Using the values \( d_{T,j,i} \) from Example 4.17, the action formula in Theorem 4.18 becomes

\[
(\text{4.35}) \\
s_i.w(j) = \begin{cases} 
  w(j) & \text{for } j \neq i, i + 1 \\
  \frac{1}{2}w(i) + \sqrt{1 - \frac{1}{i^2}}w(i + 1) & \text{for } j = i \\
  \sqrt{1 - \frac{1}{i^2}}w(i) - \frac{1}{i}w(i + 1) & \text{for } j = i + 1
\end{cases}
\]
Exercises for Section 4.5

4.5.1 (Realization over the integers). Let $K/\mathbb{Q}$ be a finite field extension and assume that $O = \{ \alpha \in K \mid \alpha \text{ is integral over } \mathbb{Z} \}$ is a PID. Show that every finite subgroup of $\text{GL}_n(K)$ is conjugate in $\text{GL}_n(K)$ to a subgroup of $\text{GL}_n(O)$. (Use the structure theorem of modules over PIDs; e.g., Jacobson \[111\], Section 3.8.) In particular, every finite subgroup of $\text{GL}_n(\mathbb{Q})$ is conjugate to a subgroup of $\text{GL}_n(\mathbb{Z})$.

4.5.2 (Bases of the standard representation $V_{n-1}$). (a) What is the relationship between the GZ-bases of $V_{n-1} = V(n-1,1)$ that are exhibited in Examples 4.3, 4.17 and 4.19? Note that the bases in Examples 4.17 and 4.19 are only determined up to a common scalar factor because of the choices of $v(\lambda)$ and of an invariant inner product in Theorems 4.15 and 4.18.

(b) Using the notation of (3.19), put $a_i = b_i - b_{i+1} (i = 1, \ldots, n-1)$. Show that $(a_1, \ldots, a_{n-1})$ is a basis of $V_{n-1}$ such that the resulting matrices of all $s \in S_n$ have integer entries.

4.6. The Murnaghan-Nakayama Rule

The Murnaghan-Nakayama Rule is an efficient method for determining the irreducible characters $\chi^\lambda = \chi_{V^\lambda}$ of $S_n$. Before presenting its statement and proof, we need to extend some earlier material from flag-shaped partitions $\lambda$ to more general shapes $\lambda/\mu$, called skew shapes.

4.6.1. Skew Shapes

For given partitions $\lambda, \mu \in \mathcal{P}$, let us write $\mu \leq \lambda$ if $\mu_i \leq \lambda_i$ for all $i$ (as in §4.3.2). Thus, $\mu \leq \lambda$ means that the Young diagram of $\lambda$ is obtained by adding boxes to the diagram of $\mu$ or, equivalently, there is a path $\mu \rightarrow \cdots \rightarrow \lambda$ in $\mathcal{Y}$. In this case, the skew shape $\lambda/\mu$ is defined to be the shape that is obtained by removing the boxes belonging to $\mu$ from $\lambda$. The notation and basic material developed earlier for partitions extend rather easily to skew shapes $\lambda/\mu$, partitions being the special case where $\mu = \emptyset$.

Paths and Skew Tableaux

We will write $|\lambda/\mu| = |\lambda| - |\mu|$ for the number of boxes of $\lambda/\mu$. A $\lambda/\mu$-tableau is a filling of the boxes of $\lambda/\mu$ with the numbers $1, 2, \ldots, |\lambda/\mu|$ in such a way that the numbers increase from left to right along rows and from top to bottom along columns. Generalizing the definition of $f^\lambda$ in (4.15), (4.17), we put

$$f^{\lambda/\mu} \overset{\text{def}}{=} \# \{ \lambda/\mu \text{-tableaux} \}$$
Successively adding boxes to $\mu$ in the order specified by the sequence of numbers in a given $\lambda/\mu$-tableau, we obtain a path $\mu \to \cdots \to \lambda$ in $Y$. Exactly as in (4.16), this yields a bijection, which we will again often treat as an identification:

\[(4.37) \quad \{ \text{paths } \mu \to \cdots \to \lambda \text{ in } Y \} \leftrightarrow \{ \lambda/\mu\text{-tableaux} \} \]

In fact, the bijection (4.37) can be derived from (4.16). To see this, fix a path $T_\mu: \emptyset \to \cdots \to \mu$ in $Y$ or, equivalently, a $\mu$-tableau. Evidently, the paths $\mu \to \cdots \to \lambda$ in $Y$ are in bijection with the collection of paths $\emptyset \to \cdots \to \lambda$ starting with $T_\mu$. The $\lambda$-tableaux arising in this manner are clearly those with $1, \ldots, |\mu|$ filled into the boxes occupied by $\mu$ according to $T_\mu$. In order to obtain a bijection of these $\lambda$-tableaux with the set of all $\lambda/\mu$-tableaux, start with a $\lambda/\mu$-tableau $T$, replace the number $i$ in each of its boxes by $|\mu| + i$, and then fill in the missing $\mu$-region northeast of the skew shape $\lambda/\mu$ with the $\mu$-tableau $T_\mu$. This yields a $\lambda$-tableau, which we shall denote by $T_\mu \sqcup T$.

**Example 4.20.** For $\lambda = (6, 4, 1)$ and $\mu = (3, 2)$, we obtain $\lambda/\mu = \begin{array}{ccc} & & 1 \\ & 2 & 3 \\ 4 & 5 & \end{array}$.

Furthermore, for example, \begin{array}{ccc} 2 & 4 & 6 \\ 1 & 3 & 5 \\ \end{array} \sqcup \begin{array}{ccc} 2 & 4 & 6 \\ 1 & 3 & 5 \\ \end{array} = \begin{array}{ccc} 2 & 4 & 6 \\ 1 & 3 & 5 \\ 8 & 9 & 10 \\ \end{array}.

In summary, the bijection (4.37) fits into the following commutative diagram of injections and bijections:

\[(4.38) \quad \begin{array}{ccc} \{ \text{paths } \mu \to \cdots \to \lambda \text{ in } Y \} & \xrightarrow{\text{start with } T_\mu} & \{ \text{paths } \emptyset \to \cdots \to \lambda \text{ in } Y \} \\
(4.37) & \downarrow \cong & \downarrow \cong \\
\{ \lambda/\mu\text{-tableaux} \} & \xleftarrow{T_\mu \sqcup} & \{ \lambda\text{-tableaux} \} \end{array} \]

**Content**

Let $T$ be a $\lambda/\mu$-tableau with $m = |\lambda/\mu|$ boxes. As for ordinary $\lambda$-tableaux, we define the **content** of $T$ to be the $m$-tuple $c_T = (c_{T,1}, c_{T,2}, \ldots, c_{T,m})$, where $c_{T,i} = a$ means that the box of $T$ containing the number $i$ lies on the $a$-diagonal $y = x + a$. The earlier argument applies verbatim to show that, when $\mu$ is given, the content $c_T$ determines $T$. In fact, one can also simply quote the fact for ordinary $\lambda$-tableaux to see this: just note that, evidently,

\[(4.39) \quad c_{T_\mu \sqcup T} = (c_{T_\mu}, c_T), \]

where $(c_{T_\mu}, c_T)$ denotes the concatenation of the content vectors $c_{T_\mu}$ and $c_T$. Thus, for any given $\mu$-tableau $T_\mu$, the content $c_T$ determines the $\lambda$-tableau $T_\mu \sqcup T$, and hence $T$ is determined by (4.38). Note also that $c_{T,i+1} \neq c_{T,i}$ for all $i$—this is easy to see directly, and it has already been observed for $\lambda$-tableaux and consequently
also follows from (4.39). Therefore, as in (4.33), we may define
\[ d_{T,i} := (c_{T,i+1} - c_{T,i})^{-1}. \]

### Admissible Transpositions

For any \( \lambda/\mu \)-tableau \( T \) and any \( s \in S_m \) with \( m = |\lambda/\mu| \), we may consider the tableau \( sT \), not necessarily a \( \lambda/\mu \)-tableau, that arises after replacing the entries in all boxes of \( T \) by their images under \( s \). If \( sT \) is in fact a \( \lambda/\mu \)-tableau again, then the transposition \( s_i = (i, i + 1) \in S_m \) is called an **admissible transposition** for \( T \). If follows from (4.25) that
\[ s_i \text{ is admissible for } T \iff d_{T,i} \neq \pm 1. \]

To see this, fix some \( \mu \)-tableau \( T_\mu \) and put \( k = |\mu| \). Clearly, \( s_i \) is admissible for \( T \) if and only if \( s_i' = (k + i, k + i + 1) \) is admissible for \( T_\mu \sqcup T \), which in turn is equivalent to \( d_{T_\mu \sqcup T,k+i} = d_{T,i} \).

#### 4.6.2. Representations Associated to Skew Shapes

Let us start over with given partitions \( \lambda \vdash n \) and \( \mu \vdash k \) and consider the representations \( V^\lambda \in \textup{Irr}S_n \) and \( V^\mu \in \textup{Irr}S_k \). By the branching rule (4.13), \( \mu \leq \lambda \) if and only if \( S_k \leq S_n \) and there exists an embedding \( V^\mu \hookrightarrow V^\lambda \). Equivalently, the following vector space is nonzero:

\[ V^{\lambda/\mu} \overset{\text{def}}{=} \text{Hom}_{S_k}(V^\mu, V^\lambda|_{S_k}) \]

From now on, let us assume that \( \mu \leq \lambda \) and put \( m = |\lambda/\mu| = n - k \). By Frobenius reciprocity (3.8), we also have \( V^{\lambda/\mu} \cong \text{Hom}_{S_n}(V^{\mu\uparrow S_n}, V^\lambda) \).

#### Dimension

By Proposition 1.32, \( \dim V^{\lambda/\mu} = m(V^\mu, V^\lambda|_{S_k}) \), the length of the \( V^\mu \)-homogeneous component of \( V^\lambda|_{S_k} \). This number is equal to \( f^{\lambda/\mu} = \# \{ \text{ paths } \mu \to \cdots \to \lambda \in \mathcal{Y} \} \); see (4.36), (4.37). Indeed, the equality is clear if \( k = n \); so assume that \( k < n \). Since \( V^\lambda|_{S_{n,k-1}} \cong \bigoplus_{\nu \rightarrow \lambda} V^\nu \) by (4.13), the length of the \( V^\mu \)-homogeneous component of \( V^\lambda|_{S_k} \) is equal to the sum of the lengths of the \( V^\mu \)-homogeneous components of all \( V^\nu|_{S_k} \). The assertion now follows by induction on \( m = n - k \) from the obvious equality \( f^{\lambda/\mu} = \sum_{\nu \rightarrow \lambda} f^{\nu/\mu} \). To summarize,
\[ \dim V^{\lambda/\mu} = f^{\lambda/\mu}. \]

#### Irreducibility

The following general lemma shows that \( V^{\lambda/\mu} \) is an irreducible representation of the centralizer algebra \( \mathbb{C}[S_n] \) (Theorem 4.1). The lemma does not need \( \text{char } \mathbb{C} = 0 \), but we continue to assume that \( \mathbb{C} \) is algebraically closed.
Lemma 4.21. Let $A \subseteq B$ be finite-dimensional semisimple $k$-algebras and let $W \in \text{Irr} B$ and $V \in \text{Irr} A$ be such that $H := \text{Hom}_A(V, W_\downarrow_A) \neq 0$. Then the centralizer algebra $C = C_P(A)$ acts on $H$ via $c \phi = c_W \circ \phi$ for $c \in C$ and $\phi \in H$ and $H$ is irreducible for this action.

Proof. The given action is clearly well defined, but we need to prove irreducibility. Using the Wedderburn isomorphism (1.46), $B \cong \prod_{S \in \text{Irr} B} \text{End}_k(S)$, we may replace $A, B$ and $C$ by their images in $\text{End}_k(W)$. (Note that centralizing in direct products is a componentwise condition and semisimplicity is preserved under taking isomorphic images.) Thus, $B = \text{End}_k(W)$ and $C = \text{End}_A(W)$. By assumption, the $V$-homogeneous component, $W(V)$, of $W_\downarrow_A$ is nonzero; so $W_\downarrow_A = W(V) \oplus X$ and $W(V) \cong V^\oplus m$ for some positive integer $m$. Therefore, $H \cong \text{Hom}_A(V, W(V)) \cong k^\oplus m$ and $\text{End}_A(W(V)) \cong \text{Mat}_m(k)$ (Lemma 1.4). Finally, note that $\text{End}_A(W(V))$ is naturally contained in $C = \text{End}_A(W) \cong \text{End}_A(W(V)) \times \text{End}_A(X)$ and that $H$ is irreducible under the action of $\text{End}_A(W(V))$ (Example 1.12). 

Basis and Actions

The $(kS_n)^{S_k}$-action in Lemma 4.21 makes $V^{\lambda/\mu}$ a (not necessarily irreducible) representation of the subgroup $S'_m \leq S_n$ consisting of the permutations that are the identity on $[k]$. We shall regard $V^{\lambda/\mu}$ as a representation of $S'_m$ via the isomorphism $\cdot': S_m \widetilde{\to} S'_m$ that comes from the bijection $[n] \setminus [k] \cong [m]$, $k+i \leftrightarrow i$. Thus, the transposition $s_i = (i, i+1) \in S_m$ acts as $s'_i = (i+k, k+i+1) \in S'_m$ on $V^{\lambda/\mu}$.

The following result extends Young’s orthogonal form (Theorem 4.18) to skew shapes. The action formulae for the Coxeter generators $s_i$ in part (a) make it clear that the structure of $V^{\lambda/\mu}$ as $S'_m$-representation only depends on the geometric shape of $\lambda/\mu$, not on the particular realization with $\lambda$ and $\mu$.

Theorem 4.22. Let $\lambda/\mu$ be a skew shape with $|\lambda/\mu| = m$ and $|\mu| = k$. Then there exists a basis $(w_T)_T$ a $\lambda/\mu$-tableau of $V^{\lambda/\mu}$ having the following properties:

(a) The action of the transposition $s_i = (i, i+1) \in S_m$ is given by

$$s_i.w_T = d_{T,i} w_T + \sqrt{1 - d^2_{T,i}} w_{s_iT} .$$

(Recall (4.40); if $d_{T,i} = \pm 1$, then the last term is missing.)

(b) Each $w_T$ is an eigenvector for the JM-elements $X_{k+j}$ ($j = 1, \ldots, m$) with eigenvalue $c_{T,j}$, and $V^{\lambda/\mu} = kS_m.w_T$.

Proof. (a) If $\mu = \emptyset$, then Theorem 4.18 provides us with a GZ-basis $(w_T)$ of $V^{\lambda/\mu} = V^\lambda$ such that the action of $s_i$ is given by the formula in the theorem. Moreover, any GZ-basis consists of eigenvectors for the GZ-algebra $G_Z$, with the eigenvalues of the JM-elements given by contents (Theorem 4.12 and (4.28)):

$$X_j.w_T = c_{T,j} w_T.$$
Now let $\mu \vdash k$ be arbitrary. For every $\lambda/\mu$-tableau $T$, consider the linear operator
\begin{equation}
 w_T : V^\mu \rightarrow V^\lambda
\end{equation}
(4.44)
where $T_\mu$ runs over all standard $\mu$-tableaux and $(w_{T_\mu})$ is the GZ-basis of $V^\mu$ as in Theorem 4.18. The images $w_{T_\mu \sqcup T}$ are part of the GZ-basis of $V^\lambda$ in Theorem 4.18. Therefore, the various maps $(\mu$-tableaux) $\rightarrow \{\lambda$-tableaux$\}$ that are given by $\cdot \sqcup T$ and $\cdot \sqcup T'$ have disjoint images. Note the obvious identity
\begin{equation}
 s't'(T_\mu \sqcup T) = (sT_\mu) \sqcup (tT) \quad (s \in S_k, t \in S_m),
\end{equation}
where $\cdot ' : S_m \rightarrow S_m'$ is the above isomorphism. Therefore, for $i = 1, \ldots, k - 1$, $s_i \cdot w_T (w_{T_\mu}) = s_i \cdot w_{T_\mu \sqcup T}$
\begin{align*}
 &\overset{\text{Theorem 4.18}}{=} d_{T_\mu \sqcup T, i} w_{T_\mu \sqcup T} + \sqrt{1 - d_{T_\mu \sqcup T, i}^2} w_{s_i(T_\mu \sqcup T)} \\
 &\overset{(4.39),(4.45)}{=} d_{T_\mu, i} w_{T_\mu} + \sqrt{1 - d_{T_\mu, i}^2} w_{s_i(T_\mu)} \\
 &\overset{\text{Theorem 4.18}}{=} w_T(s_i \cdot w_{T_\mu}).
\end{align*}
This shows that $w_T \in V^{A/\mu}$. Consequently, $(w_T)$ is a basis of $V^{A/\mu}$ by (4.42). Each $s_i \in S_m$ acts on $V^{A/\mu}$ via $(s_i \cdot w_T)(w_{T_\mu}) = s'_i \cdot w_T(w_{T_\mu})$. Thus,
\begin{align*}
 (s_i \cdot w_T)(w_{T_\mu}) &= s'_i \cdot w_{T_\mu \sqcup T} \\
 &\overset{\text{Theorem 4.18}}{=} d_{T_\mu \sqcup T, k+i} w_{T_\mu \sqcup T} + \sqrt{1 - d_{T_\mu \sqcup T, k+i}^2} w_{s'_i(T_\mu \sqcup T)} \\
 &\overset{(4.39),(4.45)}{=} d_{T, i} w_{T_\mu} + \sqrt{1 - d_{T, i}^2} w_{s_i(T)} \\
 &= (d_{T, i} w_T + \sqrt{1 - d_{T, i}^2} w_{s_i(T)})(w_{T_\mu}).
\end{align*}
proving the action formula for the Coxeter generators.

(b) The formula for the action of the JM-elements follows from the computation
\begin{equation}
 (X_{k+j} \cdot w_T)(w_{T_\mu}) = X_{k+j} \cdot w_{T_\mu \sqcup T} \overset{(4.43)}{=} c_{T_\mu \sqcup T, k+j} w_{T_\mu \sqcup T} \overset{(4.39)}{=} c_{T, j} w_{T_\mu \sqcup T} \overset{(4.39)}{=} c_{T, j} w_T(w_{T_\mu}).
\end{equation}
It remains to show that $V^{A/\mu} = \mathbb{K} S_m \cdot w_T$ for each $T$. But $V^{A/\mu}$ is an irreducible representation of the centralizer algebra $(\mathbb{K} S_m)^{S_k} = \mathbb{K}[Z_k, S_m, X_{k+1}, \ldots, X_n]$ by
Lemma 4.21 and we have also seen that $w_T$ is an eigenvector for the JM-elements $X_{k+1}, \ldots, X_n$. Therefore,

$$V^{λ/μ} = k[Z_k, S_m', X_{k+1}, \ldots, X_n], \quad w_T = k[Z_k, S_m']. \quad (\lambda)$$

Furthermore, $Z_k$ acts by scalar multiplications on $V^{μ} \in \text{Irr } S_k$ (Schur’s Lemma), and so $Z_k$ acts as scalar operators on $V^{λ/μ} = \text{Hom}_{S_k}(V^{μ}, V^{λ/μ}_{S_k})$ as well. Therefore, $V^{λ/μ} = k[S_m']. w_T = kS_m.w_T$, completing the proof of the theorem. □

4.6.3. The Murnaghan-Nakayama Rule: Statement and Examples

In this subsection, we determine the character of the representation $V^{λ/μ}$ of $S_m$ that is associated to the skew shape $λ/μ$ with $|λ/μ| = m$; this character will be denoted by $χ^{λ/μ} = χ_{V^{λ/μ}}$. We will also write $χ^{λ/μ}_α$ for the value of $χ^{λ/μ}$ on the conjugacy class of $S_m$ corresponding to the partition $α ⊢ m$ (§3.5.2). If $μ = φ$, then we will simply write $χ^{λ/μ}$ and $χ^{λ/μ}_α$; so the matrix $(χ^{λ/μ}_α, λ, α) ⊢ m$ is the character table of $S_m$.

We need a few more definitions. For each skew shape $λ/μ$, we put $h(λ/μ) \overset{\text{def}}{=} \#\{\text{rows occupied by the diagram of } λ/μ\} - 1$.

The skew shape $λ/μ$ is said to be connected if the interior of its diagram is connected in the usual sense or, equivalently, the integers $a$ such that $λ/μ$ has boxes on the $a$-diagonal $y = x + a$ form an interval in $\mathbb{Z}$. (The skew shapes $λ/μ$ is the picture in §4.6.1 and in Example 4.20 are disconnected.) A connected skew shape $λ/μ$ is called a skew hook if there is at most one box on all diagonals, that is, $λ/μ$ has “thickness” 1 throughout its extent. The number of boxes in a skew hook will be referred to as its length. A skew hook of length $r$ will also be called an $r$-hook.

Murnaghan-Nakayama Rule. Let $λ/μ$ be a skew shape with $|λ/μ| = m$ and let $α = (α_1 ≥ α_2 ≥ \cdots ≥ α_l > 0) ⊢ m$. Then

$$χ^{λ/μ}_α = \sum_\Lambda (-1)^{h(\Lambda)}$$

Here, $\Lambda$ runs over all sequences of partitions $\Lambda = (μ = λ^0 < λ^1 < \cdots < λ^l = λ)$, where $λ^i/λ^{i-1}$ is an $α_i$-hook and $h(\Lambda) = \sum_i h(λ^i/λ^{i-1})$.

The proof will be given in §4.6.4. Here, we just offer some comments and examples.
The northwest and southeast rims of any skew shape \( \lambda/\mu \) are skew hooks having the same length, because the rims can be projected onto each other by moving boxes along diagonals. In this way, each of the skew hooks \( \lambda^i/\lambda^{i-1} \) in the Murnaghan-Nakayama Rule projects into either of the rims. Thus, if the given partition \( \alpha \) has a part \( \alpha_i \) that is larger than the length of the rims, then there is no sequence \( \Lambda \) satisfying the requirement of the rule and so we must have \( \chi^{\lambda/\mu}_\alpha = 0 \) in this case.

Also, for \( \alpha = (m) \), there are no sequences \( \Lambda \) as in the Murnaghan-Nakayama Rule unless \( \lambda/\mu \) is a skew hook, in which case \( \Lambda = (\mu = \lambda^0 < \lambda^1 = \lambda) \) is the unique sequence. Therefore, for any skew shape \( \lambda/\mu \) with \( |\lambda/\mu| = m \), we obtain

\[
\chi^{\lambda/\mu}_{(m)} = \begin{cases} (-1)^h(\lambda/\mu) & \text{if } \lambda/\mu \text{ is a skew hook} \\ 0 & \text{otherwise.} \end{cases}
\]

(4.46)

Let us now discuss some concrete examples focusing on the case where \( \mu = \emptyset \). For further examples, see Exercise 4.6.1.

**Example 4.23.** Let us evaluate \( \chi^{\lambda}_\alpha \) for \( \lambda = (5, 4, 3) \) and \( \alpha = (6, 3, 2, 1) \); so \( m = 12 \) and \( l = 4 \). All sequences \( \Lambda \) in the Murnaghan-Nakayama Rule must start with a partition \( \lambda^1 \) that is also a skew hook of length 6; hence \( \lambda^1 \) is in fact an ordinary hook. The second skew hook, \( \lambda^2/\lambda^1 \), has length 3 and it must be attached to the southeast rim of \( \lambda^1 \) in such a way that the resulting shape represents a partition. Continuing in this way with the remaining two skew hooks, we must fill up the shape of \( \lambda \). The possible sequences \( \Lambda \) arising in this way are depicted below, with the \( i^{th} \) skew hook \( \lambda^i/\lambda^{i-1} \) labeled by \( i \).

![Sequence Diagram](image)

The \( h \)-values of these sequences are \( 1 + 1 + 1 + 0, 1 + 1 + 0 + 0, 1 + 0 + 0 + 0 \) and \( 2 + 0 + 0 + 0 \). Thus, \( \chi^{(5, 4, 3)}_{(6, 3, 2, 1)} = (-1)^3 + (-1)^2 + (-1)^1 + (-1)^0 = 0 \).

**Example 4.24.** Now let \( \lambda = (m - k, 1^k) \) be a hook partition with \( k \geq 1 \). We already know from Example 4.14(b) that \( V^\lambda = \Lambda^k V^m_{m-1} \) is the \( k^{th} \) exterior power of the standard representation of \( S_m \). Hence \( \chi^\lambda \) can be determined from the known character of \( V^m_{m-1} \) and the Newton identities; see Example 3.13(d) and (3.63). Nevertheless, let us also comment on the values \( \chi^\lambda_{\alpha} \) from the point of view of the Murnaghan-Nakayama Rule. If \( \alpha_i > k \) for all \( i \), then the only possible sequence \( \Lambda \) looks like this:

![Possible Sequence](image)
The $h$-value of this sequence is $k$. Thus, putting $m_\alpha(j) = \#\{i \mid \alpha_i = j\}$ as in §3.5.2, we have $\chi_\alpha^j = (-1)^k$ in case $m_\alpha(j) = 0$ for all $j \leq k$. If $\alpha_1 > k$ and $m_\alpha(k) \neq 0$ but $m_\alpha(j) = 0$ for $j < k$, then we also have the sequences

for all $i$ with $\alpha_i = k$. Since the $h$-value of these sequences is $k - 1$, we obtain $\chi_\alpha^j = (-1)^k + m_\alpha(k)(-1)^{k-1} = (-1)^{k-1}(m_\alpha(k) - 1)$ in this case.

Instead of building up the sequences $\Lambda$ in the Murnaghan-Nakayama Rule by attaching skew hooks as in the examples above, we may alternatively proceed by successively peeling off skew hooks from the southeast rims of $\lambda$ and of all subsequent partitions, always taking care that the requisite flag shape is preserved at each step. In this connection, it is useful to observe that the sum on the right in the Murnaghan-Nakayama Rule certainly makes sense if the parts $\alpha_i$ of $\alpha$ are not necessarily weakly decreasing. In fact, though it may not a priori be obvious, the proof will show that the sum takes the same value for all arrangements of the order of the parts $\alpha_i$; see the remarks at the end of Step 1 in §4.6.4. It is often useful to start the computation of $\chi_\alpha^j$ by removing $\alpha_i$-hooks from $\lambda$, where $\alpha_i$ is not necessarily the last part of $\alpha$. The remainder of the computation then determines the various $\chi_\alpha^{\lambda/\mu}$, where $\alpha \setminus \alpha_i$ is the partition $\alpha$ without the part $\alpha_i$ and $\mu$ runs over all partitions obtained by peeling an $\alpha_i$-hook from $\lambda$. Thus we have the following corollary, which highlights the inductive nature of the process; it reduces the calculation of $\chi_\alpha^j$ to characters of smaller symmetric groups and eventually to the trivial group $S_\infty$. The corollary is also referred to as the Murnaghan-Nakayama Rule. Note that the case $r = 1$ amounts to the branching rule.

**Corollary 4.25.** Let $\lambda \vdash m$ and let $s, c \in S_m$, where $c$ is an $r$-cycle and $s$ is a permutation of the remaining $m - r$ elements of $[m]$. Then

$$\chi_\lambda^{sc} = \sum_\mu (-1)^{h(\lambda/\mu)} \chi_\mu(s)$$

where $\mu$ runs over all partitions $\leq \lambda$ such that $\lambda/\mu$ is an $r$-hook.

**Example 4.26.** Let $\lambda = (5, 4, 3)$ be as in Example 4.23 but now let us consider the rearrangement $(1, 2, 3, 6)$ of $\alpha = (6, 3, 2, 1)$. There is only one way to remove a 6-hook from $\lambda$ so that the result is a partition; then there are two ways to remove a 3-hook from this partition etc. Thus the only possible sequences of partitions $\Lambda = (\lambda^1 < \cdots < \lambda^4 = \lambda)$ with $\lambda^1$ a 1-hook, $\lambda^2/\lambda^1$ a 2-hook, $\lambda^3/\lambda^2$ a 3-hook and $\lambda^4/\lambda^3$ a 6-hook are as follows:
Since the $h$-values for these two sequences are $2 + 1 + 1 + 0$ and $2 + 1 + 0 + 0$, respectively, we obtain $(-1)^3 + (-1)^3 = 0$ for the value of the sum in the Murnaghan-Nakayama Rule, which agrees with what we had found in Example 4.23. The calculation of $\chi^{(5,4,3)}_{(6,3,2,1)}$ from the viewpoint of Corollary 4.25 proceeds like this:

$$
\chi^{(5,4,3)}_{(6,3,2,1)} = (-1)^2 \chi^{(3,2,1)}_{(3,2,1)} = \chi^{(3,2,1)}_{(3,2,1)}
= (-1) \chi^{(1^3)} + (-1) \chi^{(2,1)}
= (-1)(-1)\chi^{(1)}_{(1)} + (-1)\chi^{(1)}_{(1)} = 0.
$$

4.6.4. Proof of the Murnaghan-Nakayama Rule

**Step 1: reduction to formula (4.46).** Let $\lambda/\mu$ be a skew shape with $|\lambda/\mu| = m$. We first describe the restriction of $V^{\lambda/\mu} \in \text{Rep} S_m$ to Young subgroups of $S_m$. Recall that Young subgroups are associated to compositions of $m$, that is, sequences $\mathbf{m} = (m_1, m_2, \ldots, m_k)$ of positive integers with $|\mathbf{m}| = \sum m_j = m$; the Young subgroup $S_{\mathbf{m}} \leq S_m$ consists of all $s \in S_m$ having the form $s = g_1 g_2 \cdots g_l$, where the factor $g_i$ only permutes the elements $k_{i-1} + 1, \ldots, k_i$ with $k_i = \sum_{j<i} m_j$.

Thus, $S_{\mathbf{m}} \cong S_{m_1} \times S_{m_2} \times \cdots \times S_{m_k}$. For given representations $V_i \in \text{Rep} S_{m_i}$, we may form the outer tensor product $V = V_1 \otimes V_2 \otimes \cdots \otimes V_l$ in $\text{Rep} S_m$ ($\S 1.5.1$): $V \cong \bigotimes_{i=1}^l V_i$ as $k$-vector spaces and $S_m$ acts by

$$g_1 g_2 \cdots g_l \cdot (v_1 \otimes v_2 \otimes \cdots \otimes v_l) = g_1 v_1 \otimes g_2 v_2 \otimes \cdots \otimes g_l v_l.$$

**Lemma 4.27.** Let $\lambda/\mu$ be a skew shape and let $\mathbf{m} = (m_1, m_2, \ldots, m_k)$ be a composition of $m = |\lambda/\mu|$. Then

$$
V^{\lambda/\mu} \downarrow_{S_{\mathbf{m}}} \cong \bigoplus_{\Lambda} \left( V^{\lambda^0/\lambda^0} \otimes V^{\lambda^1/\lambda^1} \otimes \cdots \otimes V^{\lambda^k/\lambda^k} \right),
$$

where $\Lambda$ runs over all sequences of partitions $\Lambda = (\mu = \lambda^0 < \lambda^1 < \cdots < \lambda^k = \lambda)$ with $|\lambda^i/\lambda^{i-1}| = m_i$. Moreover, for each such $\Lambda$, there is an epimorphism

$$
\left( V^{\lambda^0/\lambda^0} \otimes V^{\lambda^1/\lambda^1} \otimes \cdots \otimes V^{\lambda^k/\lambda^k} \right) \bigg|_{S_{\mathbf{m}}} \rightarrow V^{\lambda/\mu}.
$$

**Proof.** The basis $(w_T)_{T a \lambda/\mu}$-tableau of $V^{\lambda/\mu}$ in Theorem 4.22 is in bijection with the collection of all paths $\lambda \rightarrow \cdots \rightarrow \mu$ in $Y$ by (4.37). Each such path has $m$ arrows; so we may split it into an initial subpath $\mu \rightarrow \cdots \rightarrow \lambda^1$ given by the first $m_1$ arrows, then the path $\lambda^1 \rightarrow \cdots \rightarrow \lambda^2$ given by the next $m_2$ arrows and so on. This amounts to writing the corresponding $\lambda/\mu$-tableau $T$ as $T = T^1 \sqcup T^2 \sqcup \cdots \sqcup T^k$, where each $T^i$ is a $\lambda^i/\lambda^{i-1}$-tableau. By (4.44), the basis vector $w_T \in V^{\lambda/\mu}$ is the composite

$$w_T = \left( V^{\mu} \xrightarrow{w_{T^1}} V^{\lambda^1} \xrightarrow{w_{T^2}} V^{\lambda^2} \xrightarrow{w_{T^3}} \cdots \xrightarrow{w_{T^k}} V^{\lambda^k} \right).$$
In this way, the basis \((w_T)\) of \(V^{\Lambda/\mu}\) becomes the disjoint union of parts labeled by the sequences \(\Lambda\) in the lemma, the \(\Lambda\)-part being in bijection with the product of the bases \((w_T^i)_{T^i} \lambda^i/\lambda^{i-1}\)-tableau of the various \(\lambda^i/\lambda^{i-1}\). An inspection of the action of the Coxeter generators \(s_j \in S_m\) in Theorem 4.22 reveals that if \(s_j\) belongs to the factor \(S_{m_i}\) of \(S_m\), then \(s_j\) only affects the \(i\)th map \(w_T\) in the above composite, acting via the usual \(S_{m_i}\)-action on \(V^{\lambda^i/\lambda^{i-1}}\). Since \(S_m\) is generated by the Coxeter generators that belong to one of the factors \(S_{m_i}\), the assertion about the structure of \(V^{\Lambda/\mu}_{\downarrow S_m}\) follows.

Finally, for each \(\Lambda = (\mu = \lambda^0 < \lambda^1 < \cdots < \lambda^l = \lambda)\), the foregoing gives an embedding \(V^{\lambda^1/\mu} \otimes V^{\lambda^2/\mu} \otimes \cdots \otimes V^{\lambda^l/\mu} \hookrightarrow V^{\Lambda/\mu}_{\downarrow S_m}\). The image contains all members of the basis \((w_T)\) of \(V^{\lambda^i/\mu}\) such that the path \(T: \mu \to \cdots \to \lambda\) passes through \(\lambda^1, \lambda^2, \ldots\). By Frobenius reciprocity (3.8), the embedding corresponds to a map \((V^{\lambda^1/\mu} \otimes V^{\lambda^2/\mu} \otimes \cdots \otimes V^{\lambda^l/\mu})_{\uparrow S_m} \to V^{\Lambda/\mu}\). Since each \(w_T\) generates \(V^{\Lambda/\mu}\) (Theorem 4.22), this map is onto. 

With the lemma in hand, we now deduce the Murnaghan-Nakayama Rule from the special case (4.46). To this end, let \(s \in S_m\) be given and write \(s\) as a product of disjoint cycles, say \(s = g_1 g_2 \cdots g_l\) with \(g_i\) an \(m_i\)-cycle. Replacing \(s\) by a conjugate, we may assume that \(s \in S_m\) with \(m = (m_1, \ldots, m_l)\) and \(g_i \in S_{m_i}\). It follows from Lemma 4.27 in conjunction with (1.52) that

\[
\chi^{\Lambda/\mu}(s) = \sum_{\Lambda} \prod_i \chi^{\lambda^i/\lambda^{i-1}}(g_i).
\]

By (4.46) we further know that \(\chi^{\lambda^i/\lambda^{i-1}}(g_i) = (-1)^{h(\lambda^i/\lambda^{i-1})}\) if \(\lambda^i/\lambda^{i-1}\) is a skew hook and \(\chi^{\lambda^i/\lambda^{i-1}}(g_i) = 0\) otherwise. Therefore, only sequences of partitions \(\Lambda = (\mu = \lambda^0 < \lambda^1 < \cdots < \lambda^l = \lambda)\) with \(\lambda^i/\lambda^{i-1}\) an \(m_i\)-hook contribute to the sum, and

\[
\chi^{\Lambda/\mu}(s) = \sum_{\Lambda} (-1)^{h(\Lambda)} \quad \text{with} \quad h(\Lambda) = \sum_i h(\lambda^i/\lambda^{i-1}).
\]

This is the formula in the Murnaghan-Nakayama Rule. Since partitions are the same as non-increasing compositions, the foregoing also justifies our remarks in the previous section concerning rearrangements of the given partition \(\alpha\). It remains to prove (4.46).

**Step 2: proof of (4.46) for \(\lambda/\mu\) disconnected.** Let \(\lambda/\mu\) be a disconnected skew shape. Then \(\lambda/\mu\) is certainly not a skew hook and so (4.46) asserts that \(\chi^{\lambda/\mu}(s) = 0\) for every \(m\)-cycle \(s \in S_m\), where \(m = |\lambda/\mu|\). In order to prove this, we will show that \(V^{\lambda/\mu}\) is induced from a representation of some proper Young subgroup of \(S_m\). Since no such subgroup can contain a conjugate of \(s\), the desired equality \(\chi^{\lambda/\mu}(s) = 0\) holds.
0 then follows from a standard fact about characters of induced representations (Exercise 3.1.3).

In order to show that $V^{\lambda/\mu}$ is induced, write $\lambda/\mu$ as the union of two skew shapes $\lambda_1/\mu$ and $\lambda_2/\mu$ occupying disjoint columns and rows. Put $|\lambda_1/\mu| = m_1$ and consider the Young subgroup $S_m \equiv S_{m_1} \times S_{m_2} \leq S_m$ for the composition $m = (m_1, m_2)$ of $m$. Since $\lambda/\lambda_1 = \lambda_2/\mu$, we have an epimorphism $(\psi_1 : V^{\lambda_1/\mu} \boxtimes V^{\lambda_2/\mu}) \hookrightarrow V^{\lambda/\mu}$ (Lemma 4.27).

It suffices to show that both spaces have the same dimension. By (4.42) and (4.36), we know that $\dim V^{\lambda_1/\mu} = f_{\lambda_1/\mu}$, the number of $\lambda_1/\mu$-tableaux, while

$$\dim \left( V^{\lambda_1/\mu} \boxtimes V^{\lambda_2/\mu} \right)_{S_m} = \left[ S_m : S_m \right] (\dim V^{\lambda_1/\mu})(\dim V^{\lambda_2/\mu}) = \left( \begin{array}{c} m_1 \\ m_1 \end{array} \right) f_{\lambda_1/\mu} f_{\lambda_2/\mu}.$$  

But, to create a $\lambda/\mu$-tableau, we must choose $m_1$ elements from $[m]$ and fill them into the boxes of the skew shape $\lambda_1/\mu$ in increasing order along rows and columns. This can be done in $f_{\lambda_1/\mu}$ many ways. The remaining numbers must then be filled in the same manner into $\lambda_2/\mu$, for which there are $f_{\lambda_2/\mu}$ possibilities. Since $\lambda_1/\mu$ and $\lambda_2/\mu$ share no columns or rows, this will always result in a $\lambda/\mu$-tableau. Consequently, $f_{\lambda/\mu} = \left( \begin{array}{c} m_1 \\ m_1 \end{array} \right) f_{\lambda_1/\mu} f_{\lambda_2/\mu}$, as desired. Therefore,

(4.47)  

$$V^{\lambda/\mu} \cong \left( V^{\lambda_1/\mu} \boxtimes V^{\lambda_2/\mu} \right)_{S_m}$$

This completes Step 2. Henceforth, we may assume $\lambda/\mu$ to be connected.

**Step 3: proof of (4.46) for $\mu = \emptyset$.** We need to show that, for a partition $\lambda \vdash m$ and an $m$-cycle $s \in S_m$,

$$\chi^A(s) = \begin{cases} (-1)^h & \text{if } \lambda = (m - h, 1^h) \text{ is a hook} \\ 0 & \text{if } \lambda \text{ is not a hook} \end{cases}$$

The product of the JM-elements $X_2X_3 \ldots X_m$ is the sum of all $m$-cycles in $S_m$ (Exercise 4.1.2). Since there are $(m - 1)!$ such cycles and all are conjugates of $s$, we have

$$\chi^A(X_2X_3 \ldots X_m) = (m - 1)! \chi^A(s).$$

In order to determine $\chi^A(X_2X_3 \ldots X_m)$, fix a GZ-basis $(w_T)_{T \in \lambda}$-tableau of $V^\lambda$. Then $X_iw_T = c_T,iw_T$ by (4.43), where $c_T$ is the content of the $\lambda$-tableau $T$. Thus,

$$X_2X_3 \ldots X_mw_T = c_{T,2}c_{T,3} \ldots c_{T,m}w_T.$$  

If $\lambda$ is not a hook, then there must be at least two boxes on the 0-diagonal of $\lambda$, whence $a_{T,i} = 0$ for some $i \neq 1$ and so $X_2X_3 \ldots X_mw_T = 0$. Therefore, $\chi^A(s) = 0$ holds in this case, as was to be shown.
Now assume that $\lambda$ is a hook, say $\lambda = (m - h, 1^h)$. Then we know from Example 4.14(b) that $V^\lambda = \Lambda^h V_{m-1}$ and $\chi^\lambda(s)$ is not hard to evaluate from this. Instead, let us proceed directly and observe that there are \( \binom{m-1}{h} \) possible $\lambda$-tableaux: they are obtained by selecting $h$ numbers from \( \{2, 3, \ldots, m\} \) and filling them into the boxes in the “leg” of $\lambda$ while the remaining numbers occupy the “arm” of $\lambda$, with numbers increasing along both arm and leg. Moreover, the content of each $\lambda$-tableau $T$ has entries $-h, -h + 1, \ldots, m - h - 1$ in some order. Therefore, \( c_{T,1}c_{T,3} \ldots c_{T,m} = (-1)^h h!(m - h - 1)! \) for each $T$ and so

\[
\chi^\lambda(X_2X_3 \ldots X_m) = (-1)^h h!(m - h - 1)! \binom{m-1}{h} = (-1)^h (m - 1)!,
\]

which amounts to the desired formula for $\chi^\lambda(s)$.

**Step 4: proof of (4.46) for $\lambda/\mu$ connected.** For this step, which will complete the proof of the Murnaghan-Nakayama Rule, we will determine certain multiplicities of $V^\nu \in \text{Irr} S_m$ in $V^{\lambda/\mu}$. In light of Step 3, we are specifically interested in the case when the partition $\nu + m$ is a hook.

**Proposition 4.28.** Let $\lambda/\mu$ be connected with $|\lambda/\mu| = m$ and let $\nu = (m - h, 1^h)$ be a hook partition. Then

\[
m(V^\nu, V^{\lambda/\mu}) = \begin{cases} 
1 & \text{if } \lambda/\mu \text{ is a skew hook with } h(\lambda/\mu) = h \\
0 & \text{otherwise.}
\end{cases}
\]

Let us grant this proposition for now and proceed to derive (4.46). So let $\lambda/\mu$ be as in the proposition. Then, for every $m$-cycle $s \in S_m$,

\[
\chi^{\lambda/\mu}(s) = \sum_{\alpha+m} m(V^\alpha, V^{\lambda/\mu}) \chi^\alpha(s)
\]

\[
= \text{Step 3 } \sum_{h=0}^{m-1} m(V^{(m-h,1^h)}, V^{\lambda/\mu}) (-1)^h
\]

\[
= \begin{cases} 
(-1)^{h(\lambda/\mu)} & \text{if } \lambda/\mu \text{ is a skew hook} \\
0 & \text{otherwise.}
\end{cases}
\]

This is exactly what (4.46) states. It remains to prove the proposition.

**Proof of Proposition 4.28.** First assume that $\lambda/\mu$ is connected but not a skew hook. Then $\lambda/\mu$ must contain some $2 \times 2$-square $\lambda^0$. It follows that there exists a sequence of partitions $\Lambda = (\mu = \lambda^0 < \cdots < \lambda^{s-1} < \lambda^s < \cdots < \lambda^l = \lambda)$ with \( |\lambda^i/\lambda^{i-1}| = 1 \) for $i \neq s$ and $\lambda^s/\lambda^{s-1}$ the $2 \times 2$-square; see the picture for a possible choice of $\lambda^{s-1}$ and $\lambda^s$. In the notation of Lemma 4.27, $\Lambda$ gives a representation $W$ of the Young subgroup $S_m$, where $m = (1, \ldots, 1, 4, 1, \ldots, 1)$,
and an epimorphism $W^\uparrow_{S_m} \rightarrow V^{\lambda/\mu}$. The subgroup $S_m \leq S_m$ is conjugate to the standard subgroup $S_4 \leq S_m$ and $W$ is isomorphic to $V^{(2,2)} \in \text{Irr} S_4$. Thus, we obtain an epimorphism $V^{(2,2)} \uparrow_{S_4} \rightarrow V^{\lambda/\mu}$. Now let $V^\alpha (\alpha \vdash m)$ be an irreducible constituent of $V^{\lambda/\mu}$. Then we have an epimorphism $V^{(2,2)} \uparrow_{S_4} \rightarrow V^{\lambda/\mu} \rightarrow V^\alpha$. By the branching rule (4.14), this implies that $(2, 2) \leq \alpha$. Therefore, $\alpha$ is a non-hook, which proves the proposition when $\lambda/\mu$ is not a skew hook.

Now let $\lambda/\mu$ be a skew hook. As we have remarked earlier (see the paragraph before Theorem 4.22), the structure of $V^{\lambda/\mu}$ as $S_m$-representation only depends on the geometric shape of $\lambda/\mu$, not on the particular realization with $\lambda$ and $\mu$. Thus, choosing $\lambda$ so that $|\lambda|$ is minimal, we may assume that the skew hook $\lambda/\mu$ touches both axes as in the picture. If the given hook partition $\nu = (m - h, 1^h)$ satisfies $h \neq h(\lambda/\mu)$, then $\nu \nleq \lambda$ and so $V^\nu$ is not a constituent of $V^{\lambda/\mu}$ either; so $m(V^\nu, V^{\lambda/\mu}) = 0$ as asserted in the proposition. Finally, assume that $h = h(\lambda/\mu)$. Then we may transform $\lambda/\mu$ into $\nu$ by moving boxes along diagonals and $\mu$ has the same shape as $\lambda/\nu$. Therefore, $V^{\lambda/\nu}$ and $V^\mu$ are isomorphic irreducible representations of $S_m$, and hence $V^\nu \boxtimes V^{\lambda/\nu}$ and $V^\nu \boxtimes V^\mu$ are isomorphic irreducible representations of $S_m \times S_k$, where $k = |\mu| = |\lambda| - m$ (Exercise 1.5.9). By Lemma 4.27,

$$V^{\lambda \downarrow}_{S_m \times S_k} \cong \bigoplus_{\alpha \vdash m, \alpha \leq \lambda} V^\alpha \boxtimes V^{\lambda/\alpha}.$$ 

Thus, $X^{\lambda \downarrow}_{S_m \times S_k} = \sum_{\alpha} X^\alpha X^{\lambda/\alpha}$ by (1.52). The multiplicity of the irreducible representation $V^\nu \boxtimes V^{\lambda/\nu}$ in $V^{\lambda \downarrow}_{S_m \times S_k}$ is given by the inner product,

$$\left( X^{\nu \downarrow}_{S_m \times S_k}, X^{\lambda \downarrow}_{S_m \times S_k} \right) = \sum_{\alpha} \left( X^\nu X^{\lambda/\nu}, X^\alpha X^{\lambda/\alpha} \right) + \sum_{\alpha} \left( X^\nu X^\alpha, X^{\lambda/\nu} X^{\lambda/\alpha} \right) = \sum_{\alpha} \delta_{\nu, \alpha} \left( X^{\lambda/\nu}, X^{\lambda/\alpha} \right) = 1 .$$

Therefore, $m(V^\nu \boxtimes V^\mu, V^{\lambda \downarrow}_{S_m \times S_k}) = m(V^\nu \boxtimes V^{\lambda/\nu}, V^{\lambda \downarrow}_{S_m \times S_k}) = 1$. A similar calculation shows that $m(V^\nu \boxtimes V^\mu, V^{\lambda \downarrow}_{S_m \times S_k}) = m(V^\nu, V^{\lambda/\mu})$, whence $m(V^\nu, V^{\lambda/\mu}) = 1$ as desired. This completes the proof of the proposition, and hence of the Murnaghan-Nakayama Rule as well. \qed
Exercises for Section 4.6

4.6.1 (Some character values). Prove the following character formulae from [75, §2], where \( \alpha \vdash m \) and \( m_\alpha(j) = \# \{ i \mid \alpha_i = j \} \):

(a) \( \chi_\alpha^{(m-2,2)} = \frac{1}{2}(m_\alpha(1) - 1)(m_\alpha(1) - 2) - m_\alpha(2) \)

(b) \( \chi_\alpha^{(m-2,2)} = \frac{1}{2}(m_\alpha(1) - 1)(m_\alpha(1) - 2) + m_\alpha(2) - 1 \)

(c) \( \chi_\alpha^{(m-3,3)} = \frac{1}{6}m_\alpha(1)(m_\alpha(1) - 1)(m_\alpha(1) - 5) + (m_\alpha(1) - 1)m_\alpha(2) + m_\alpha(3) \)

4.6.2 (Some more character calculations). Compute \( \chi_{(5,4,3)}^{(6,3,2,1)} \) again (as in Examples 4.23 and 4.26), using the arrangement \((3,6,2,1)\) of the partition \((6,3,2,1)\).
Also compute \( \chi_{(5,4,2)}^{(5,4,3)} \) and \( \chi_{(5,4,3,1)}^{(5,4,3)} \).

4.7. Schur-Weyl Duality

We will now use our knowledge of the representations of the symmetric groups \( S_n \) to construct certain irreducible representations of the general linear group \( \text{GL}(V) \) for a finite-dimensional \( \k \)-vector space \( V \). The material presented in this section is but a modest introduction to the representation theory of classical groups. For a more thorough coverage of this territory, the reader is referred to Weyl’s classic [205] as well as to the more recent (and more readable) monographs of Goodman and Wallach [88], [89] and Procesi [168].

Throughout this section, \( V \) denotes a finite-dimensional \( \k \)-vector space. With the exception of §4.7.1, the material below assumes a base field \( \k \) of characteristic 0 as in the earlier sections of this chapter, but \( \k \) need no longer be algebraically closed.

4.7.1. The Double Centralizer Theorem for Semisimple Algebras

This subsection revisits and expands on some general material for algebras that was first discussed in §1.2.5. The base field \( \k \) can be arbitrary for now. Recall that, for any \( A \in \text{Alg}_\k \) and any \( W \in \text{Rep}_A \), the image of \( A \) in \( \text{End}_\k(W) \) is denoted by \( A_W \). Let us now use \( ' \) to denote centralizers in \( \text{End}_\k(W) \). Besides \( A_W \), we are especially interested in the following two subalgebras of \( \text{End}_\k(W) \):

\[ A_W' = \text{End}_A(W) \quad \text{and} \quad A_W'' = \text{BiEnd}_A(W). \]

Thus, \( A_W \subseteq A_W'' \). We have already seen that equality holds if \( A \) is semisimple and \( W \) is irreducible (Corollary 1.34). Theorem 4.29 below, called the Double Centralizer Theorem, establishes this equality more generally and it also gives a bijection between \( \text{Irr} A_W \) and \( \text{Irr} A_W' \).

Note that \( \text{Irr} A_W \subseteq \text{Irr} A \) via inflation. We will continue to denote the Schur division algebra of \( S \in \text{Irr} A \) by \( D(S) \) rather than \( A'_S \). The space \( \text{Hom}_A(S,W) \) is naturally a \( (A_W',D(S)) \)-bimodule via composition of functions (Example 1.3).
Hence, we may form the tensor product $\text{Hom}_A(S, W) \otimes_D S$ and view it as a representation of $A'_W$ in the usual fashion (§B.1.2). In fact, $\text{Hom}_A(S, W) \otimes_D S$ is a representation of the algebra $A'_W \otimes A$: the “outer tensor product” action on $\text{Hom}_A(S, W) \otimes S$, with $A'_W$ acting on $\text{Hom}_A(S, W)$ and $A$ on $S$ (§1.5.1), passes down to an action on the space $\text{Hom}_A(S, W) \otimes_D S$. We will denote this representation of $A'_W \otimes A$ by $\text{Hom}_A(S, W) \otimes_D S$. The given $W$ is also a representation of $A'_W \otimes A$, via $(a' \otimes a).w = a'(a.w) = a.a'(w)$.

**Theorem 4.29.** Let $A \in \text{Alg}_k$ be semisimple, let $W \in \text{Rep} A$ be finitely generated, and let $A_W, A'_W, A''_W \subseteq \text{End}_k(W)$ be as above. Then:

(a) $A_W$ and $A'_W$ are both semisimple.

(b) $S \leftrightarrow \text{Hom}_A(S, W)$ gives a bijection between $\text{Irr} A_W$ and $\text{Irr} A'_W$.

(c) $W \cong \bigoplus_{S \in \text{Irr} A_W} \text{Hom}_A(S, W) \otimes_D S$ in $\text{Rep} (A'_W \otimes A)$; the summands are pairwise non-isomorphic irreducible representations of $A'_W \otimes A$.

(d) $A_W = A''_W$.

**Proof.** Being an image of the semisimple algebra $A$, the algebra $A_W$ is semisimple as well. Since $W$ is a faithful representation of $A_W$, the $S$-homogenous component $W(S)$ is nonzero for each $S \in \text{Irr} A_W$; in fact,

$$W(S) \cong S^\oplus_{m_S}$$

for some $m_S \in \mathbb{N}$.

Thus, $W = \bigoplus_{S \in \text{Irr} A_W} W(S) \cong \bigoplus_{S \in \text{Irr} A_W} S^\oplus_{m_S}$.

Proposition 1.33 now gives the isomorphism

$$A'_W = \text{End}_A(W) \cong \prod_{S \in \text{Irr} A_W} \text{Mat}_{m_S}(D(S)).$$

By Wedderburn’s Structure Theorem (§1.4.4), the algebra on the right is a semisimple, whence $A'_W$ is semisimple. Wedderburn’s Structure Theorem also tells us that $\text{Irr} A'_W$ is in bijection with the matrix components in (4.50), and hence with $\text{Irr} A_W$. This proves (a) and some of (b).

To finish (b), let $\pi_S: W \rightarrow W(S)$ denote the projection along the sum of the homogeneous components $W(S')$ with $S' \neq S$. Then $\pi_S \in A'_W$ and $\pi_S$ acts as the identity on $\text{Hom}_A(S, W)$ and as 0 on $\text{Hom}_A(S', W)$ with $S' \neq S$. Therefore, the various $\text{Hom}_A(S, W) \in \text{Rep} A'_W$ are non-isomorphic. Moreover, each $\text{Hom}_A(S, W)$ is irreducible. Indeed, let $f, f' \in \text{Hom}_A(S, W)$ be given with $f \neq 0$. Then $f$ is injective and $W = f(S) \oplus K$ for some $A$-subrepresentation $K \subseteq W$. Define $a': W \rightarrow W$ to be $f' \circ f^{-1}$ on the summand $f(S)$ and 0 on $K$. Then $f' = a' \circ f$, which shows that $A'_W.f = \text{Hom}_A(S, W)$. This completes the proof of (b).
consequently, \( A^\prime \)

234 symmetric groups

Therefore, \( \text{Hom} \) surely non-isomorphic in \( \text{Hom} \).

Furthermore, for each \( S \),

\[
\text{End}_{\text{Mat}_m(S)}(S^\otimes m) \xleftarrow{\sim} \text{End}_{\text{D}(S)}(S^\otimes m) \xrightarrow{\psi} \text{End}_{\text{D}(S)}(S) = A''_S
\]

Consequently, \( A''_W \equiv \prod_{S \in \text{Irr} A_W} A''_S \). Finally, \( A \equiv \prod_{S \in \text{Irr} A} A''_S \) by Corollary 1.34, and hence \( A \) maps onto \( A''_W \). This finishes the proof of the theorem. ❑

4.7.2. The Double Centralizer Theorem for \( S_n \) and GL(\( V \))

We now turn to the main topic of this section: the general linear group \( GL(V) \) of a finite-dimensional \( \mathbb{k} \)-vector space \( V \). For the remainder of this section, we assume again that the base field \( \mathbb{k} \) has characteristic 0.
For any integer \( n \geq 0 \), the tensor power \( V^\otimes n \) is a representation of \( \text{GL}(V) \) via the diagonal action,
\[
g.(v_1 \otimes v_2 \otimes \cdots \otimes v_n) = g(v_1) \otimes g(v_2) \otimes \cdots \otimes g(v_n),
\]
and we also have the place permutation action \( S_n \subset V^\otimes n \) that was already explored in §3.8.1:
\[
s.(v_1 \otimes v_2 \otimes \cdots \otimes v_n) = v_{s^{-1}1} \otimes v_{s^{-1}2} \otimes \cdots \otimes v_{s^{-1}n}.
\]
Since these two actions commute with each other, we obtain algebra maps
\[
\begin{align*}
\mathbb{k}[\text{GL}(V)] & \xrightarrow{\delta} \text{End}_{S_n}(V^\otimes n) \xrightarrow{\pi} \text{End}_{\mathbb{k}}(V^\otimes n) \\
\mathbb{k}S_n & \xrightarrow{\pi} \text{End}_{\text{GL}(V)}(V^\otimes n)
\end{align*}
(4.51)
\]

**Schur’s Double Centralizer Theorem.** Let \( V \in \text{Vect}_k \) be a finite dimensional. Then the maps \( \delta \) and \( \pi \) in (4.51) are both surjective.

**Proof.** We proceed in two steps.

**Step 1:** surjectivity of \( \delta \) implies surjectivity of \( \pi \). Put \( D = \text{Im} \delta \) and \( P = \text{Im} \pi \) and let \( . \)’ denotes centralizers in \( \text{End}_k(V^\otimes n) \). Then, certainly, \( D \subseteq P' = \text{End}_{S_n}(V^\otimes n) \) and \( P \subseteq D' = \text{End}_{\text{GL}(V)}(V^\otimes n) \). Now suppose that \( \delta \) is surjective: \( D = P' \). Then it follows that \( P'' = D' \). Since \( \mathbb{k}S_n \) is semisimple by Maschke’s Theorem (§3.4.1), Theorem 4.29(d) further tells us that \( P'' = P \). Hence, \( P = D' \) and so \( \pi \) is surjective.

**Step 2:** \( \delta \) is surjective. We use the standard \( \mathbb{k} \)-linear isomorphism,
\[
\Psi : \quad \text{End}_k(V^\otimes n) \xrightarrow{\sim} \text{End}_k(V^\otimes n)
\]
\[
\phi_1 \otimes \cdots \otimes \phi_n \mapsto (v_1 \otimes \cdots \otimes v_n \mapsto \phi_1(v_1) \otimes \cdots \otimes \phi_n(v_n))
\]
This map is \( S_n \)-equivariant: for \( \phi = \phi_1 \otimes \cdots \otimes \phi_n \), \( v = v_1 \otimes \cdots \otimes v_n \) and \( s \in S_n \),
\[
(s.\Psi(\phi))(v) = s.(\Psi(\phi)(s^{-1}.v)) = s. (\Psi(\phi)(v_{s^{-1}1} \otimes \cdots \otimes v_{s^{-1}n}))
\]
\[
= s. (\phi_1(v_{s^{-1}1}) \otimes \cdots \otimes \phi_n(v_{s^{-1}n})) = \phi_{s^{-1}1}(v_1) \otimes \cdots \otimes \phi_{s^{-1}n}(v_n)
\]
\[
= \Psi(s.\phi)(v).
\]
Therefore, \( \Psi \) restricts to an isomorphism of \( S_n \)-invariants:
\[
(\text{End}_k(V^\otimes n))^{S_n} \xrightarrow{\sim} (\text{End}_k(V^\otimes n))^{S_n} = \text{End}_{S_n}(V^\otimes n).
\]
For \( g \in \text{GL}(V) \), we have \( \Psi(g^\otimes n) = \delta(g) \). Therefore, surjectivity of \( \delta \) follows from the fact that the powers \( g^\otimes n \) with \( g \in \text{GL}(V) \) span the vector space \( (\text{End}_k(V^\otimes n))^{S_n} \), which in turn follows from Proposition 3.37: \( \text{GL}(V) \) is the principal open (and hence dense) subset \( \text{End}_k(V)_{\det} \subseteq \text{End}_k(V) \), where \( \det \) is the determinant, a polynomial
4. Symmetric Groups

function on $\text{End}_G(V)$ (homogeneous of degree $\dim_V V$). This completes the proof of Schur’s Double Centralizer Theorem.

4.7.3. Schur Functors

Assume that $V \in \text{Rep} G$ for some group $G$. Then, for any $W \in \text{Rep} S_n$, we obtain another representation of $G$ by defining

$$S^W V \overset{\text{def}}{=} \text{Hom}_{S_n}(W, V^{\otimes n})$$

with $G$-action $g.f = g^{\otimes n} \circ f$ for $g \in G$ and $f \in S^W V$. If $f : V \to V'$ is a morphism in $\text{Rep} G$, then the map $(f^{\otimes n})_\ast = f^{\otimes n} \circ \cdot : S^W V \to S^W V'$ is a morphism in $\text{Rep} G$. Thus, we obtain a functor $S^W : \text{Rep} G \to \text{Rep} G$.

Since the group algebra of $\mathbb{k}S_n$ is semisimple, the description of $S^W V$ reduces to the case where $W$ is irreducible; so $W \cong V^\lambda$ for some $\lambda \vdash n$. The functor $S^W$ will then be denoted by $S^\lambda$. Below, we assume again that $\dim V < \infty$ as in the rest of this section, and we shall focus on $G = \text{GL}(V)$. Thus, for any partition $\lambda \vdash n$, we are interested in the representation

$$S^\lambda V \overset{\text{def}}{=} \text{Hom}_{S_n}(V^\lambda, V^{\otimes n}) \in \text{Rep}_{\text{fin}} \text{GL}(V) \quad (n = |\lambda|)$$

**Examples 4.30.** By Example 4.14(a), the partition $\lambda = (n)$ gives $V^\lambda = \mathbb{1}$. Furthermore, by Example 3.18 and Lemma 3.36, $\text{Hom}_{S_n}(\mathbb{1}, V^{\otimes n}) \cong (V^{\otimes n})^S_n = S^n V \cong \text{Sym}^n V$. Thus, we obtain the following isomorphism in $\text{Rep}_{\text{fin}} \text{GL}(V)$:

$$S^{(n)} V \cong \text{Sym}^n V.$$  

Similarly, for $\lambda = (1, 1, \ldots, 1) \vdash n$, we have $V^\lambda = \text{sgn}$ and $\text{Hom}_{S_n}(\text{sgn}, V^{\otimes n}) \cong (V^{\otimes n})(\text{sgn}) = \Lambda^n V \cong \Lambda^n V$; so

$$S^{(1, \ldots, 1)} V \cong \Lambda^n V.$$  

**Main Result**

For any partition $\lambda = (\lambda_1, \lambda_2, \ldots) \in \mathcal{P}$, we let $\ell(\lambda)$ denote the number of nonzero parts $\lambda_i$ or, equivalently, the number of rows in the Young diagram of $\lambda$. Thus, $\ell(\lambda) = h(\lambda) + 1$ in the notation of §4.6.3.

Only part (a) of the following technical lemma will be needed for now; (b) will only play a role in Section 8.8, but it comes for free now. Recall that, associated to any sequence $m = (m_1, m_2, \ldots)$ of non-negative integers with $|m| = \sum_i m_i = n$, there is the Young subgroup $S_m = S_{m_1} \times S_{m_2} \times \ldots \leq S_n$ (§§3.8.2, 4.6.4). We let $\varepsilon_i$ denote the sequence with 1 in the $i$th position and 0 elsewhere.

**Lemma 4.31.**

(a) $S^\lambda V \neq 0$ if and only if $\ell(\lambda) \leq \dim V$.

(b) $(V^\lambda)^{S_m} = \mathbb{k}$ for $m = \lambda$ while $(V^\lambda)^{S_m} = 0$ for $m = \lambda + \varepsilon_i - \varepsilon_j$ with $i < j$. 


4.7. Schur-Weyl Duality

Proof. Put \( n = |\lambda| \) and \( d = \dim_k V \). The structure of \( V^\otimes n \) as \( S_n \)-representation was given in (3.70):

\[
V^\otimes n \cong \bigoplus_{m \in \mathbb{Z}_+^d : |m| = n} k[S_n/S_m] \cong \bigoplus_{m \in \mathbb{Z}_+^d : |m| = n} 1 \uparrow S_n^m.
\]

Thus,

\[
\mathcal{S}^d V \cong \bigoplus_{m \in \mathbb{Z}_+^d : |m| = n} \text{Hom}_{S_n}(V^{\lambda}, 1 \uparrow S_m^\lambda)
\]

\[
\cong (3.9) \bigoplus_{m \in \mathbb{Z}_+^d : |m| = n} \text{Hom}_{S_m}(V^{\lambda}, 1 \downarrow S_n^\lambda)
\]

\[
\cong \bigoplus_{m \in \mathbb{Z}_+^d : |m| = n} ((V^\lambda)^{S_m})^*,
\]

where the last isomorphism comes from restricting homomorphisms to invariants. Recall that \( S_m = S_m' \), where \( m' = (m'_1, \ldots, m'_l) \) with \( l \leq d \) is the composition of \( n \) that is obtained from \( m \) by removing all components with value 0. Letting \( \mathcal{P}_m \) denote the set of sequences of partitions \( \Lambda = (\emptyset = \lambda^0 < \lambda^1 < \cdots < \lambda^l = \lambda) \) with \( |\lambda^i/\lambda^{i-1}| = m'_i \), Lemma 4.27 gives

\[
V^{\lambda} \downarrow S_m \cong \bigoplus_{\Lambda \in \mathcal{P}_m} \left( V^{\lambda^0/\lambda^1} \boxtimes V^{\lambda^1/\lambda^2} \boxtimes \cdots \boxtimes V^{\lambda^l/\lambda^l} \right).
\]

Since \( (V^{\lambda^i/\lambda^{i-1}})^{S_m} \cong \bigotimes_{i=1}^l (V^{\lambda^i/\lambda^{i-1}})^{S_{m'_i}} \) (Exercise 3.3.2), we obtain a \( k \)-linear isomorphism

\[
(4.52) \quad \mathcal{S}^d V \cong \bigoplus_{n \in \mathbb{Z}_+^d} \bigoplus_{\Lambda \in \mathcal{P}_m} \bigotimes_{i=1}^l ((V^{\lambda^i/\lambda^{i-1}})^{S_{m'_i}})^*.
\]

(a) Proposition 4.28 with \( h = 0 \) tells us that \( (V^{\lambda^i/\lambda^{i-1}})^{S_{m'_i}} \neq 0 \) if and only if \( \lambda^i/\lambda^{i-1} \) is a skew hook with \( h(\lambda^i/\lambda^{i-1}) = 0 \), that is, the diagram of \( \lambda^i/\lambda^{i-1} \) consists of just one row. To summarize, \( \mathcal{S}^d V \neq 0 \) if and only if there is a composition \( m' = (m'_1, \ldots, m'_l) \) of \( n \) with \( l \leq d \) parts and a sequence of partitions \( \Lambda = (\emptyset = \lambda^0 < \lambda^1 < \cdots < \lambda^l = \lambda) \) such that the diagram of each \( \lambda^i/\lambda^{i-1} \) is a row with \( m'_i \) boxes. Observe that, for each such \( m' \) and \( \Lambda \), the partitions \( \lambda^i \) have at most \( i \) rows. Therefore, if \( m' \) and \( \Lambda \) exist, then \( \lambda \) has at most \( l \leq d \) rows; so \( \ell(\lambda) \leq d \). On the other hand, if \( \ell(\lambda) \leq d \), then the composition \( m' = \lambda \) and the sequence \( \Lambda \) with \( l = \ell(\lambda) \) and with \( \lambda^i \) consisting of the first \( i \) rows of \( \lambda \) clearly satisfy our requirements; so \( \mathcal{S}^d V \neq 0 \) in this case. This proves (a).

(b) For \( m = m' = \lambda \), it has already been observed above that \( (V^\lambda)^{S_m} \equiv \bigoplus_{\Lambda \in \mathcal{P}_m} \bigotimes_{i=1}^l (V^{\lambda^i/\lambda^{i-1}})^{S_{m'_i}} \) is nonzero if and only if the diagram of each \( \lambda^i/\lambda^{i-1} \)
is a row, in which case \( V^{\lambda^i / \lambda^{i-1}} = \mathbb{1}_{S_{\lambda_i}} \). Thus, the only sequence \( \Lambda \) contributing to \( (V^\lambda)^{S_m} \) is the one where each \( \lambda^i \) consists of the first \( i \) rows of \( \lambda \), and the contribution of this sequence is 1-dimensional. For \( m = \lambda + e_i - e_j \) with \( i < j \leq \ell \), there is no sequence \( \Lambda = (\varnothing = \lambda^0 < \lambda^1 < \cdots < \lambda^\ell = \lambda) \) with each \( \lambda^k / \lambda^{k-1} \) a row of length \( m_k \), because \( \lambda^i \) does not fit inside \( \lambda \). Therefore, \( (V^\lambda)^{S_m} = 0 \) in this case. \( \square \)

**Theorem 4.32.** Let \( V \) be a \( \mathbb{k} \)-vector space with \( d := \dim V < \infty \). Then:

(a) If \( \lambda \) is a partition with \( \ell(\lambda) \leq d \), then \( S^d V \in \text{Rep}_{\text{fin}} \text{GL}(V) \) is irreducible.

(b) If \( \lambda \neq \mu \) are partitions with \( \ell(\lambda), \ell(\mu) \leq d \), then \( S^d V \not\cong S^\mu V \) in \( \text{Rep}_{\text{fin}} \text{GL}(V) \).

(c) \( V^{\otimes n} \cong \bigoplus_{\lambda \vdash n; \ell(\lambda) \leq d} S^d V \boxtimes V^\lambda \) in \( \text{Rep}_{\text{fin}} (\text{GL}(V) \times S_n) \); the summands are pairwise non-isomorphic irreducible representations of \( \text{GL}(V) \times S_n \).

**Proof.** We will apply Theorem 4.29 with \( A = \mathbb{k}S_n \), \( W = V^{\otimes n} \) and \( S = V^\lambda \) \( (n = |\lambda|) \). Thus, \( \text{Hom}(S, W) = S^d V \) and \( A_W = (\mathbb{k}S_n)_{V^{\otimes n}} \subseteq \text{End}_{\mathbb{k}}(V^{\otimes n}) \). Moreover, \( A_W = \text{End}_{S_n}(V^{\otimes n}) \) is a homomorphic image of \( \mathbb{k} [\text{GL}(V)] \) by Schur’s Double Centralizer Theorem and

\[
\text{Irr} A_W = \{ V^\lambda \mid \lambda \vdash n, S^d V \neq 0 \} = \{ V^\lambda \mid \lambda \vdash n, \ell(\lambda) \leq d \}.
\]

(a) Since \( V^\lambda \) is an irreducible representation of \( A_W \), it follows from Theorem 4.29(b) that \( S^d V \) is an irreducible representation of \( A_W \) and hence also of \( \text{GL}(V) \), because \( \mathbb{k} [\text{GL}(V)] \) maps onto \( A_W \).

(b) Let \( \lambda \) and \( \mu \) be distinct partitions with \( \ell(\lambda), \ell(\mu) \leq d \). First, let us assume that \( |\lambda| = |\mu| = n \). Then \( V^\lambda \) and \( V^\mu \) are non-isomorphic irreducible representations of the algebra \( A_W \). Therefore, \( S^d V \) and \( S^\mu V \) are non-isomorphic irreducible representations of \( A_W \) by Theorem 4.29(b). As in (a), we obtain that \( S^d V \) and \( S^\mu V \) are non-isomorphic as representation of \( \text{GL}(V) \). Below, we will show that the same conclusion also holds if \( |\lambda| \neq |\mu| \).

(c) In view of the foregoing and the fact that \( D(V^\lambda) = \mathbb{k} \) (Corollary 4.16), Theorem 4.29(c) gives the following isomorphism in \( \text{Rep} (\text{GL}(V) \times S_n) \):

\[
V^{\otimes n} \cong \bigoplus_{\lambda \vdash n; \ell(\lambda) \leq d} \text{Hom}_{S_n} (V^\lambda, V^{\otimes n}) \boxtimes D(V^\lambda) V^\lambda \\
\cong \bigoplus_{\lambda \vdash n; \ell(\lambda) \leq d} S^d V \boxtimes V^\lambda.
\]

Theorem 4.29(c) also tells us that the summands \( S^d V \boxtimes V^\lambda \) are pairwise non-isomorphic irreducible representations of \( A_W \otimes A \), and hence of \( \text{GL}(V) \times S_n \). This proves (c).

We still have to justify the claim that \( S^d V \) and \( S^\mu V \) are non-isomorphic representations of \( \text{GL}(V) \) if \( \ell(\lambda), \ell(\mu) \leq d \) and \( |\lambda| \neq |\mu| \). But, \( S^d V \boxtimes V^\lambda \cong (S^d V)^{\otimes f^d} \)
in $\text{Rep} \ GL(V)$, since $\dim V^\lambda = f^\lambda$. So the decomposition of $V^{\otimes n}$ in part (c) shows that $\mathbb{S}^{\lambda} V$ is a direct summand of $V^{\otimes n} \downarrow_{GL(V)}$, where $n = |\lambda|$. Similarly, $\mathbb{S}^{\mu} V$ is a direct summand of $V^{\otimes m} \downarrow_{GL(V)}$ for $m = |\mu|$. Therefore, it suffices to show that $\text{Hom}_{GL(V)}(V^{\otimes m}, V^{\otimes n}) = 0$ if $m \neq n$. But, for any $c \in \mathbb{k}^X$, the scalar operator $c \id_V \in GL(V)$ acts on $V^{\otimes n}$ and $V^{\otimes m}$ as $c^n$ and $c^m$, respectively. Therefore, if $f \in \text{Hom}_{GL(V)}(V^{\otimes m}, V^{\otimes n})$, then $c^m f = c^n f$. Choosing $c$ to be a non-root of unity, this forces $f = 0$ as desired. This completes the proof of the theorem. □

4.7. Schur-Weyl Duality

The argument at the beginning of the last paragraph of the above proof gives the following decomposition of $V^{\otimes n}$ into homogeneous components for $GL(V)$:

\[(4.53)\]

\[V^{\otimes n} \cong \bigoplus_{\lambda \vdash n : \ell(\lambda) \leq \dim V} (\mathbb{S}^{\lambda} V)^{\otimes f^\lambda}\]

This formula generalizes (3.67), which dealt with the case $n = 2$. The partition $\lambda = (n)$ produces the summand $\mathbb{S}^{(n)} V \cong \text{Sym}^n V$ in (4.53) and, if $n \leq \dim V$, then $\lambda = (1, 1, \ldots, 1)$ contributes the summand $\mathbb{S}^{(1, \ldots, 1)} V \cong \Lambda^n V$, both with multiplicity $f^\lambda = 1$ (Examples 4.30). Theorem 4.32(c) also allows us to make the decomposition of $V^{\otimes n} \in \text{Rep} \ S_n$ into homogeneous components more precise than stated earlier in (3.65): since $\mathbb{S}^{\lambda} V \boxtimes V^\lambda \cong (V^\lambda)^{\oplus \dim \mathbb{S}^\lambda V}$ in $\text{Rep} \ S_n$, we obtain

\[(4.54)\]

\[V^{\otimes n} \cong \bigoplus_{\lambda \vdash n : \ell(\lambda) \leq \dim V} (V^\lambda)^{\oplus \dim \mathbb{S}^{\lambda} V}\]

The multiplicities $\dim \mathbb{S}^{\lambda} V$ will be determined later as an application of Weyl’s Character Formula (Corollary 8.40).

Exercises for Section 4.7

4.7.1 (Schur’s Double Centralizer Theorem for finite fields). Show that the proof of Schur’s Double Centralizer Theorem given in this section goes through as long as $\text{char} \ k > n$ and $|k| \geq n + \dim_V$. (Use Exercise C.3.2.)

4.7.2 (Iterated centralizers). Let $A \in \text{Alg}_k$ and let $\cdot'$ denote centralizers in $A$. Show that $B' = B'''$ holds for each subalgebra $B \subseteq A$. 

Part III

Lie Algebras
Chapter 5

Lie Algebras and Enveloping Algebras

The goal of this chapter is to deploy the fundamentals of the theory of Lie algebras and to furnish the basic tools for the investigation of their representations. Other than for groups, we do not assume any familiarity with Lie algebras. We shall however be somewhat brief in laying out the preliminaries on Lie algebras in Section 5.1, since much of this material is rooted in the theory of vector spaces or else closely mirrors similar developments for groups.

While representations of Lie algebras share many general features with group representations—our treatment will aim to highlight the parallels—Lie algebras do in fact behave very differently from groups in other ways. Indeed, as we have seen in Chapter 3, the representations of a given group $G$ over a field $k$ are governed by the group algebra $kG$. If $G$ is finite, then $kG$ is finite dimensional and, consequently, there are only finitely many non-equivalent irreducible representations, all of which are finite dimensional. The analogous associative vehicle for the representation theory of a Lie algebra $\mathfrak{g}$ is the so-called enveloping algebra $U\mathfrak{g}$, to be defined below. In contrast with group algebras, $U\mathfrak{g}$ is never finite dimensional or semisimple if $\mathfrak{g} \neq 0$. Also, it will turn out that “most” irreducible representations of a typical finite-dimensional Lie algebra over a field $k$ with $\text{char} k = 0$ are infinite dimensional and there are infinitely many non-equivalent irreducible representations. Other than in Chapters 3 and 4, primitive ideals will feature prominently and a geometric viewpoint will frequently be emphasized in our treatment of finite-dimensional Lie algebras and their enveloping algebras.

We work over an arbitrary base field $k$ in this chapter (except for Section 5.7). Additional hypotheses will be explicitly stated as they become necessary.
5.1. Lie Algebra Basics

5.1.1. Lie Algebras and their Homomorphisms

A Lie algebra\(^1\) over \(\mathbb{k}\) is a vector space \(g \in \text{Vect}_\mathbb{k}\) together with a \(\mathbb{k}\)-bilinear map,

\[
[\ldots]: g \times g \rightarrow g,
\]

satisfying the following conditions, for all \(x, y, z \in g\):

**alternating law:**  \[ [x, x] = 0, \]

**Jacobi identity:**  \[ [x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0. \]

The map \([\ldots]\) is called the Lie bracket of \(g\). As in our discussion of exterior algebras (§1.1.2), one sees that the alternating law implies the following anticommutativity rule, and is in fact equivalent to it in case \(\text{char} \mathbb{k} \neq 2\):

\[
[x, y] = -[y, x].
\]

Given two Lie algebras \(g\) and \(h\), a homomorphism from \(g\) to \(h\) is a map \(f : g \rightarrow h\) in \(\text{Vect}_\mathbb{k}\) satisfying \(f[x, y] = [fx, fy]\) for all \(x, y \in g\). In this way, we obtain the category of all Lie algebras over \(\mathbb{k}\),

\[
\text{Lie}_\mathbb{k}.
\]

Isomorphism, monomorphism, epimorphisms, endomorphisms, and automorphisms of Lie algebras are defined as in \(\text{Vect}_\mathbb{k}\): Lie algebra homomorphism that are bijective, injective, surjective, . . . .

**Example 5.1** (Abelian Lie algebras). Any \(V \in \text{Vect}_\mathbb{k}\) can be made into a Lie algebra by simply defining \([x, y] = 0\) for all \(x, y \in V\). Lie algebras of this form are called abelian. For any Lie algebra \(g\) and any subsets \(X, Y \subseteq g\), we put

\[
[X, Y] \overset{\text{def}}{=} \langle [x, y] \mid x \in X, y \in Y \rangle_\mathbb{k},
\]

where \(\langle \ldots \rangle_\mathbb{k}\) denotes the \(\mathbb{k}\)-subspace that is spanned by the indicated elements. Thus, \([X, Y] = [Y, X]\) by anticommutativity. Abelian Lie algebras are characterized by the condition \([g, g] = 0\).

**Example 5.2** (Lie algebras from associative algebras). For any \(A \in \text{Alg}_\mathbb{k}\), we have already considered the Lie commutator in earlier chapters: \([a, b] = ab - ba\) for \(a, b \in A\). The alternative law is evident for the Lie commutator and the Jacobi identity is also easily verified. In fact, for any \(a, b, c \in A\), we have the Leibniz identity,

\[
[a, bc] = [a, b]c + b[a, c].
\]

\(^1\)Lie algebras are named after the Norwegian mathematician Sophus Lie (1842–1899), who introduced them as a tool in his investigations of “continuous transformation groups,” now called, after him, Lie groups.
which in turn readily implies the Jacobi identity for the Lie commutator. Thus, viewing $A$ as a $\mathbb{k}$-vector space endowed with the commutator as Lie bracket, we obtain a Lie algebra; it will be denoted by $A_{\text{Lie}}$. This yields a functor,

$$\cdot_{\text{Lie}} : \text{Alg}_k \to \text{Lie}_k,$$

which will play a similar role as the functor $\cdot^k : \text{Alg}_k \to \text{Groups}$ (§3.1.1).

### 5.1.2. Representations of Lie Algebras

The endomorphism algebra $\text{End}_k(V)$ of a given $V \in \text{Vect}_k$ will continue to play a prominent role. For the underlying Lie algebra, we will use the special notation

$$\mathfrak{gl}^k(V) \overset{\text{def}}{=} \text{End}_k(V)_{\text{Lie}}$$

We will also write $\mathfrak{gl}_n(\mathbb{k})$ or simply $\mathfrak{gl}_n$ if $V = \mathbb{k}^n$. These notations parallel the group theoretical notations $\text{GL}(V) = \text{End}_k(V)^\times$ and $\text{GL}_n(\mathbb{k}) = \text{End}_k(\mathbb{k}^n)^\times$ as does the nomenclature: $\mathfrak{gl}(V)$ is called the \textit{general linear Lie algebra} of $V$.

A \textit{representation} of a given $\mathfrak{g} \in \text{Lie}_k$ is a homomorphism in $\text{Lie}_k$ of the form

$$\mathfrak{g} \longrightarrow \mathfrak{gl}(V)$$

for some $V \in \text{Vect}_k$. As with associative algebras and groups, we will denote the image of $x \in \mathfrak{g}$ in $\mathfrak{gl}(V)$ by $x_V$ and we will write $x_V(v) = x_V(v)$. Also, the vector space $V$ will often be referred to as the representation rather than the actual Lie homomorphism. A subspace $U \subseteq V$ that is stable under all operators $x_V$ with $x \in \mathfrak{g}$ is called a \textit{subrepresentation} of $V$, and $V$ is said to be \textit{irreducible} if 0 and $V$ are distinct and are the only subrepresentations.

The formal aspects of representations of Lie algebras will be systematically addressed in Section 5.5; in particular, we will see that above notions do in fact reduce to the corresponding notions for algebras. For starters, here is at least one important example of a representation.

**Example 5.3** (The adjoint representation of a Lie algebra). For any $\mathfrak{g} \in \text{Lie}_k$, we may define the map

$$\text{ad} : \mathfrak{g} \longrightarrow \mathfrak{gl}(\mathfrak{g})$$

$$\begin{array}{ccc}
\varnothing & \varnothing \\
x & [x, \cdot] \\
\end{array}$$

It follows from the Jacobi identity that $\text{ad}$ is in fact a homomorphism in $\text{Lie}_k$ (Exercise 5.1.5). Thus, we obtain a representation of $\mathfrak{g}$, the so-called \textit{adjoint representation}. The vector space $\mathfrak{g}$, viewed as a representation of $\mathfrak{g}$ in this way, will be denoted by $\mathfrak{g}_{\text{ad}}$. 
5.1.3. Subalgebras and Ideals

Let \( g \in \text{Lie}_k \). A **Lie subalgebra** of \( g \), unsurprisingly, is defined to be a \( k \)-subspace \( h \subseteq g \) satisfying \([h, h] \subseteq h\). The alternative law and the Jacobi identity are then clearly inherited by \( h \) from \( g \), thereby making \( h \) a Lie algebra in its own right. For example, all 1-dimensional \( k \)-subspaces of \( g \) are (abelian) Lie subalgebras of \( g \).

Here are some more interesting examples.

**Example 5.4 (Transporters).** For any two \( k \)-subspaces \( U \subseteq V \) of \( g \), the set

\[
\{ x \in g \mid [x, V] \subseteq U \}
\]

is easily seen to be a Lie subalgebra of \( g \); it is sometimes called the **transporter** of \( V \) into \( U \). Taking \( U = 0 \) and \( U = V \), one in particular obtains the following Lie subalgebras, called respectively the **centralizer** and the **normalizer** of \( V \):

\[
C_g(V) \overset{\text{def}}{=} \{ x \in g \mid [x, V] = 0 \} \quad \text{and} \quad N_g(V) \overset{\text{def}}{=} \{ x \in g \mid [x, V] \subseteq V \}.
\]

The centralizer of \( V = g \) is called the **center** of \( g \):

\[
Z_g \overset{\text{def}}{=} \{ x \in g \mid [x, g] = 0 \}.
\]

A \( k \)-subspace \( a \subseteq g \) is called an **ideal** of \( g \) if \([g, a] \subseteq a\). Note that this condition is right-left symmetric. Ideals of \( g \) can also be characterized as the subrepresentations of \( g_{\text{ad}} \); they are in particular Lie subalgebras of \( g \), but the converse generally fails. For any ideal \( a \) of \( g \), the quotient \( g/a \in \text{Vect}_k \) inherits the structure of a Lie algebra by defining

\[
[x + a, y + a] := [x, y] + a \quad (x, y \in g).
\]

Well-definedness and the Lie algebra axioms are straightforward to verify as is the fact that the vector space epimorphism \( g \to g/a, x \mapsto x + a \), is map in \( \text{Lie}_k \).

**The Isomorphism Theorem.** We also leave it to the reader to check the details of the following Lie algebra version of the standard First Isomorphism Theorem for vector spaces. Recall that a homomorphism \( f : g \to h \) of Lie algebras is in the first place a \( k \)-linear map. So \( \text{Ker} f \) and \( \text{Im} f \) may be defined as in \( \text{Vect}_k \) and they are certainly subspaces of \( g \) and \( h \), respectively. In fact:

**Proposition 5.5.** Let \( f : g \to h \) be a homomorphism in \( \text{Lie}_k \). Then \( \text{Im} f \) is a Lie subalgebra of \( h \) and \( \text{Ker} f \) is an ideal of \( g \). If \( a \) is any ideal of \( g \) such that \( a \subseteq \text{Ker} f \), then there is a unique Lie algebra homomorphism \( \overline{f} : g/a \to h \) such that the following diagram commutes:

\[
\begin{array}{ccc}
g & \xrightarrow{f} & h \\
\downarrow \text{can} & & \downarrow \exists \overline{f} \\
g/a & \xrightarrow{\text{can}} & h/a
\end{array}
\]

In particular, \( g/\text{Ker} f \cong \text{Im} f \) in \( \text{Lie}_k \).
In the following, we will tacitly use this proposition as well as the Lie theoretic versions of the other isomorphism theorems; their formulation requires little imagination and they are consequences of Proposition 5.5, adapting the familiar proofs of the isomorphism theorems from linear algebra or group theory.

**Example 5.6** (Special linear Lie algebras). Let \( V \in \text{Vect}_k \) be nonzero and finite dimensional. Then the trace map is a Lie algebra epimorphism,

\[
\text{trace} = \text{trace}_V : \mathfrak{gl}(V) = \text{End}_k(V)_{\text{Lie}} \twoheadrightarrow k_{\text{Lie}},
\]

because \( \text{trace}([x,y]) = \text{trace}(xy) - \text{trace}(yx) = 0 \) for all \( x, y \in \mathfrak{gl}(V) \). Thus,

\[
\mathfrak{sl}(V) \overset{\text{def}}{=} \text{Ker}(\text{trace}_V)
\]

is an ideal of \( \mathfrak{gl}(V) \) such that \( \mathfrak{gl}(V)/\mathfrak{sl}(V) \cong k_{\text{Lie}} \). Viewing \( \mathfrak{sl}(V) \) as a Lie algebra in its own right, it is called the **special linear Lie algebra** of \( V \).

**Example 5.7** (The adjoint representation). Recall that the adjoint representation of a Lie algebra \( g \) gives a Lie algebra homomorphism \( \text{ad} : g \rightarrow \mathfrak{gl}(g) \) (Example 5.3). Clearly, \( \text{Ker}(\text{ad}) = Z(g) \), the center of \( g \). Thus, \( \text{ad} g \) is a Lie subalgebra of \( \mathfrak{gl}(g) \) and we have an isomorphism of Lie algebras,

\[
\text{ad} g \cong g / Z(g) \tag{5.2}
\]

**5.1.4. Examples of Lie Algebras**

**Dimensions 1 and 2.** By the alternative law, any 1-dimensional Lie algebra must be abelian; so there is only one such Lie algebra up to isomorphism.

In dimension 2, however, it turns out that there is also a unique (up to isomorphism) non-abelian Lie algebra \( g \). Indeed, the matrices \( x = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \) and \( y = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \) span the 2-dimensional \( k \)-subspace \( \begin{pmatrix} k & k \\ 0 & 0 \end{pmatrix} = kx \oplus ky \subseteq \mathfrak{gl}_2 \), which is in fact a non-abelian Lie subalgebra of \( \mathfrak{gl}_2 \), because \( [x,y] = xy - yx = y \). Conversely, let \( g \) be any non-abelian Lie algebra of dimension 2 and let \( x, y \) be any \( k \)-basis of \( g \). Bilinearity of the Lie bracket together with the alternative law imply that \( [g,g] = k[x,y] \); so \( \dim_k [g,g] = 1 \). Choosing our basis so that that \( x \notin [g,g] \) but \( y \in [g,g] \) we have \( [x,y] = \lambda y \) for some \( \lambda \in k^\times \). We may assume that \( \lambda = 1 \) by rescaling \( x \) if necessary. Thus, \( g = kx \oplus ky \) and

\[
[x,y] = y \tag{5.3}
\]

In view of bilinearity of \([\cdot,\cdot,\cdot]\) and the alternative law, the relation (5.3) determines the entire Lie bracket of \( g \). Therefore, \( g \) is isomorphic to the Lie subalgebra of \( \mathfrak{gl}_2 \) exhibited above.
Dimension 3: \( \mathfrak{sl}_2 \). The special linear Lie algebra of \( V = \mathbb{k}^n \) will be denoted by \( \mathfrak{sl}_n(\mathbb{k}) \) or just \( \mathfrak{sl}_n \) rather than \( \mathfrak{sl}(\mathbb{k}^n) \). Thus, \( \dim_k \mathfrak{sl}_n = n^2 - 1 \). The Lie algebra \( \mathfrak{sl}_2 \) will be particularly important later on. Note that 
\[
\mathfrak{sl}_2 = \mathbb{k}f \oplus \mathbb{k}h \oplus \mathbb{k}e
\]
with 
\[
f = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \text{and} \quad e = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.
\]
The Lie bracket of \( \mathfrak{sl}_2 \) is determined by the rules
\[
[h, f] = -2f, \quad [h, e] = 2e \quad \text{and} \quad [e, f] = h.
\]

Dimension 3: Heisenberg. Another example of a 3-dimensional Lie algebra is the so-called Heisenberg Lie algebra, which is defined to be the Lie subalgebra \( \mathfrak{h} \subseteq \mathfrak{gl}_3 \) consisting of all strictly upper triangular \( 3 \times 3 \)-matrices over \( \mathbb{k} \). Putting 
\[
x = e_{1,2}, \quad y = e_{2,3} \quad \text{and} \quad z = e_{1,3},
\]
where \( e_{i,j} \) denotes the matrix with 1 in the \( (i, j) \)-position and 0 elsewhere, we have 
\[
\mathfrak{h} = \mathbb{k}x \oplus \mathbb{k}y \oplus \mathbb{k}z.
\]
The Lie bracket of \( \mathfrak{h} \) is given by 
\[
[x, y] = z \quad \text{and} \quad [x, z] = [y, z] = 0.
\]
It is easy to check that \( \mathfrak{z} \mathfrak{h} = \mathbb{k}z \). If \( \text{char} \mathbb{k} = 2 \), then \( \mathfrak{h} \) is isomorphic to \( \mathfrak{sl}_2 \) but not otherwise (Exercise 5.1.4).

Other Linear Lie Algebras. Lie subalgebras of general linear Lie algebras are called linear Lie algebras. All examples presented above are of this kind. In fact, every finite-dimensional Lie algebra \( \mathfrak{g} \) is isomorphic to a Lie subalgebra of \( \mathfrak{gl}(V) \) for some finite-dimensional vector space \( V \) (Ado’s Theorem). For the proof of this result, see [26, §7], for example.

We mention the following linear Lie algebras; the verification that they are indeed Lie subalgebras of \( \mathfrak{gl}_n \) is straightforward in all cases:
\[
\mathfrak{n}_n(\mathbb{k}) \overset{\text{def}}{=} \{ \text{all strictly upper triangular } n \times n\text{-matrices over } \mathbb{k} \},
\]
\[
\mathfrak{t}_n(\mathbb{k}) \overset{\text{def}}{=} \{ \text{all upper triangular } n \times n\text{-matrices over } \mathbb{k} \},
\]
\[
\mathfrak{d}_n(\mathbb{k}) \overset{\text{def}}{=} \{ \text{all diagonal } n \times n\text{-matrices over } \mathbb{k} \}.
\]
As with \( \mathfrak{gl}_n \) and \( \mathfrak{sl}_n \), the base field \( \mathbb{k} \) is often omitted from the notation. Thus, \( \mathfrak{t}_n = \mathfrak{n}_n \oplus \mathfrak{d}_n \) as vector spaces, \( \mathfrak{n}_n \) is an ideal of \( \mathfrak{t}_n \), and \( \mathfrak{d}_n \) is an abelian Lie subalgebra of \( \mathfrak{t}_n \). The Lie algebra \( \mathfrak{n}_3 \) is the same as the Heisenberg Lie algebra \( \mathfrak{h} \).

5.1.5. Derivations
Let \( \mathfrak{M} \) be a \( \mathbb{k} \)-vector space that is equipped with a \( \mathbb{k} \)-bilinear “multiplication” map,
\[
\mathfrak{M} \times \mathfrak{M} \longrightarrow \mathfrak{M}
\]
\[
\psi \quad \psi
\]
\[
(a, b) \longmapsto ab
\]
For example, $\mathfrak{A}$ could be an associative algebra with the usual multiplication or a Lie algebra with the Lie bracket $[,]$ as multiplication. In any case, we may define a derivation of $\mathfrak{A}$ to be a $k$-linear map $\delta \in \text{End}_k(\mathfrak{A})$ satisfying the Leibniz product rule,

$$\delta(ab) = (\delta a)b + a(\delta b) \quad (a, b \in \mathfrak{A}).$$

If $\mathfrak{A} = A \in \text{Alg}_k$, then the Leibniz identity (5.1) says that, for any $a \in A$, the Lie commutator $[a, ,]$ is a derivation of $A$. For $\mathfrak{A} = g \in \text{Lie}_k$, we may rewrite the Jacobi identity in the following form, which shows that $\text{ad} x$ is a derivation of $g$:

$$(\text{ad} x).[y, z] = [(\text{ad} x).y, z] + [y, (\text{ad} x).z] \quad (x, y, z \in g).$$

For a general $\mathfrak{A}$, it is readily verified that $k$-linear combinations of derivations are again derivations. It is not hard to check that the derivations of $\mathfrak{A}$ do in fact form a Lie subalgebra of $\text{gl}(\mathfrak{A})$ (Problem 5.1.5); this Lie algebra will be denoted by

$$\text{Der} \mathfrak{A} \overset{\text{def}}{=} \{\text{all derivations of } \mathfrak{A}\}$$

In particular, for $\mathfrak{A} = g \in \text{Lie}_k$, the adjoint representation can be written in the following refined form:

$$\begin{align*}
\text{ad}: & \quad g \longrightarrow \text{Der} g \subseteq \text{gl}(g) \\
& \quad \psi \longmapsto \psi \quad \psi \\
& \quad x \longmapsto [x, ,]
\end{align*}$$

Thus, $\text{ad} g$ is a Lie subalgebra of $\text{Der} g$; it is called the Lie algebra of inner derivations of $g$. In fact, $\text{ad} g$ is an ideal of $\text{Der} g$ (Exercise 5.1.5). The Lie algebra $\text{Der} g/\text{ad} g$ is called the Lie algebra of outer derivations of $g$.

5.1.6. Actions and Semidirect Products

Let $a, b \in \text{Lie}_k$. An action of $b$ on $a$ is given by a Lie homomorphism,

$$b \longmapsto \text{Der} a.$$

For example, the adjoint representation (5.6) gives an action of $g$ on itself and on any of its ideals. As for groups, we will write $b \subseteq a$ to indicate an action of $b$ on $a$ and $b.a \in a$ will denote the image of $a \in a$ under the derivation given by $b \in b$. The action is called trivial if $b.a = 0$ for all $a \in a$ and $b \in b$.

Given an action $b \subseteq a$, we may define a Lie bracket on the cartesian product $a \times b \in \text{Vect}_k$ by using the Lie brackets of $a$ and $b$ along with the $b$-action on $a$:

$$[(a, b), (a', b')] := ([a, a'] + b.a' - b'.a, [b, b']).$$

A routine check shows that the Lie algebra axioms are satisfied. The resulting Lie algebra is called the semidirect product of $a$ and $b$ and denoted by $a \rtimes b$. 
The subspace $a \times 0$ is an ideal of $a \times b$ that is isomorphic to $a$, the subspace $0 \times b$ is a Lie subalgebra isomorphic to $b$, and $a \times b \cong a \oplus b$ as vector spaces. Conversely, if $g \in \text{Lie}_k$ decomposes as $g = a \oplus b$ for an ideal $a$ and a Lie subalgebra $b$, then $g \cong a \times b$, with action $b \subset a$ given by the restriction of the adjoint representation.

For the trivial action $b \subset a$, the above Lie bracket becomes $[(a, b), (a', b')] = ([a, a'], [b, b'])$. The semidirect product is then denoted by $a \times b$ and called the direct product of the Lie algebras $a$ and $b$.

**Example 5.8** (Lie algebras in dimension 2 and the Heisenberg Lie algebra). All linear endomorphisms of an abelian Lie algebra $a$ are in fact derivations. If $a = \mathbb{k}y$ is 1-dimensional, then there are two derivations up to scalar multiples: $d = 0$ and $d = \text{Id}$. Letting the 1-dimensional Lie algebra $b = \mathbb{k}x$ act on $a$ via $x \mapsto 0$, we obtain the 2-dimensional abelian Lie algebra, $\mathbb{k}y \times \mathbb{k}x$; the possibility $x \mapsto \text{Id}$ gives the 2-dimensional non-abelian Lie algebra:

$$\mathbb{k}y \times \mathbb{k}x \cong \mathbb{k}x \oplus \mathbb{k}y \quad \text{with} \quad [x, y] = y.$$ 

Starting with the 2-dimensional abelian Lie algebra $a = \mathbb{k}z \times \mathbb{k}y \cong \mathbb{k}z \oplus \mathbb{k}y$ and letting $b = \mathbb{k}x$ act on $a$ via $x \mapsto \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, we obtain the Heisenberg Lie algebra (5.5):

$$\mathfrak{h} = (\mathbb{k}z \times \mathbb{k}y) \rtimes \mathbb{k}x \cong \mathbb{k}x \oplus \mathbb{k}y \oplus \mathbb{k}z \quad \text{with} \quad [x, y] = z \text{ and } z \text{ central}.$$ 

**Example 5.9** (The diamond Lie algebra). The Heisenberg Lie algebra $\mathfrak{h} = \mathbb{k}x \oplus \mathbb{k}y \oplus \mathbb{k}z$ has a derivation $t \in \text{Der} \mathfrak{h}$ that is given by $t.x = -x$, $t.y = y$ and $t.z = 0$ (Exercise 5.1.5). This defines an action $\mathbb{k}x \subset \mathfrak{h}$. The semidirect product

$$\mathfrak{h} \rtimes \mathbb{k}t \cong \mathbb{k}x \oplus \mathbb{k}y \oplus \mathbb{k}z \oplus \mathbb{k}t$$

is called the **diamond Lie algebra**; its Lie bracket is determined by the relations

$$(5.7) \quad [x, y] = z, \quad [t, x] = -x \quad \text{and} \quad [t, y] = y$$

and the condition that $z$ is central.

### Exercises for Section 5.1

**5.1.1** (Ideals). Show that if $a$ and $b$ are ideals of a Lie algebra $g$, then so are $a + b$, $a \cap b$, $[a, b]$, and $\{ x \in g \mid [x, a] \subseteq b \}$.

**5.1.2** (Heisenberg Lie algebra). Recall that any $V \in \text{Rep} A_1(\mathbb{k})$ is given by a pair $(a_V, b_V) \in \text{End}_k(V)^2$ satisfying the relation $b_V a_V = a_V b_V + \text{Id}_V$ (Example 1.8). Show that $V$ becomes a representation of the Heisenberg Lie algebra $\mathfrak{h} = \mathbb{k}x \oplus \mathbb{k}y \oplus \mathbb{k}z$, with $[x, y] = z$ and $[x, z] = [y, z] = 0$, by sending $x \mapsto b_V, y \mapsto a_V$ and $z \mapsto \text{Id}_V$. Moreover, $V$ is irreducible for $A_1(\mathbb{k})$ if and only if $V$ is irreducible for $\mathfrak{h}$.

**5.1.3** (gl($V$) and sl($V$)). Let $V \in \text{Vect}_k$ with $n = \dim_k V < \infty$. Prove:

(a) If char $\mathbb{k} \neq 2$ or $n \neq 2$, then $sl(V) = [gl(V), gl(V)] = [sl(V), sl(V)]$.

(b) If char $\mathbb{k} \nmid n$, then $gl(V) \cong sl(V) \times \mathbb{k}Lie$ and $\mathcal{Z} sl(V) = 0$. 


5.1. Lie Algebra Basics

5.1.4 (Dimension 3). (a) Let \( \mathfrak{g} \in \text{Lie}_k \) be such that \( \dim_k \mathfrak{g} = 3 \) and \( [\mathfrak{g}, \mathfrak{g}] = \mathfrak{g} \). Show that the only ideals of \( \mathfrak{g} \) are 0 and \( \mathfrak{g} \) and that \( \mathfrak{g} \neq A_{\text{Lie}} \) for any \( A \in \text{Alg}_k \). In particular, this holds for \( \mathfrak{s}_3 \) if \( \text{char } k \neq 2 \).

(b) Show that \( \mathfrak{s}_3 \cong \mathfrak{h} \), the Heisenberg Lie algebra, if and only if \( \text{char } k = 2 \).

(c) Let \( x \times y = (x_2 y_3 - x_3 y_2, x_3 y_1 - x_1 y_3, x_1 y_2 - x_2 y_1) \) denote the cross product of \( x = (x_1, x_2, x_3) \), \( y = (y_1, y_2, y_3) \in \mathbb{k}^3 \). Show that the bracket \( [x, y] = x \times y \) makes \( \mathbb{k}^3 \) into a Lie algebra. [Use the formula \( x \times y \times z = (x \cdot z)y - (x \cdot y)z \) for \( x, y, z \in \mathbb{k}^3 \), where \( \cdot \) denotes the usual dot product of \( \mathbb{k}^3 \).]

(d) Assume that \( \text{char } k \neq 2 \) and that there is an element \( \lambda \in \mathbb{k} \) with \( \lambda^2 = -1 \). Show that the Lie algebra \( (\mathbb{k}^3, \times) \) in (c) is isomorphic to \( \mathfrak{s}_3 \).

5.1.5 (Derivations). (a) For \( \mathfrak{g} \in \text{Lie}_k \), check that \( \text{ad}: \mathfrak{g} \to \text{Der} \mathfrak{g} \) is a map in \( \text{Lie}_k \) and that \( [d, \text{ad } x] = \text{ad } d(x) \) for all \( d \in \text{Der} \mathfrak{g} \) and \( x \in \mathfrak{g} \). Thus, \( \text{ad } \mathfrak{g} \) is an ideal of \( \text{Der} \mathfrak{g} \).

(b) Show that the non-abelian 2-dimensional Lie algebra has no outer derivations and neither does \( \mathfrak{s}_3 \) if \( \text{char } k \neq 2 \), but the endomorphism \( t \) in Example 5.9 is an outer derivation of the Heisenberg Lie algebra.

(c) For \( \mathfrak{A} \) as in §5.1.5, check that \( \text{Der} \mathfrak{A} \) is a Lie subalgebra of \( \mathfrak{gl}(\mathfrak{A}) \). For \( a, b \in \mathfrak{A} \), \( \lambda, \mu \in \mathbb{k} \) and \( d \in \text{Der} \mathfrak{A} \), prove the Leibniz formula,

\[
(d - (\lambda + \mu))^{n}(ab) = \sum_{i=0}^{n} \binom{n}{i} (d - \lambda)^{i}(a)(d - \mu)^{n-i}(b).
\]

(d) For \( A \in \text{Alg}_k \), show that \( \text{Der} A \subseteq \text{Der}(A_{\text{Lie}}) \).

5.1.6 (Witt algebra). Assume that \( \text{char } k = 0 \) and put \( w = \text{Der} \mathbb{k}[x^\pm 1] \). The Lie algebra \( w \) is called the Witt algebra.

(a) Show that the derivations \( e_n = x^{n+1} \frac{d}{dx} \) \((n \in \mathbb{Z})\) form a \( k \)-basis of \( w \).

(b) Show that \( [e_n, e_m] = (m - n)e_{n+m} \). Conclude that \( e_{-1} \) and \( e_2 \) generate the Lie algebra \( w \).

5.1.7 (Diamond algebra). Let \( \mathfrak{b} = k \times \mathfrak{t} \) be the diamond Lie algebra (Example 5.9).

(a) Show that \( \mathfrak{b} \) is isomorphic to a Lie subalgebra of \( \mathfrak{t}_3 \).

(b) Show that the following diagram is the complete lattice of ideals of \( \mathfrak{b} \):

\[
\begin{array}{c}
\mathfrak{b} \\
\downarrow \\
[\mathfrak{b}, \mathfrak{b}] \\
\downarrow \\
\mathbb{k}x \oplus \mathbb{k}z \\
\downarrow \\
\mathbb{k}y \oplus \mathbb{k}z \\
\downarrow \\
\mathbb{k}z = \mathbb{Z} \mathfrak{b} \\
\downarrow \\
0
\end{array}
\]
5.2. Types of Lie Algebras

5.2.1. Nilpotent and Solvable Lie Algebras

The Lie theoretic versions of nilpotency and solvability are analogous to the corresponding notions from group theory, which are presumably familiar to the reader. We will therefore be rather brief below.

**Nilpotency.** A Lie algebra \( \mathfrak{g} \) is called **nilpotent** if there exists a series of \( k \)-subspaces

\[
\mathfrak{g} = \mathfrak{g}_0 \supseteq \mathfrak{g}_1 \supseteq \mathfrak{g}_2 \supseteq \cdots \supseteq \mathfrak{g}_k = 0
\]

such that \( [\mathfrak{g}, \mathfrak{g}_i] \subseteq \mathfrak{g}_{i+1} \) for all \( i \). This condition states that all \( \mathfrak{g}_i \) are in fact ideals of \( \mathfrak{g} \) and \( \mathfrak{g}_i / \mathfrak{g}_{i+1} \subseteq \mathfrak{z}(\mathfrak{g} / \mathfrak{g}_{i+1}) \).

In order to test for the existence of a series (5.8), it suffices to consider either the **descending central series** \( \{ C^i \mathfrak{g} \}_{i \geq 0} \) or the **ascending central series** \( \{ C^i \mathfrak{g} \}_{i \geq 0} \); these series are defined by

\[
C^0 \mathfrak{g} = \mathfrak{g}, \quad C^{i+1} \mathfrak{g} = [\mathfrak{g}, C^i \mathfrak{g}]
\]

and

\[
C^0 \mathfrak{g} = 0, \quad C_{i+1} \mathfrak{g} = \{ x \in \mathfrak{g} | [x, \mathfrak{g}] \subseteq C^i \mathfrak{g} \}.
\]

An easy induction using Exercise 5.1.1 shows that all \( C^i \mathfrak{g} \) and \( C_i \mathfrak{g} \) are ideals of \( \mathfrak{g} \).

An easy induction using Exercise 5.1.1 shows that all \( C^i \mathfrak{g} \) and \( C_i \mathfrak{g} \) are ideals of \( \mathfrak{g} \).

In particular, the series \( \{ C^i \mathfrak{g} \}_{i \geq 0} \) is indeed descending: \( C^i \mathfrak{g} \supseteq C^{i+1} \mathfrak{g} \) for all \( i \); and \( \{ C_i \mathfrak{g} \}_{i \geq 0} \) is ascending. Note also that \( C_{i+1} \mathfrak{g} / C_i \mathfrak{g} = \mathfrak{z}(\mathfrak{g} / C_i \mathfrak{g}) \).

**Lemma 5.10.** The following are equivalent for any \( \mathfrak{g} \in \text{Lie}_k \):

(i) \( \mathfrak{g} \) is nilpotent, having a series (5.8); (ii) \( C^k \mathfrak{g} = 0 \); (iii) \( C_k \mathfrak{g} = \mathfrak{g} \).

**Proof.** For (i) \( \Rightarrow \) (ii), one shows by a straightforward induction that \( \mathfrak{g}_i \supseteq C^i \mathfrak{g} \) holds for all \( i \). Conversely, (ii) implies that \( \mathfrak{g}_i := C^i \mathfrak{g} \) yields a series of the form (5.8). Thus, (i) and (ii) are equivalent. In order to prove the equivalence of (ii) and (iii), observe the following equivalences, using the definition of \( C^{k-i} \mathfrak{g} \) for the first equivalence and the definition of \( C_{i+1} \mathfrak{g} \) for the second: \( C^{k-i} \mathfrak{g} \subseteq C_i \mathfrak{g} \iff [\mathfrak{g}, \mathfrak{g}^{k-i-1} \mathfrak{g}] \subseteq C_i \mathfrak{g} \iff C^{k-i-1} \mathfrak{g} \subseteq C_{i+1} \mathfrak{g} \). Consequently, if \( \mathfrak{g}^{k-i} \subseteq C_i \mathfrak{g} \) holds for some \( i \), then it also holds for all subsequent \( i \) and for all previous \( i \). Now, (ii) says that the inclusion holds for \( i = 0 \); hence it also holds for \( i = k \), giving (iii). Similarly, (iii) implies (ii). \( \square \)

If \( \mathfrak{g} \) is nilpotent, then the smallest \( k \) such that the equivalent conditions of Lemma 5.10 are satisfied is called the **nilpotency class** of \( \mathfrak{g} \). Thus, class 0 means \( \mathfrak{g} = 0 \) and class 1 means that \( \mathfrak{g} \) is nonzero abelian. The Heisenberg Lie algebra is nilpotent of class 2. More generally, the Lie algebra of \( n \times n \) matrices over \( k \) is nilpotent of class \( n - 1 \) (Exercise 5.2.1).
Solvability. The Lie algebra \( g \) is called **solvable** if there is a series of \( k \)-subspaces (5.9)
\[
g = g_0 \supseteq g_1 \supseteq g_2 \supseteq \cdots \supseteq g_k = 0
\]
such that \([g_i, g_j] \subseteq g_{i+j+1}\) for all \( i \). Equivalently, each \( g_i \) is a Lie subalgebra of \( g \) and \( g_{i+1} \) is an ideal of \( g_i \) such that \( g_i / g_{i+1} \) is an abelian Lie algebra.

The standard test series for solvability is the so-called **derived series**, which is defined by
\[
D^0 g = g, \quad D^{i+1} g = [D^i g, D^i g].
\]
As for the two central series considered above, one shows that the derived series consists of ideals of \( g \). Moreover, the proof of the equivalence \((i) \iff (ii)\) in Lemma 5.10 carries over verbatim to give the following criterion.

**Lemma 5.11.** \( g \) is solvable, having a series (5.9), if and only if \( D^k g = 0 \).

The class of solvable Lie algebras clearly includes all nilpotent Lie algebras, but it is in fact far more extensive. For example, the 2-dimensional non-abelian Lie algebra, the diamond algebra and the Lie algebras \( t_n \) of all upper triangular matrices are readily checked to be solvable but not nilpotent (Exercise 5.2.4).

**Some Properties of Nilpotent and Solvable Lie Algebras.** The first proposition below addresses the stability properties. By part (b), solvability is very well-behaved under extensions, whereas nilpotency is a more fragile property.

**Proposition 5.12.**
(a) Homomorphic images and Lie subalgebras of solvable Lie algebras are again solvable; likewise for nilpotent Lie algebras.
(b) Let \( a \) be an ideal of the Lie algebra \( g \). If \( a \cap g/a \) are solvable, then so is \( g \). For nilpotency, the corresponding statement holds provided \( a \subseteq Z g \).

**Proof.** (a) Given an epimorphism \( f : g \rightarrow \overline{g} \) in \( \text{Lie}_k \) and a series \( \{g_i\} \) of type (5.8) or (5.9) for \( g \), the images \( \{f g_i\} \) form a series of the corresponding type for \( \overline{g} \). Similarly, if \( h \subseteq g \) is a Lie subalgebra, then the series \( \{h \cap g_i\} \) works for \( h \).

(b) Splicing together series \( \{a_i\} \) and \( \{\overline{g}_i\} \) of type (5.9) for \( a \) and \( \overline{g} = g / a \), we obtain the series \( g = g_0 \supseteq \cdots \supseteq g_k = a = a_0 \supseteq \cdots \supseteq a_l = 0 \), where \( g_j \) denotes the preimage of \( \overline{g}_j \) in \( g \). This series is readily checked to be of type (5.9) again, which proves the first claim. The same process, with \( l = 1 \), works for series of type (5.8) under the given proviso.

**Proposition 5.13.**
(a) If \( g \in \text{Lie}_k \) is nilpotent, then \( a \cap \mathcal{Z} g \neq 0 \) for every ideal \( 0 \neq a \) of \( g \) and \( h \subseteq N(a) \) holds for every Lie subalgebra \( h \subseteq g \).
(b) If \( a \) is a nonzero solvable ideal of an arbitrary \( g \in \text{Lie}_k \), then \( a \) contains a nonzero abelian ideal of \( g \).

**Proof.** (a) Let \( \{g_i\} \) be as in (5.8) and let \( t \) be chosen maximal so that \( a \cap g_t \neq 0 \). Then \([g, a \cap g_t] \subseteq a \cap g_{t+1} = 0\) and so \( a \cap g_t \subseteq a \cap \mathcal{Z}(g) \). This proves the assertion
concerning ideals. As for subalgebras, we clearly have \( h \subseteq N_{g}(h) \). Choosing \( s \) maximal so that \( g_{s} \not\subseteq h \), we obtain \([g_{s}, h] \subseteq g_{s+1} \subseteq h\) and so \( g_{s} \subseteq N_{g}(h) \).

(b) The last nonzero term of the derived series \( \mathcal{D}^{l}a \) of \( a \) is abelian, and all \( \mathcal{D}^{l}a \) are easily seen to be ideals of \( g \) using Exercise 5.1.1 and induction. \( \square \)

5.2.2. Simple and Semisimple Lie Algebras

It is an immediate consequence of Proposition 5.12(b) that the sum of any two solvable ideals of any Lie algebra \( g \) is again a solvable ideal of \( g \). Therefore, assuming \( g \) to be finite-dimensional, there is a unique largest solvable ideal of \( g \) that contains all others; this ideal is called the **radical** of \( g \):

\[
\text{rad } g \overset{\text{def}}{=} \text{the sum of all solvable ideals of } g
\]

A finite-dimensional Lie algebra \( g \) such that \( \text{rad } g = 0 \) is called **semisimple**. Any finite-dimensional \( g \in \text{Lie}_k \) is “solvable-by-semisimple:” \( \text{rad } g \) is solvable and the quotient \( g/\text{rad } g \) is semisimple by Proposition 5.12(b). Note also that, by Proposition 5.13(b), \( \text{rad } g = 0 \) is equivalent to the condition that \( g \) has no nonzero abelian ideals.

A finite-dimensional \( g \in \text{Lie}_k \) is called **simple** if \( g \) is non-abelian and the only ideals of \( g \) are 0 and \( g \). Thus, simple Lie algebras are certainly semisimple. Since abelian Lie algebras \( g \) whose only ideals are \( g \) and 0 are exactly those of dimension \( \leq 1 \), the non-abelian hypothesis in the definition can be replaced by the requirement that \( \dim_k g \geq 2 \). The condition that 0 and \( g \) are the only ideals of \( g \), and distinct, is equivalent to irreducibility of the adjoint representation \( g_{ad} \).

Simple and semisimple Lie algebras will be studied in detail in Chapter 6. In particular, the reason for the name “semisimple” will be revealed there. For now, we content ourselves by offering the following example; see also Exercise 5.2.6.

**Example 5.14** (Simplicity of \( \mathfrak{sl}_2 \)). Let \( \text{char } k \neq 2 \). Recall from §5.1.4 that \( \mathfrak{sl}_2 = \mathbb{k}f \oplus \mathbb{k}h \oplus \mathbb{k}e \), with Lie bracket given by \([h, f] = -2f, [h, e] = 2e \) and \([e, f] = h \). The operator \( \text{ad } h \in \text{Der } \mathfrak{sl}_2 \) is diagonalizable, with eigenvalues \(-2, 0 \) and 2 and corresponding eigenspaces \( \mathbb{k}f, \mathbb{k}h \) and \( \mathbb{k}e \). Thus, \( \mathfrak{sl}_2 \) is a completely reducible representation of the polynomial algebra \( \mathbb{k}[x] \), with \( x \) acting via \( \text{ad } h \), and the above eigenspaces are the homogeneous components (cf. Example 1.27). Any nonzero ideal \( a \) of \( \mathfrak{sl}_2 \) is a subrepresentation, and hence \( a \) must contain one of the homogeneous components. Successive application of \( \text{ad } e \) or \( \text{ad } f \) then yields that \( a \) contains all of \( \mathfrak{sl}_2 \), proving simplicity of \( \mathfrak{sl}_2 \). See also Exercise 5.1.4 for a different argument.

\[\text{The corresponding fact for nilpotent ideals is also true but less obvious; see Exercise 5.2.2.}\]

\[\text{If char } k = 0, \text{ then any finite-dimensional } g \in \text{Lie}_k \text{ is the semidirect product of rad } g \text{ and a semisimple Lie subalgebra } s \subseteq g, \text{ which is essentially unique [26, §6 Thm. 5]. The decomposition } g = \text{rad } g \ltimes s \text{ is called the Levi decomposition of } g.\]
Exercises for Section 5.2

5.2.1 (Nilpotency of $n_n$). Show that that the descending central series of the Lie algebra $n_n$ of all strictly upper triangular $n \times n$-matrices over $k$ is given by $C^i n_n = \{(\lambda_{j,k}) \in gl_n | \lambda_{j,k} = 0 \text{ for } k - j \leq i \}$. In particular, $n_n$ is nilpotent of class $n - 1$.

5.2.2 (Nilradical). Let $a$ and $b$ be ideals of a Lie algebra $g$. Show that the terms of the descending central series $C^i a$ are ideals of $g$ and that $C^{i+1} (a + b) \subseteq C^i a + C^i b$. Conclude that if $a$ and $b$ are nilpotent, then so is the ideal $a + b$. Consequently, any finite-dimensional Lie algebra $g$ has a unique largest nilpotent ideal, namely the sum of all nilpotent ideals; this ideal is called the nilradical of $g$.

5.2.3 (Outer derivations). Let $g \in \text{Lie}_k$ be finite dimensional and assume that $g, [g, g]$ and $Z g, 0$. Show that $g$ has an outer derivation by completing the following steps; this result is from [198].

(a) Show that $g$ has an ideal $a$ of codimension 1 and that, for any such $a$, the centralizer $C := C_g(a)$ satisfies $[g, C] \subseteq C$.

(b) With $a$ as in (a), write $g = a \oplus k b$ and choose $c \in C \setminus [g, C]$. Show that the endomorphism of $g$ that vanishes on $a$ and sends $b$ to $c$ is an outer derivation of $g$.

5.2.4 (Non-nilpotent solvable Lie algebras). Let $g$ be the non-abelian Lie algebra of dimension 2, and let $b$ be the diamond algebra (5.7).

(a) Show that $g$ and $b$ are both solvable but not nilpotent. Likewise for the Lie algebras $t_n$ of upper triangular $n \times n$-matrices over $k$.

(b) Construct Lie epimorphisms $b \rightarrow g$ and $t_n \rightarrow g$.

5.2.5 (Lie algebras and field extensions). Let $g \in \text{Lie}_k$ and let $K/k$ be a field extension. The unique $K$-bilinear extension of the Lie bracket of $g$ makes $g \otimes K$ becomes a Lie algebra over $K$. Show that $g$ is nilpotent if and only if $g \otimes K$ is nilpotent; similarly for solvability.

5.2.6 (Simplicity of $sl_n$). Assuming that $\text{char} k \neq 2$ and $\text{char} k \nmid n \geq 2$, show that $sl_n$ is a simple Lie algebra.

5.3. Three Theorems about Linear Lie Algebras

The image of any representation $g \rightarrow gl(V)$ of a Lie algebra $g$ is a Lie subalgebra $g_V \subseteq gl(V)$. In this section, we will consider some structural properties of Lie subalgebras of $gl(V)$ for a finite-dimensional $V \in \text{Vect}_k$.

5.3.1. Engel’s Theorem

The term “nilpotent,” when applied to an element $x$ of an associative algebra, means that some power $x^n = 0$. It is an elementary fact from linear algebra that, for any nilpotent operator $f \in \text{End}_k(V)$ on a finite-dimensional $k$-vector space $V$,
there exists a basis of $V$ such that the matrix of $f$ has strict upper triangular form. For background, we mention two generalizations and variations of this fact; see Kaplansky [117, pages 100 and 135] for complete details. Let $S \subseteq \text{End}_k(V)$ be a multiplicative subsemigroup. Then:

- If $S$ consists of nilpotent operators, then $V$ has a basis such that the matrices of all elements of $S$ have strict upper triangular form (Levitzki’s Theorem).
- If all $s \in S$ are unipotent, that is, $s - \text{Id}_V$ is nilpotent, then there exists a basis of $V$ such that the matrices of all elements of $S$ have upper triangular form with 1s on the diagonal (Kolchin’s Theorem).

Engel’s Theorem$^4$ is a Lie theoretic variant of these results. The following simple observation will be useful in connecting associative nilpotency with Lie nilpotency.

**Lemma 5.15.** Let $A \in \text{Alg}_k$ and let $x \in A$ be nilpotent. Then the Lie commutator $[x, \cdot] \in \text{End}_k(A)$ is nilpotent.

**Proof.** Note that $[x, \cdot] = l_x - r_x$, where $l_x, r_x \in \text{End}_k(A)$ denote left and right multiplication with $x$, respectively. Both $l_x$ and $r_x$ are nilpotent and $l_x r_x = r_x l_x$ by associativity. The binomial theorem therefore shows that $l_x - r_x$ is nilpotent. □

Given a representation $\mathfrak{g} \rightarrow \mathfrak{gl}(V)$, we define the space of $\mathfrak{g}$-invariants in $V$ by

$$V^\mathfrak{g} \overset{\text{def}}{=} \{ v \in V \mid x.v = 0 \text{ for all } x \in \mathfrak{g} \}$$

**Engel’s Theorem.** Let $\mathfrak{g} \in \text{Lie}_k$ and let $V$ be a finite-dimensional representation of $\mathfrak{g}$ such that $x_V \in \text{End}_k(V)$ is nilpotent for all $x \in \mathfrak{g}$. Then $V$ has a $k$-basis such that the matrices of all $x_V$ are strictly upper triangular. In particular, the Lie algebra $\mathfrak{g}_V \subseteq \mathfrak{gl}(V)$ is nilpotent.

**Proof.** In view of nilpotency of the Lie algebra of all strictly upper triangular matrices (Exercise 5.2.1), nilpotency of $\mathfrak{g}_V$ will follow from the asserted basis of $V$. For the construction of this basis, we proceed by induction on $\dim_k V$. The case $V = 0$ being vacuous, we may assume that $V \neq 0$. It suffices to show that

$$V^\mathfrak{g} \neq 0.$$ (5.10)

For, then the quotient representation $\overline{V} = V/V^\mathfrak{g}$ has smaller dimension and, clearly, all $x_V$ are nilpotent. By induction, $\overline{V}$ has a basis of the required form. The preimages in $V$ of this basis together with any basis of $V^\mathfrak{g}$ will then yield a basis of $V$ of the kind that we are looking for.

$^4$The result is due to Friedrich Engel (1861 -1941).
For the proof of (5.10) we may replace \( \mathfrak{g} \) by \( \mathfrak{g}_V \). Thus, \( \mathfrak{g} \subseteq \mathfrak{gl}(V) \) and so \( \mathfrak{g} \) is finite dimensional. Since nilpotent operators have only zero eigenvalues, (5.10) is obvious if \( \dim_k \mathfrak{g} \leq 1 \). So let us assume that \( \dim_k \mathfrak{g} > 1 \) and that (5.10) holds for all Lie algebras of smaller dimension. The normalizer \( N_{\mathfrak{g}}(\mathfrak{h}) = \{ x \in \mathfrak{g} | [x, \mathfrak{h}] \subseteq \mathfrak{h} \} \) of every Lie subalgebra \( \mathfrak{h} \subseteq \mathfrak{g} \) is a Lie subalgebra of \( \mathfrak{g} \) (Example 5.4) such that \( \mathfrak{h} \subseteq N_{\mathfrak{g}}(\mathfrak{h}) \).

**Claim.** If \( \mathfrak{h} \not\subseteq \mathfrak{g} \), then \( \mathfrak{h} \not\subseteq N_{\mathfrak{g}}(\mathfrak{h}) \).

To prove this, note that \( [x, \cdot] \in \mathrm{End}_k(\mathfrak{gl}(V)) \) is nilpotent for all \( x \in \mathfrak{g} \) by Lemma 5.15. Hence, \( \mathrm{ad} x = [x, \cdot] \) is nilpotent as well. If \( x \in \mathfrak{h} \), then \( \mathrm{ad} x \) stabilizes \( \mathfrak{h} \) and so \( \mathrm{ad} x \) yields an endomorphism \( \mathrm{ad}_{\mathfrak{g}/\mathfrak{h}} x \) of \( \mathfrak{g}/\mathfrak{h} \) which is clearly nilpotent. Since (5.10) holds for \( \mathfrak{h} \) by induction, we conclude that \( (\mathfrak{g}/\mathfrak{h})_0 \). This is equivalent to the claim.

Now we can complete the proof of the theorem as follows. Choose a maximal Lie subalgebra \( \mathfrak{h} \subseteq \mathfrak{g} \). By the claim, \( \mathfrak{h} \) is in fact an ideal of \( \mathfrak{g} \). Moreover, \( \mathfrak{h}, 0 \), because any \( 1 \)-dimensional subspace of \( \mathfrak{g} \) is a proper Lie subalgebra of \( \mathfrak{g} \). By induction, \( V^\mathfrak{h} \neq 0 \). Furthermore, \( x.V^\mathfrak{h} \subseteq V^\mathfrak{h} \) holds for all \( x \in \mathfrak{g} \): if \( y \in \mathfrak{h} \) and \( v \in V^\mathfrak{h} \), then

\[
y.x.v = x.y.v + [y, x].v = 0
\]

because \( y.v \) and \( [y, x].v \) both belong to \( \mathfrak{h}, v = \{0\} \). Therefore, the Lie algebra \( \mathfrak{g}/\mathfrak{h} \) acts on \( V^\mathfrak{h} \), clearly by nilpotent operators. By induction, \( (V^\mathfrak{h})_{0/\mathfrak{h}} \neq 0 \). Inasmuch as \( (V^\mathfrak{h})_{0/\mathfrak{h}} = V^\mathfrak{h} \), this proves (5.10) and so the theorem is proved.

We note that the sufficient condition for nilpotency of linear Lie algebras given in Engel’s Theorem is far from necessary. In fact, the Lie algebra \( \mathfrak{d}_n \) of all diagonal \( n \times n \)-matrices over \( \mathbb{R} \) is abelian, and hence certainly nilpotent, but \( \mathfrak{d}_n \) plays a very different role from the Lie algebra \( \mathfrak{n}_n \) of all strictly upper triangular matrices that is addressed Engel’s Theorem.

The following corollary, which is also often referred to as Engel’s Theorem, characterizes nilpotency for finite-dimensional Lie algebras.

**Corollary 5.16.** A finite-dimensional Lie algebra \( \mathfrak{g} \) is nilpotent if and only if the operators \( \mathrm{ad} x \in \mathrm{End}_k(\mathfrak{g}) \) are nilpotent for all \( x \in \mathfrak{g} \).

**Proof.** Since the descending central series \( \{ \mathfrak{c}^k \}_{k=0} \) of any nilpotent \( \mathfrak{g} \in \mathfrak{Lie}_k \) satisfies \( \mathfrak{c}^k \mathfrak{g} = 0 \) for some \( k \), it follows that \( (\mathrm{ad} x)^k = 0 \) for all \( x \in \mathfrak{g} \). Conversely, if all \( \mathrm{ad} x \) are nilpotent, then the Lie algebra \( \mathrm{ad} \mathfrak{g} \subseteq \mathfrak{gl}(\mathfrak{g}) \) is nilpotent by Engel’s Theorem. Since \( \mathrm{ad} \mathfrak{g} \cong \mathfrak{g}/\mathfrak{Z} \mathfrak{g} \) by (5.2), Proposition 5.12(b) further gives nilpotency of \( \mathfrak{g} \).
5.3.2. Lie’s Theorem

If \( V \) is a finite-dimensional representation of \( g \in \text{Lie}_k \), then any subrepresentation \( 0 \subseteq U \subseteq V \) gives rise, in the familiar fashion, to a \( k \)-basis of \( V \) such that the matrices of all operators \( x_V \) with \( x \in g \) have block upper triangular form,

\[
\begin{pmatrix}
* & * \\
0 & *
\end{pmatrix}
\]

A chain \( 0 = V_0 \subseteq V_1 \subseteq \cdots \subseteq V_k = V \) of subrepresentations of \( V \) is called a **complete flag** if \( \dim_k V_i = i \) for all \( i \). The existence of a complete flag of subrepresentations is equivalent to the existence of a \( k \)-basis of \( V \) such that the matrices of all \( x_V \) are upper triangular.

**Lie’s Theorem.** Let \( g \in \text{Lie}_k \) be solvable and assume that \( k \) is algebraically closed with \( \text{char } k = 0 \). Then:

(a) All finite-dimensional irreducible representations of \( g \) are 1-dimensional.

(b) Every finite-dimensional representation of \( g \) has a complete flag of subrepresentations.

Note that (b) follows from (a) by induction on the dimension of the given representation, say \( V \). Indeed, (a) handles the case where \( V \) is irreducible. If there is a subrepresentation \( 0 \subseteq U \subseteq V \), then splicing together complete flags of subrepresentations for \( U \) and for \( V/U \), which we may assume to exist by induction, we obtain the desired complete flag for \( V \). For the proof of (a), we isolate the main technical point in a separate lemma; it elaborates on the following easy fact that was established towards the end of the proof of Engel’s Theorem: for any representation \( V \) of an arbitrary \( g \in \text{Lie}_k \) and any ideal \( a \) of \( g \), the \( a \)-invariants \( V^a \) are a subrepresentation of \( V \). In place of invariants, the lemma considers more general **weight spaces** that are associated to Lie algebra homomorphisms \( \lambda: a \to \mathbb{K}_{\text{Lie}} \):

\[
V_\lambda = \{ v \in V \mid x.v = \lambda(x)v \text{ for all } x \in a \}
\]

**Lemma 5.17.** Let \( g \in \text{Lie}_k \) and assume that \( \text{char } k = 0 \). Let \( V \) be a finite-dimensional representation of \( g \), let \( a \) be an ideal of \( g \), and let \( \lambda: a \to \mathbb{K}_{\text{Lie}} \) be a Lie algebra homomorphism. Then the weight space \( V_\lambda \) is a subrepresentation of \( V \).

**Proof.** Let \( x \in g \) and \( v \in V_\lambda \). We need to show that \( y.x.v = \lambda(y)x.v \) for all \( y \in a \), and we may clearly assume that \( v \neq 0 \). Using the fact that \( [y, x] \in a \) we compute

\[
y.x.v = x.y.v + [y, x].v = \lambda(y)x.v + \lambda([y, x])v.
\]

Thus, the issue is to show that \( \lambda([y, x]) = 0 \) for all \( x \in g \) and \( y \in a \).
5.3. Three Theorems about Linear Lie Algebras

Fix \( x \in \mathfrak{g} \) and put \( v_j := x^j(v) \) and \( U_i = \sum_{j=0}^{i-1} \mathbb{K} v_j \). This gives a chain of subspaces of \( V \),

\[
0 = U_0 \subseteq U_1 = \mathbb{K} v \subseteq \cdots \subseteq U_i \subseteq U_{i+1} = U_i + \mathbb{K} v_i \subseteq \cdots \subseteq U := \bigcup_i U_i,
\]
such that \( x.U_i \subseteq U_{i+1} \) for all \( i \) and \( x.U \subseteq U \). Since \( V \) is finite dimensional, we may choose \( m \) to be minimal such that \( U_m = U_{m+1} \). Then \( \dim_k U_i = i \) for all \( i \leq m \) and \( U_i = U \) for \( i \geq m \). Thus, \( v = v_0, v_1, \ldots, v_{m-1} \) is a basis of \( U \).

Claim. \( U \) is stable under the action of every \( y \in \mathfrak{a} \) and \( \text{trace}(y_U) = m \lambda(y) \).

To prove this, we will show that \( y.v_i \equiv \lambda(y)v_i \mod U_i \) for all \( i \); this will imply that \( y.U \subseteq U \) and that the matrix of \( y_U \) with respect to the basis \( (v_i)_{0}^{m-1} \) of \( U \) has the form \( \lambda(y)1_{m \times m} + N \) with \( N \) strictly upper triangular, giving the claimed trace. To start, \( y.v_0 = \lambda(y)v_0 \) for all \( y \in \mathfrak{a} \), because \( v_0 = v \in V_1 \). So let \( i > 0 \) and assume that \( y.v_j \equiv \lambda(y)v_j \mod U_j \) for all \( j < i \) and all \( y \in \mathfrak{a} \). Then \( U_i \) is stable under the operators in \( \mathfrak{a}_V \). In particular, \( [y,x].v_{i-1} \in U_i \) and \( y.v_{i-1} = \lambda(y)v_{i-1} \mod U_{i-1} \). Therefore, \( x.y.v_{i-1} - \lambda(y)x.v_{i-1} \in xU_{i-1} \subseteq U_i \) and so

\[
y.v_i = x.y.v_{i-1} - \lambda(y)x.v_{i-1} = \lambda(y)v_i \mod U_i,
\]
proving the claim.

Since \( [y,x] \in \mathfrak{a} \), the Claim gives \( \text{trace}([y,x]_U) = m \lambda([y,x]) \). Furthermore, since \( U \) is stable under the actions of both \( x \) and \( y \), we also have \( \text{trace}([y,x]_U) = \text{trace}(y_U x_U) - \text{trace}(x_U y_U) = 0 \). Thus \( m \lambda([y,x]) = 0 \), which forces \( \lambda([y,x]) = 0 \) by our hypothesis on \( \text{char} \mathbb{K} \). This completes the proof of the lemma. \( \square \)

With the lemma in hand, we are now ready to prove Lie’s Theorem.

Proof of Lie’s Theorem. As we have already remarked, it suffices to prove (a). Thus, we may assume that \( V \) is a finite-dimensional irreducible representation of \( \mathfrak{g} \) and we need to show that \( \dim_k V = 1 \).

In proving this, we may replace \( \mathfrak{g} \) by \( \mathfrak{g}_V \subseteq \mathfrak{gl}(V) \), thereby reducing to the case where \( \mathfrak{g} \) is finite dimensional. The case \( \mathfrak{g} = 0 \) being trivial, we may assume that \( \dim_k \mathfrak{g} \geq 1 \) and proceed by induction on \( \dim_k \mathfrak{g} \). Note that \( \mathfrak{g} \) has an ideal \( \mathfrak{a} \) of codimension 1. Indeed, \( [\mathfrak{g}, \mathfrak{g}] \subseteq \mathfrak{g} \) by solvability of \( \mathfrak{g} \), and any subspace of \( \mathfrak{g} \) containing \( [\mathfrak{g}, \mathfrak{g}] \) is automatically an ideal. So we may choose \( \mathfrak{a} \) to be any codimension-1 subspace of \( \mathfrak{g} \) containing \( [\mathfrak{g}, \mathfrak{g}] \). View \( \mathfrak{a} \) as a Lie algebra in its own right and choose an irreducible \( \mathfrak{a} \)-subrepresentation \( U \subseteq V \); this surely exists as \( V \) is finite dimensional. Then \( \dim_k U = 1 \) by induction, because \( \mathfrak{a} \) is solvable (Proposition 5.12). Thus, the representation \( U \) amounts to a homomorphism of Lie algebras \( \lambda: \mathfrak{a} \rightarrow \mathfrak{gl}(U) \equiv \mathfrak{gl}_k \). Thus, using Lemma 5.17 and its notation, \( U \subseteq V_1 \) and \( V_1 \) is a subrepresentation of \( V \). Write \( \mathfrak{g} = \mathfrak{a} \oplus \mathbb{K} x \) for some \( x \in \mathfrak{g} \) and choose an eigenvector \( 0 \neq v \in V_\lambda \) for \( x \). Then \( v \) is a common eigenvector for all of \( \mathfrak{g} \) and
so \( k \mathfrak{v} \) is a subrepresentation of \( V \). By irreducibility, we conclude that \( V = k \mathfrak{v} \) as desired.

Unsurprisingly, Lie’s Theorem fails if \( k \) is not algebraically closed, as it may not be possible to find the requisite eigenvalues in \( k \) then. It is perhaps more surprising that Lie’s Theorem also fails in positive characteristics.

**Example 5.18** (Failure of Lie’s Theorem in characteristic \( p > 0 \)). Irreducible representation of the Weyl algebra \( A_1(\mathbb{k}) \) can be viewed as irreducible representations of the Heisenberg Lie algebra \( \mathfrak{h} \) (Exercise 5.1.2). For \( \mathbb{k} \) algebraically closed with \( \text{char} \, \mathbb{k} = p > 0 \), all irreducible representations of \( A_1(\mathbb{k}) \) have dimension \( p \) (Exercise 1.2.9). Since \( \mathfrak{h} \) is nilpotent, hence solvable, these representations give rise to counterexamples to Lie’s Theorem in characteristic \( p \).

Since \( A_1(\mathbb{k}) \) has no nonzero finite-dimensional representations if \( \text{char} \, \mathbb{k} = 0 \) (Example 1.8), the above example also points to the fact that many irreducible representations of solvable (or even nilpotent) Lie algebras are infinite dimensional in characteristic 0. For other examples and counterexamples, see the Exercises.

### 5.3.3. Jordan Canonical Form

Before proceeding to the last main result of this section, we recall some facts from linear algebra. Let \( V \) be a finite-dimensional vector space over an algebraically closed field \( \mathbb{k} \) and let \( \phi \in \text{End}_\mathbb{k}(V) \). Then, besides the ordinary eigenspaces \( V_\lambda = \{ v \in V \mid \phi(v) = \lambda v \} \) for \( \lambda \in \mathbb{k} \), there are the **generalized eigenspaces**

\[
V^\lambda \overset{\text{def}}{=} \{ v \in V \mid (\phi - \lambda)^t v = 0 \text{ for some } t \geq 0 \}
\]

Thus, \( V_\lambda \subseteq V^\lambda \) and \( V^\lambda \neq 0 \) if and only if \( V_\lambda \neq 0 \). Furthermore, \( \dim_\mathbb{k} V^\lambda \) is equal to the multiplicity of \( \lambda \) as a root of the characteristic polynomial of \( \phi \) and

\[
V = \bigoplus_{\lambda \in \mathbb{k}} V^\lambda.
\]

The endomorphism \( \phi \) is said to be **diagonalizable** if \( V \) has a \( \mathbb{k} \)-basis consisting of \( \phi \)-eigenvectors. This happens if and only if \( V_\lambda = V^\lambda \) for all \( \lambda \), which in turn is equivalent to the minimal polynomial of \( \phi \) being separable (Example 1.27). The following proposition gives an abstract description of the familiar Jordan canonical form.

**Proposition 5.19.** Let \( V \) be a finite-dimensional vector space over an algebraically closed field \( \mathbb{k} \) and let \( \phi \in \text{End}_\mathbb{k}(V) \). Then \( \phi = \sigma + \nu \) with unique \( \sigma, \nu \in \text{End}_\mathbb{k}(V) \) such that \( \sigma \) is diagonalizable, \( \nu \) is nilpotent and \( \sigma \nu = \nu \sigma \). Moreover, \( \sigma = p(\phi) \) and \( \nu = q(\phi) \) for suitable polynomials \( p(t), q(t) \in \mathbb{k}[t] \).
5.3. Three Theorems about Linear Lie Algebras

Proof. Consider the decomposition (5.11) and write the characteristic polynomial of $\phi$ as $\prod A(t-\lambda)^{m_{\lambda}}$ with $m_{\lambda} = \dim V_{\lambda}$. Then $V_{\lambda} = \ker(\phi - \lambda)^{m_{\lambda}}$. By the Chinese Remainder Theorem, there is a polynomial $p(t) \in \mathbb{K}[t]$ satisfying the congruences $p(t) \equiv \lambda \mod (t-\lambda)^{m_{\lambda}}$ for all eigenvalues $\lambda$ of $\phi$. Put $q(t) = t - p(t)$ and let $\sigma = p(\phi)$ and $\nu = q(\phi)$. Then, clearly, $\sigma = \sigma + \nu$ and $\sigma \nu = \nu \sigma$. We show that $\sigma$ is diagonalizable and $\nu$ nilpotent. Since $\sigma$ and $\nu$ both commute with $\phi$, they both stabilize all generalized eigenspaces $V_{\lambda}$. Therefore, it suffices to check the desired properties for the restrictions $\sigma|_{V_{\lambda}}$ and $\nu|_{V_{\lambda}}$. But $(\phi - \lambda)^{m_{\lambda}}|_{V_{\lambda}} = 0$ implies $\sigma|_{V_{\lambda}} = p(\phi)|_{V_{\lambda}} = \lambda \text{Id}_{V_{\lambda}}$ by the defining congruences for $p(t)$, and $\nu^{m_{\lambda}}|_{V_{\lambda}} = (\phi - \sigma)^{m_{\lambda}}|_{V_{\lambda}} = (\phi - \lambda)^{m_{\lambda}}|_{V_{\lambda}} = 0$. This proves the proposition except for the uniqueness assertion.

For uniqueness, assume that $\phi = \sigma' + \nu'$ in $\mathfrak{gl}(V)$, with $\sigma'$ diagonalizable, $\nu'$ nilpotent and $\sigma' \nu' = \nu' \sigma'$. Then $\sigma'$ and $\nu'$ both commute with $\phi$, and hence also with $\sigma = p(\phi)$ and $\nu = q(\phi)$. Therefore, $\nu - \nu'$ is nilpotent, being the sum of two commuting nilpotent operators, and $\sigma - \sigma'$ is diagonalizable, being the sum of two commuting diagonalizable operators. Hence, the operator $\sigma - \sigma' = \nu' - \nu$ is diagonalizable as well as nilpotent, which forces it to be zero. \hfill \Box

The decomposition $\phi = \sigma + \nu$ in Proposition 5.19 is called the Jordan decomposition (or Jordan-Chevalley decomposition) of $\phi \in \text{End}_{\mathbb{K}}(V)$.

Lemma 5.20. Let $V$ be a finite-dimensional vector space over an algebraically closed field $\mathbb{K}$. If $x \in \mathfrak{gl}(V)$ has Jordan decomposition $x = s + n$, then $\text{ad}x = \text{ad}s + \text{ad}n$ is the Jordan decomposition of $\text{ad}x \in \mathfrak{gl}(\mathfrak{gl}(V))$.

Proof. Clearly, $\text{ad}x = \text{ad}s + \text{ad}n$ and $[\text{ad}s, \text{ad}n] = \text{ad}[s, n] = 0$. It remains to show that $\text{ad}n$ is a nilpotent operator on $\mathfrak{gl}(V)$ while $\text{ad}s$ is diagonalizable. The former was already done in Lemma 5.15. For $\text{ad}s$, fix a $\mathbb{K}$-basis $(v_i)$ of $V$ consisting of $s$-eigenvectors, say $sv_i = \lambda_i v_i$. Let $e_{i,j} \in \mathfrak{gl}(V)$ be the operator given by $e_{i,j}(v_k) = \delta_{j,k} v_i$. Then $(e_{i,j})$ is a $\mathbb{K}$-basis for $\mathfrak{gl}(V)$ and

\[
((\text{ad}s)(e_{i,j}))(v_k) = (s e_{i,j} - e_{i,j} s)(v_k) = s(\delta_{j,k} v_i) - e_{i,j}(\lambda_k v_k) = \delta_{j,k} \lambda_i v_i - \lambda_i \delta_{j,k} v_i = (\lambda_i - \lambda_j) e_{i,j}(v_k).
\]

Therefore, we have the following relation, which shows that $\text{ad}s$ is diagonalizable:

\[
(\text{ad}s)(e_{i,j}) = (\lambda_i - \lambda_j) e_{i,j}.
\] \hfill \Box

5.3.4. Cartan’s Criterion

We close this section with a solvability criterion\footnote{The criterion is due to the French mathematician Élie Cartan (1869 - 1951), the father of Henri Cartan, who was himself an eminent mathematician and a founding member of Bourbaki.} that will play an important role in our investigation of semisimple Lie algebras in Chapter 6.
Cartan’s Criterion. Let $V \in \text{Vect}_k$ be finite dimensional and assume that char $k = 0$. Then a Lie subalgebra of $\mathfrak{g} \subseteq \mathfrak{gl}(V)$ is solvable if and only if $\text{trace}(xy) = 0$ for all $x \in [\mathfrak{g}, \mathfrak{g}]$ and $y \in \mathfrak{g}$.

Proof. The base field $k$ may be assumed to be algebraically closed without loss of generality. For, if $\overline{k}$ is an algebraic closure of $k$, then it is easy to see that the Lie $\overline{k}$-algebra $\mathfrak{g} \otimes \overline{k}$ is solvable if and only if $\mathfrak{g}$ is so (Exercise 5.2.5). Furthermore, the condition $\text{trace}(xy) = 0$ for all $x \in [\mathfrak{g}, \mathfrak{g}]$ and $y \in \mathfrak{g}$ is evidently equivalent to the same condition for $\mathfrak{g} \otimes \overline{k}$. Therefore, we may use the above facts about Jordan decomposition as well as Lie’s Theorem below.

First, assume that $\mathfrak{g}$ is solvable. Then we know by Lie’s Theorem that there is a basis of $V$ such that the matrices of all operators from $\mathfrak{g}$ are upper triangular. The matrices of all $xy$ with $x \in [\mathfrak{g}, \mathfrak{g}]$ and $y \in \mathfrak{g}$ are thus strictly upper triangular, and hence $\text{trace}(xy) = 0$.

Conversely, assume that $\text{trace}([\mathfrak{g}, \mathfrak{g}]\mathfrak{g}) = \{0\}$. In order to show that $\mathfrak{g}$ is solvable, it will suffice to show that $[\mathfrak{g}, \mathfrak{g}]$ is solvable; for, $\mathfrak{g}$ will then be solvable-by-abelian and hence solvable (Proposition 5.12). Replacing $\mathfrak{g}$ by $[\mathfrak{g}, \mathfrak{g}]$, we may thus assume that $\text{trace}(\mathfrak{g}\mathfrak{g}) = \{0\}$ and our goal is to deduce solvability of $\mathfrak{g}$ from this. To this end, let $\mathfrak{n} = N_{\mathfrak{gl}(V)}(\mathfrak{g})$ denote the normalizer of $\mathfrak{g}$ in $\mathfrak{gl}(V)$; this is a Lie subalgebra of $\mathfrak{gl}(V)$ with $\mathfrak{n} \supseteq \mathfrak{g}$. Since the trace of any product of operators is invariant under cyclic permutations of the factors, we have the following identity for $x, y, z \in \mathfrak{gl}(V)$:

\begin{equation}
\text{trace}([x, y]z) = \text{trace}(x[y, z]).
\end{equation}

Thus, $\text{trace}([\mathfrak{g}, \mathfrak{g}]\mathfrak{n}) = \text{trace}(\mathfrak{g}[\mathfrak{g}, \mathfrak{n}]) \subseteq \text{trace}(\mathfrak{g}\mathfrak{g}) = \{0\}$ by our hypothesis; so

\begin{equation}
\text{trace}(xy) = 0 \text{ for all } x \in [\mathfrak{g}, \mathfrak{g}] \text{ and } y \in \mathfrak{n}.
\end{equation}

We will show that (5.14) implies that all $x \in [\mathfrak{g}, \mathfrak{g}]$ are nilpotent operators. By Engel’s Theorem, it will then follow that the Lie algebra $[\mathfrak{g}, \mathfrak{g}]$ is nilpotent and, consequently, $\mathfrak{g}$ is nilpotent-by-abelian and hence solvable, as desired.

So assume (5.14) and let $x = s + n$ be the Jordan decomposition of $x$ as in Proposition 5.19. Our goal is to show that $s = 0$. Fix a $k$-basis $(v_i)$ of $V$ such that the matrix of $x$ is upper triangular and the matrix of $s$ diagonal, say $s.v_i = \lambda_i v_i$ with $\lambda_i \in k$. We need to show that all $\lambda_i = 0$ or, equivalently, $L = \sum_i \mathbb{Q}\lambda_i = 0$. For this, we will show that $L^* = \text{Hom}_\mathbb{Q}(L, \mathbb{Q}) = 0$. So let $f \in L^*$ be given and let $y = y_f \in \mathfrak{gl}(V)$ be the diagonal operator that is defined by $y.v_i = f(\lambda_i)v_i$. Observe that $\text{trace}(xy) = \sum_i \lambda_i f(\lambda_i)$. It will suffice to prove the following

Claim. $y \in \mathfrak{n}$.

Indeed, (5.14) will then give that $\sum_i \lambda_i f(\lambda_i) = 0$. Applying $f$ to this expression, we further obtain $\sum_i f(\lambda_i)^2 = 0$. Since all $f(\lambda_i)$ are rational, this forces $f(\lambda_i) = 0$ for all $i$, and so $f = 0$ as needed.
To prove the claim, we will show that $ad\ y = r(ad\ x)$ for some polynomial $r(t) \in \mathbb{k}[t]$; this will imply that $[y, g] = r(ad\ x)(g) \subseteq g$ as needed. To construct $r$, recall from (5.12) that $(ad\ x)(e_{i,j}) = (\lambda_i - \lambda_j)e_{i,j}$ and $(ad\ y)(e_{i,j}) = (f(\lambda_i) - f(\lambda_j))e_{i,j}$ for the basis vectors $e_{i,j}$ of $gl(V)$. By Lagrange interpolation, we may choose a polynomial $q(t) \in \mathbb{k}[t]$ satisfying $q(\lambda_i - \lambda_j) = f(\lambda_i) - f(\lambda_j)$ for all $i, j$—note that $\lambda_i - \lambda_j = \lambda_k - \lambda_i$ implies $f(\lambda_i) - f(\lambda_j) = f(\lambda_k) - f(\lambda_i)$ by linearity of $f$. Thus, $ad\ y = q(ad\ x)$. By Proposition 5.19 and Lemma 5.20 we further know that $ad\ x = p(ad\ x)$ for some $p(t) \in \mathbb{k}[t]$. Therefore, the polynomial $r(t) = q(p(t))$ satisfies $r(ad\ x) = q(ad\ x) = ad\ y$, proving the claim and completing the proof of Cartan’s Criterion. □

**Exercises for Section 5.3**

5.3.1 (Some representations of the 2-dimensional non-abelian Lie algebra). Let $g = \mathbb{k}x \oplus \mathbb{k}y$, $[x, y] = y$, be the non-abelian Lie algebra of dimension 2 and let $V = \mathbb{k}[t]$.

(a) Define $\rho(x), \rho(y) \in gl(V)$ by $\rho(x) = t \frac{\partial}{\partial t}$ and $\rho(y) = t$ (multiplication by $t$). Show that this yields a representation $\rho : g \to gl(V)$.

(b) For char $\mathbb{k} = 0$, show that the ideals $(t^n)$ for $n \geq 0$ are the only nonzero subrepresentations of $V$. In particular, $V$ has no irreducible subrepresentation.

(c) For char $\mathbb{k} = p > 0$, show that $U = (t^p - 1)$ is a maximal proper subrepresentation of $V$. Thus, $V/U$ is a $p$-dimensional irreducible representation of $g$. This gives another counterexample to Lie’s Theorem in positive characteristics.

5.3.2 (Structure of solvable Lie algebras). In (a)–(c), let $g \in Lie_\mathbb{k}$ be finite dimensional and solvable.

(a) Show that there exists a chain of Lie subalgebras $0 = g_0 \subseteq g_1 \subseteq \cdots \subseteq g_n = g$ such that each $g_i$ is an ideal of $g_{i+1}$ and $\dim_\mathbb{k} g_i = i$. In particular, $g_i \cong g_{i-1} \rtimes \mathbb{k}x_i$ for all $i$.

(b) Assume that $\mathbb{k}$ is algebraically closed with char $\mathbb{k} = 0$. Use Lie’s Theorem to show that the chain in (a) may be chosen so that all $g_i$ are ideals of $g$.

(c) Assume that char $\mathbb{k} = 0$. Use (b) (and a field extension) to show that $\mathcal{D}g = [g, g]$ is nilpotent.

By Proposition 5.12(b), nilpotency of the derived subalgebra $\mathcal{D}g$ of an arbitrary $g \in Lie_\mathbb{k}$ certainly implies solvability of $g$. The remaining parts of this exercise show that (c) fails if $g$ is infinite dimensional or char $\mathbb{k} \neq 0$.

(d) Let $\mathbb{k}[t]$ be the standard representation of the Weyl algebra $A_1(\mathbb{k})$ (Example 1.8), viewed as a representation of the Heisenberg Lie algebra $h = \mathbb{k}x \oplus \mathbb{k}y \oplus \mathbb{k}z$, $[x, y] = z$, as in Exercise 5.1.2: $x, y$ and $z$ act as $\frac{\partial}{\partial t}, t \cdot$ and $Id_{\mathbb{k}[t]}$, respectively. With the abelian Lie algebra structure on $\mathbb{k}[t]$, form the semidirect product $g = \mathbb{k}[t] \rtimes h$. 

Show that \( g \) is solvable but \( \mathcal{D}g \) is not nilpotent. This shows that the conclusion of (c) fails for infinite-dimensional Lie algebras.

(e) Let \( \text{char} \ k = p > 0 \). With \( g \) as in (d), show that \( a = (t^p) \times \{0\} \) is an ideal of \( g \) such that \( \mathcal{A} = g/a \) is finite dimensional and solvable, but \( \mathcal{D}g \) is not nilpotent. Thus (c) also fails in positive characteristics.

5.3.3 (Solvable non-nilpotent Lie algebras). (a) Let \( k \) be algebraically closed with \( \text{char} k = 0 \). Show that every finite-dimensional solvable Lie algebra \( g \) that is not nilpotent has the the 2-dimensional non-abelian Lie algebra \( k x \oplus k y, [x, y] = y \), as a homomorphic image.

(b) Using the Lie algebras \( g \) from Exercise 5.3.2(d)(e), show that (a) fails for infinite-dimensional \( g \) or when \( \text{char} k \neq 0 \).

5.3.4 (Jordan decomposition). Let \( V \in \text{Vect}_{k} \) be finite dimensional and assume that \( k \) algebraically closed. For any \( x \in gl(V) \), let \( x_s \) and \( x_n \) denote the diagonalizable and nilpotent parts, respectively, of the Jordan decomposition of \( x \). If \( x, y \in gl(V) \) commute, show that \( (x + y)_s = x_s + y_s \) and \( (x + y)_n = x_n + y_n \). Give an example to show that this can fail if \( x \) and \( y \) do not commute.

5.4. Enveloping Algebras

In the representation theory of Lie algebras, enveloping algebras serve a purpose analogous to that of group algebras in group representation theory: representations of a Lie algebra are the “same” as representations of its enveloping algebra. This section is devoted to the construction of enveloping algebras and to the development of their basic formal and structural properties, including the above connection with representations of Lie algebras.

5.4.1. The Enveloping Algebra of a Lie Algebra

The Goal. In analogy with (3.2), which establishes an adjointness relation between the unit group functor \( \times: \text{Alg}_{k} \rightarrow \text{Groups} \) and the group algebra functor \( k \_ : \text{Groups} \rightarrow \text{Alg}_{k} \), we are now looking for a functor

\[
U: \text{Lie}_{k} \rightarrow \text{Alg}_{k}
\]

that is left adjoint to the functor \( \times_{\text{Lie}}: \text{Alg}_{k} \rightarrow \text{Lie}_{k} \) (Example 5.2). Explicitly, for any \( g \in \text{Lie}_{k} \), we want an algebra \( Ug \in \text{Alg}_{k} \) satisfying the following universal property: for any \( A \in \text{Alg}_{k} \), there is a bijection, functorial in both \( g \) and \( A \),

\[
\text{Hom}_{\text{Alg}_{k}}(Ug, A) \cong \text{Hom}_{\text{Lie}_{k}}(g, A_{\text{Lie}})
\]

To realize (5.15), we will construct the algebra \( Ug \) along with a “canonical” Lie algebra map

\[
i_{g} : g \rightarrow (Ug)_{\text{Lie}}
\]
such that the following condition is satisfied: for any Lie algebra homomorphism $f: \mathfrak{g} \to A_{\text{Lie}}$ with $A \in \text{Alg}_k$, there is a unique $k$-algebra map $\tilde{f}: U_\mathfrak{g} \to A$ making the following diagram commute:

$$
\begin{array}{ccc}
\mathfrak{g} & \xrightarrow{f} & A \\
\downarrow \iota_\mathfrak{g} & & \downarrow \exists \tilde{f} \\
U_\mathfrak{g} & \xrightarrow{\iota} & \end{array}
$$

(5.16)

The desired bijection $\text{Hom}_{\text{Lie}_k}(\mathfrak{g}, A_{\text{Lie}}) \cong \text{Hom}_{\text{Alg}_k}(U_\mathfrak{g}, A)$ will then be given by $f \mapsto \tilde{f}$; the inverse map sends $h \in \text{Hom}_{\text{Alg}_k}(U_\mathfrak{g}, A)$ to $h_{\text{Lie}} \circ \iota_\mathfrak{g} \in \text{Hom}_{\text{Lie}_k}(\mathfrak{g}, A_{\text{Lie}})$. In fact, it is not hard to see that (5.15) necessitates the existence of a map $\iota_\mathfrak{g}$ as in the foregoing (Exercise 5.4.1).

**The Construction.** Let $\mathfrak{g} \in \text{Lie}_k$. Starting with the tensor algebra $T_\mathfrak{g} = \bigoplus_{n \geq 0} \mathfrak{g}^\otimes n$ of $\mathfrak{g}$, regarded as a $k$-vector space, we define the *enveloping algebra* $U_\mathfrak{g}$ by

$$
U_\mathfrak{g} \overset{\text{def}}{=} T_\mathfrak{g}/L \quad \text{with} \quad L = L(\mathfrak{g}) = \langle x \otimes y - y \otimes x - [x, y] \mid x, y \in \mathfrak{g} \rangle
$$

Note the formal similarity to the earlier constructions of symmetric and exterior algebras (§1.1.2). In fact, for abelian $\mathfrak{g}$, we have $U_\mathfrak{g} = \text{Sym}_\mathfrak{g}$, because the ideal $L$ is identical to the relation ideal $I(\mathfrak{g})$ of the symmetric algebra in this case. In particular, the enveloping algebra of $\mathfrak{g} = 0$ is $U_\mathfrak{g} = k$. For a general Lie algebra $\mathfrak{g}$, however, a typical generator of $L$ is a combination of the degree-2 term $x \otimes y - y \otimes x \in \mathfrak{g}^\otimes 2$ and the degree-1 term $[x, y] \in \mathfrak{g} = \mathfrak{g}^\otimes 1$, which may both be nonzero. Thus, other than for symmetric and exterior algebras, the standard grading of $T_\mathfrak{g}$ generally does not pass down to $U_\mathfrak{g}$. However, we do at least know that the relation ideal $L$ is contained in the proper ideal $T^+_\mathfrak{g} = \bigoplus_{n > 0} \mathfrak{g}^\otimes n$ of the tensor algebra; so always $U_\mathfrak{g} \neq 0$.

The desired map $\iota_\mathfrak{g}: \mathfrak{g} \to U_\mathfrak{g}$ for the bijection (5.15) is the composite of the embedding $\mathfrak{g} = \mathfrak{g}^\otimes 1 \hookrightarrow T_\mathfrak{g}$ with the canonical epimorphism $T_\mathfrak{g} \twoheadrightarrow U_\mathfrak{g}$:

$$
\iota = \iota_\mathfrak{g}: \mathfrak{g} \to U_\mathfrak{g}, \quad x \mapsto x + L
$$

The relations generating the ideal $L$ are exactly what is needed to make $\iota$ a Lie algebra homomorphism from $\mathfrak{g}$ to the underlying Lie algebra $(U_\mathfrak{g})_{\text{Lie}}$ of $U_\mathfrak{g}$. Indeed, the fact that $x \otimes y - y \otimes x - [x, y] \in L$ translates into the following equation in $U_\mathfrak{g}$:

$$
[lx, ly] = lx ly - ly lx = l[x, y] \quad (x, y \in \mathfrak{g})
$$

(5.17)
The bijection (5.15) also follows:

\[ \text{Hom}_{\text{Alg}_k}(U\mathfrak{g}, A) \cong \left\{ f \in \text{Hom}_{\text{Alg}_k}(T\mathfrak{g}, A) \mid (fx)(fy) - (fy)(fx) = f[x, y] \text{ for all } x, y \in \mathfrak{g} \right\} \]

Proposition 1.1

\[ \text{Hom}_{\text{Lie}_k}(\mathfrak{g}, A_{\text{Lie}}) \cong \left\{ f \in \text{Hom}_k(\mathfrak{g}, A_{\text{Vect}_k}) \mid (fx)(fy) - (fy)(fx) = f[x, y] \text{ for all } x, y \in \mathfrak{g} \right\} \]

**Functoriality.** Any Lie homomorphism \( f : \mathfrak{g} \to \mathfrak{h} \) gives rise to the map \( \iota_{\mathfrak{h}} \circ f : \mathfrak{g} \to \mathfrak{h} \) in \( \text{Lie}_k \), which in turn by (5.16) corresponds to a unique algebra homomorphism \( Uf := \iota_{\mathfrak{h}} \circ f : U\mathfrak{g} \to U\mathfrak{h} \) making the following diagram commute:

\[
\begin{array}{ccc}
\mathfrak{g} & \xrightarrow{f} & \mathfrak{h} \\
\downarrow{\iota_{\mathfrak{g}}} & & \downarrow{\iota_{\mathfrak{h}}} \\
U\mathfrak{g} & \xrightarrow{Uf} & U\mathfrak{h}
\end{array}
\]

It is a trivial matter to check that \( U \cdot \) respects identity maps and composites as is required for a functor, and it is equally straightforward to verify that the bijection (5.15) is functorial in both inputs, \( \mathfrak{g} \) and \( A \). So we have a functor \( U : \text{Lie}_k \to \text{Alg}_k \) that is left adjoint to \( \cdot_{\text{Lie}} : \text{Alg}_k \to \text{Lie}_k \).

### 5.4.2. Representations of \( \mathfrak{g} \) and \( U\mathfrak{g} \)

Let \( \mathfrak{g} \in \text{Lie}_k \). Recall (§5.1.2) that a representation of \( \mathfrak{g} \) is given by some \( V \in \text{Vect}_k \) and a Lie homomorphism \( \mathfrak{g} \to \mathfrak{gl}(V), x \mapsto x_V \). Let us also define a morphism of representations of \( \mathfrak{g} \) to be a map \( f : V \to W \) in \( \text{Vect}_k \) satisfying \( f \circ x_V = x_W \circ f \) for all \( x \in \mathfrak{g} \). The representations of \( \mathfrak{g} \) are now a category, \( \text{Rep}_\mathfrak{g} \).

With the aid of the enveloping algebra \( U\mathfrak{g} \), all this can be placed into the context of Chapter 1. Indeed, since \( \mathfrak{gl}(V) = \text{End}_k(V)_{\text{Lie}} \), there is a natural bijection,

\[ \text{Hom}_{\text{Lie}_k}(\mathfrak{g}, \mathfrak{gl}(V)) \cong \text{Hom}_{\text{Alg}_k}(U\mathfrak{g}, \text{End}_k(V)) \]

The set on the right consists of the representations \( U\mathfrak{g} \to \text{End}_k(V) \) of the algebra \( U\mathfrak{g} \), which in turn correspond to left \( U\mathfrak{g} \)-module action \( U\mathfrak{g} \otimes V \to V \) (§1.1.3). Viewing representations of \( \mathfrak{g} \) as representations of \( U\mathfrak{g} \) or left \( U\mathfrak{g} \)-modules in this manner, morphisms \( f : V \to W \) in \( \text{Rep}_\mathfrak{g} \) are the same as maps in \( \text{Rep}_U\mathfrak{g} \) or \( U\mathfrak{g} \)-module maps, because the algebra \( U\mathfrak{g} \) is generated by the image of the canonical map \( \iota : \mathfrak{g} \to U\mathfrak{g} \). To summarize, we have equivalences of categories,

\[
\text{Rep}_\mathfrak{g} \equiv \text{Rep}_U\mathfrak{g} \equiv U\mathfrak{g}\text{Mod}
\]
Frequently, our focus will be on the full subcategory $\text{Rep}_\text{fin} \mathfrak{g} = \text{Rep}_\text{fin} \mathcal{U}_\mathfrak{g}$ consisting of the finite-dimensional representations of $\mathfrak{g}$. We will also write $\text{Irr} \mathfrak{g} = \text{Irr} \mathcal{U}_\mathfrak{g}$ for the isomorphism classes of irreducible representations of $\mathfrak{g}$, or a full representative set thereof; similarly for $\text{Irr}_\text{fin} \mathfrak{g} = \text{Irr}_\text{fin} \mathcal{U}_\mathfrak{g}$.

The $\mathfrak{g}$-kernel of a representation $V \in \text{Rep} \mathfrak{g}$ is the ideal of $\mathfrak{g}$ that is defined by

$$\text{Ker}_\mathfrak{g} V \overset{\text{def}}{=} \{ x \in \mathfrak{g} \mid x_V = 0 \}.$$ If $\text{Ker}_\mathfrak{g} V = 0$, then $V$ will be called $\mathfrak{g}$-faithful. Viewing $V$ as a representation of $\mathcal{U}_\mathfrak{g}$, we have $\text{Ker}_\mathfrak{g} V = \mathfrak{g} \cap \text{Ker} V$. Faithfulness for $\mathcal{U}_\mathfrak{g}$ is thus stronger than $\mathfrak{g}$-faithfulness. The relationship between these two notions will be discussed in Proposition 5.28.

5.4.3. The Poincaré-Birkhoff-Witt Theorem

While the construction of enveloping algebras and the verification of their basic functorial properties have been rather effortless, it requires honest work to understand what enveloping algebras actually look like. At this point, it is not clear, for example, whether the canonical map $\iota: \mathfrak{g} \to \mathcal{U}_\mathfrak{g}$ is injective or how to obtain a $k$-basis for $\mathcal{U}_\mathfrak{g}$. These issues are easily settled for tensor algebras, symmetric algebras and group algebras, but the case of enveloping algebras requires a substantial theorem, the celebrated Poincaré-Birkhoff-Witt Theorem.

**Ring Theoretic Intermezzo: Filtrations**

We have already noted that the standard grading of the tensor algebra $T_\mathfrak{g}$ fails to give a grading of $\mathcal{U}_\mathfrak{g}$, but we shall see that it does at least provide us with a very useful filtration of $\mathcal{U}_\mathfrak{g}$.

A filtration\(^6\) of a $k$-vector space $V$ is an increasing chain of subspaces $(V_n)_{n \in \mathbb{Z}}$ such that $V_m = 0$ for some $m$ and $V = \bigcup_n V_n$. Thus,

$$0 = V_m \subseteq V_{m+1} \subseteq \cdots \subseteq V_n \subseteq V_{n+1} \subseteq \cdots \subseteq V = \bigcup_n V_n.$$ A filtration $(A_n)$ of a $k$-algebra $A$ is also required to respect the multiplication and the unit of $A$ in the sense that

$$1 \in A_0 \quad \text{and} \quad A_n A_{n'} \subseteq A_{n+n'} \quad (n, n' \in \mathbb{Z}).$$

Given a filtration $(A_n)$ of $A \in \text{Alg}_k$, the **associated graded algebra** is defined by

$$\text{gr} A \overset{\text{def}}{=} \bigoplus_n \text{gr}^n A \quad \text{with} \quad \text{gr}^n A \overset{\text{def}}{=} A_n / A_{n-1}$$

---

\(^{*}\)More precisely, a filtration as considered above is called increasing, discrete and exhaustive, but we shall be exclusively concerned with filtrations having these properties and hence not refer to them explicitly.
The \( k \)-linear structure and grading of \( \text{gr} A \) are clear. The unit map is the composite \( \mathbb{k} \hookrightarrow A_0 \to A_0/A_{-1} \) and the multiplication of \( \text{gr} A \) is given by

\[
(x + A_{n-1})(x' + A_{n'-1}) := xx' + A_{n+n'-1} \in A_{n+n'}/A_{n+n'-1}
\]

for \( x \in A_n \) and \( x' \in A_{n'} \). With this, \( \text{gr} A \) does indeed become a graded \( \mathbb{k} \)-algebra. For \( 0 \neq x \in A \), we may choose \( n \) minimal with \( x \in A_n \) and define the symbol of \( x \) with respect to the given filtration by

\[
x := x + A_{n+1} \in \text{gr}^n A \setminus \{0\}.
\]

**Example 5.21** (Filtrations given by generators). If \( X = (x_i)_{i \in I} \) is a fixed family of generators of \( A \in \text{Alg}_k \), then we obtain a filtration of \( A \) by defining \( A_n \) \((n \geq 0)\) to be the \( k \)-linear span of all monomials of length at most \( n \) with factors from \( X \):

\[
A_n := \langle x_i x_{i_1} \ldots x_{i_m} \mid i_j \in I, m \leq n \rangle_k.
\]

Thus, \( A_0 = \mathbb{k} \) and we also put \( A_n := 0 \) for \( n < 0 \). We will refer to this filtration as the \emph{\( X \)-filtration} of \( A \) and denote the associated graded algebra by \( \text{gr}_X A \).

**Example 5.22** (Filtrations from gradings). If \( A = \bigoplus_{n \geq 0} A^n \) is a graded algebra (§1.1.2), then we may filter \( A \) by putting \( A_n := \bigoplus_{k \leq n} A^k \). Manifestly, \( A \cong \text{gr} A \) as graded algebras and the symbol of any \( 0 \neq x \in A \) is just the homogeneous component of \( x \) having largest degree.

For general filtered algebras \( A \), the structure of \( \text{gr} A \) will be quite different from that of \( A \). Nonetheless, some properties of \( \text{gr} A \) transfer nicely to \( A \). To wit:

**Lemma 5.23.** Let \( A \) be a filtered algebra. If \( \text{gr} A \) is a domain, then \( A \) is a domain as well; similarly for the properties of being right or left noetherian.

**Proof.** First assume that \( \text{gr} A \) is a domain, that is, products of any two nonzero elements are nonzero. Then, for any \( 0 \neq x, y \in A \), the symbols \( x, y \in \text{gr} A \) satisfy \( x \cdot y \neq 0 \), which in turn readily implies \( xy \neq 0 \) (Exercise 5.4.2).

For noetherianess (left, say), note that any \( \mathbb{k} \)-subspace \( V \subseteq A \) inherits a filtration from \( A \) by putting \( V_n = V \cap A_n \). Then

\[
\text{gr} V = \bigoplus_n V_n/V_{n-1} \cong \bigoplus_n (V_n + A_{n-1})/A_{n-1}
\]

is a subspace of \( \text{gr} A \). Furthermore, \( V \subseteq V' \) clearly implies \( \text{gr} V \subseteq \text{gr} V' \); and if \( V \subseteq V' \), then, choosing \( n \) minimal with \( V_n \subseteq V'_n \), we readily obtain \( V_n/V_{n-1} \subseteq V'_n/V'_{n-1} \) and so \( \text{gr} V \subseteq \text{gr} V' \). Finally, if \( V \) is a left ideal of \( A \), then \( \text{gr} V \) is easily seen to be a left ideal of \( \text{gr} A \). Therefore, any ascending chain of left ideals of \( A \) gives rise to an ascending chain of left ideals of \( \text{gr} A \), and when the latter chain stabilizes, then so does the former. \( \square \)
The Standard Filtration of $U\mathfrak{g}$

Returning to the case of a Lie algebra $\mathfrak{g}$, consider the filtration of the tensor algebra $T\mathfrak{g}$ that comes from the standard grading as in Example 5.22:

$$T_n\mathfrak{g} = \bigoplus_{k \leq n} \mathfrak{g}^\otimes k.$$ 

Taking images under the canonical epimorphism $T\mathfrak{g} \twoheadrightarrow U\mathfrak{g} = T\mathfrak{g}/L$, we obtain the standard filtration of the enveloping algebra $U\mathfrak{g}$:

$$U_n = U_n\mathfrak{g} \overset{\text{def}}{=} (T_n\mathfrak{g} + L)/L.$$ 

So $U_{−1} = 0 \subseteq U_0 = \mathbb{k} \subseteq U_1 = \mathbb{k} + \mathfrak{g} \subseteq \cdots \subseteq U_0 = \bigcup_n U_n$ and $U_n U_{n'} \subseteq U_{n+n'}$. Any $\mathbb{k}$-basis $(x_i)_{i \in I}$ of $\mathfrak{g}$ generates the tensor algebra $T\mathfrak{g}$, and hence the algebra $U\mathfrak{g}$ is generated by $X = (x_i)_{i \in I}$. The standard filtration $(U_n)$ is the same as the $X$-filtration of $U\mathfrak{g}$ in the sense of Example 5.21. We will be particularly interested in the associated graded algebra,

$$\text{gr } U\mathfrak{g} = \bigoplus_n U_n/U_{n−1},$$

which provides a useful link between the (commutative) symmetric algebra $\text{Sym } \mathfrak{g}$ and the (generally noncommutative) enveloping algebra $U\mathfrak{g}$.

To pinpoint the connection, we remark that it is easy to see (Exercise 5.4.3) that $\text{gr } U\mathfrak{g}$ is generated by the elements $\hat{x} := \iota x + U_0 \in \text{gr}^1 U\mathfrak{g}$ for $x \in \mathfrak{g}$. (We shall soon identify $x$ and $\iota x$ and see that $\hat{x}$ is the symbol of $x$ if $x \neq 0$, but this is not essential for now.) These elements do in fact commute. Indeed, the relations (5.17) give the following equations in $\text{gr}^2 U\mathfrak{g}$:

$$\hat{x} \hat{y} − \hat{y} \hat{x} = \iota[x, y] + U_1 = 0 \quad (x, y \in \mathfrak{g}).$$

By the universal property (1.8) of the symmetric algebra, the linear map $\mathfrak{g} \to \text{gr}^1 U\mathfrak{g}$, $x \mapsto \hat{x}$, lifts uniquely to a homomorphism of graded algebras, which is surjective, because $\text{gr } U\mathfrak{g}$ is generated by its 1-component:

$$\phi: \text{Sym } \mathfrak{g} \longrightarrow \text{gr } U\mathfrak{g}$$

(5.18)

$$\quad \omega \quad \omega$$

$$\begin{array}{ccc}
  x & \longmapsto & \hat{x} \\
  (x \in \mathfrak{g})
\end{array}$$

The Poincaré-Birkhoff-Witt Theorem will also give injectivity of $\phi$ (Corollary 5.25), but this is not needed in order to establish the following important properties of enveloping algebras of finite-dimensional Lie algebras.
**Proposition 5.24.** The enveloping algebra \( U_\mathfrak{g} \) of a finite-dimensional Lie algebra \( \mathfrak{g} \) is affine and (right and left) noetherian.\(^7\)

**Proof.** The algebra \( U_\mathfrak{g} \) is clearly affine, being generated by the elements \( tx_i \) for any \( \mathbb{k} \)-basis \( (x_i)_{\mathfrak{g}} \) of \( \mathfrak{g} \). Furthermore, \( \text{Sym} \mathfrak{g} \cong \mathbb{k}[x_1, \ldots, x_d] \) is noetherian by the Hilbert Basis Theorem. It follows from (5.18) that \( \text{gr} U_\mathfrak{g} \) is noetherian as well, and hence \( U_\mathfrak{g} \) is noetherian by Lemma 5.23. \( \Box \)

**Standard \( \mathbb{k} \)-Bases of \( U_\mathfrak{g} \)**

Fix a \( \mathbb{k} \)-basis \( X = (x_i)_{i \in I} \) of \( \mathfrak{g} \). We may identify \( T_\mathfrak{g} \) with the free algebra \( \mathbb{k}\langle X \rangle \) (§1.1.2). The Lie algebra \( \mathfrak{g} \subseteq T_\mathfrak{g} \) then becomes the \( \mathbb{k} \)-span \( \langle X \rangle_{\mathbb{k}} \subseteq \mathbb{k}\langle X \rangle \) and the Lie bracket of \( \mathfrak{g} \) can be viewed inside \( \mathbb{k}\langle X \rangle \) as a bracket \( [\ldots] : \langle X \rangle_{\mathbb{k}} \times \langle X \rangle_{\mathbb{k}} \to \langle X \rangle_{\mathbb{k}} \). Since the generators \( x \otimes y - y \otimes x - [x, y] (x, y \in \mathfrak{g}) \) of the relations ideal \( L \subseteq T_\mathfrak{g} \) are linear in both \( x \) and \( y \), the enveloping algebra \( U_\mathfrak{g} \) has the following presentation:

\[
U_\mathfrak{g} \cong \mathbb{k}\langle X \rangle / (x_i x_j - x_j x_i - [x_i, x_j] \mid i, j \in I)
\]

Finite products \( x_{i_1} x_{i_2} \ldots x_{i_n} \in \mathbb{k}\langle X \rangle \) will simply be called monomials below; similarly for their images \( tx_{i_1} tx_{i_2} \ldots tx_{i_n} = x_{i_1} x_{i_2} \ldots x_{i_n} + L \in U_\mathfrak{g} \). Since \( U_\mathfrak{g} \) is generated by the elements \( tx_i \), the latter monomials form a \( \mathbb{k} \)-linear spanning set of \( U_\mathfrak{g} \). Moreover, by the above presentation, \( tx_i tx_j \equiv tx_j tx_i \mod \sum_{k \in I} \mathbb{k}tx_k \) for all \( i, j \in I \). Thus, an arbitrary monomial \( tx_{i_1} tx_{i_2} \ldots tx_{i_n} \in U_\mathfrak{g} \) can be rewritten in any desired order of the factors at the cost of introducing an additional \( \mathbb{k} \)-linear combination of monomials with fewer than \( n \) factors. Fixing a preferred order of the basis \( X \) by choosing a total order \( \preceq \) on \( I \), it follows by induction that every monomial with \( n \) factors can be expressed as a \( \mathbb{k} \)-linear combination of monomials \( tx_{i_1} tx_{i_2} \ldots tx_{i_m} \) with \( m \leq n \) and \( i_1 \leq i_2 \leq \cdots \leq i_m \). Thus, the family of all such monomials spans the \( \mathbb{k} \)-vector space \( U_\mathfrak{g} \). Linear independence is far less obvious; it is the essence of the following theorem.

**Poincaré-Birkhoff-Witt Theorem.** Let \( X = (x_i)_{i \in I} \) be a \( \mathbb{k} \)-basis of \( \mathfrak{g} \in \text{Lie}_\mathbb{k} \) and let \( \preceq \) be a total order on \( I \). Then a \( \mathbb{k} \)-basis of \( U_\mathfrak{g} \) is given by the images in \( U_\mathfrak{g} \) of the monomials \( x_{i_1} x_{i_2} \ldots x_{i_n} \in \mathbb{k}\langle X \rangle \) with \( n \geq 0 \) and \( i_1 \leq i_2 \leq \cdots \leq i_n \).

It is not necessary to see the proof of the theorem in order to appreciate the statement. We delegate this task to the appendix (Section D.5) and proceed to give some applications. First, applying the theorem to monomials with one factor, we see in particular that the canonical map \( \iota = \iota_{\mathfrak{g}} : \mathfrak{g} \to U_\mathfrak{g} \) is injective. Henceforth, we shall altogether dispense with the notation \( \iota \) and simply view the canonical map

\(^7\)It is an open problem whether there exist infinite-dimensional Lie algebras \( \mathfrak{g} \) such that \( U_\mathfrak{g} \) is noetherian; the answer is generally believed to be negative. The enveloping algebra of the Witt algebra (Exercise 5.1.6) is affine, but it is known to be non-noetherian \([185]\).
as an embedding,
\[ g \hookrightarrow U g. \]

We will refer to ordered monomials \( x_{i_1}x_{i_2} \ldots x_{i_n} \) with \( i_1 \leq i_2 \leq \cdots \leq i_n \) as in the Poincaré-Birkhoff-Witt Theorem as standard monomials (for the given ordered basis \( X \)). The number of factors will also be called the length of the monomial. Note that we can view \( x_{i_1}x_{i_2} \ldots x_{i_n} \in \k(X) \) and also \( x_{i_1}x_{i_2} \ldots x_{i_n} \in U g \), the latter monomial being the image of the former under the epimorphism \( \k(X) = T g \rightarrow U g \).

The monomial \( x_{i_1}x_{i_2} \ldots x_{i_n} \) also has meaning in the symmetric algebra \( \Sym g \). In fact, we know by (1.10) that the standard monomials of length \( n \) form a basis of the \( n \)-th homogeneous component \( \Sym^n g \).

**Corollary 5.25.**

(a) For any \( g \in \Lie k \), the map (5.18) is an isomorphism \( \phi : \Sym g \xrightarrow{\cong} \gr U g \). Consequently, \( U g \) is a domain.

(b) For any monomorphism \( f : h \hookrightarrow g \) in \( \Lie k \), the map \( U f : U h \rightarrow U g \) in \( \Alg k \) is injective and it makes \( U h \) a free left and right \( U h \)-module.

**Proof.** (a) Fix an ordered \( k \)-basis \( X = (x_i)_{i \in I} \) for \( g \) and let \((U_n)\) denote the standard filtration of \( U g \). By the Poincaré-Birkhoff-Witt Theorem, the standard monomials of length \( m \leq n \) form a \( k \)-basis of \( U_n \); indeed, they are linearly independent and, as we have observed above, any monomial \( x_{i_1}x_{i_2} \ldots x_{i_m} \) with \( m \leq n \) can be written as a linear combination of standard monomials of length \( \leq m \). It follows that the residue classes modulo \( U_{n-1} \) of the standard monomials \( x_{i_1}x_{i_2} \ldots x_{i_n} \) of length exactly \( n \) form a basis of \( \gr^n U g = U_n/U_{n-1} \). Letting \( \hat{x} = x + \k \in \gr U g \) denote the symbol of \( x \in X \) as before, the following equalities hold in \( \gr U g \):

\[
x_{i_1}x_{i_2} \ldots x_{i_n} + U_{n-1} = \hat{x}_{i_1}\hat{x}_{i_2} \ldots \hat{x}_{i_n} = \phi(x_{i_1}x_{i_2} \ldots x_{i_n}),
\]

where the last monomial is of course computed in \( \Sym^n g \). Thus \( \phi \) sends a basis of each \( \Sym^n g \) to a basis of \( \gr^n U g \), and hence \( \phi \) is bijective. The last assertion in (a) now follows from Lemma 5.23, since \( \Sym g \) is a polynomial algebra and so a domain.

(b) Replacing \( h \) by its image in \( g \), we may assume that \( h \) is a Lie subalgebra of \( g \). Choose a \( k \)-basis of \( g \) of the form \( X = Y \sqcup Z \), where \( Y \) is a \( k \)-basis of \( h \), and order \( X \) by ordering \( Y \) and \( Z \) and by declaring \( y < z \) for each \( y \in Y \) and \( z \in Z \). Then the ordered monomials with factors from \( Y \) are part of the \( k \)-basis of \( U h \) that is provided by the Poincaré-Birkhoff-Witt Theorem. Therefore, the algebra map \( U h \rightarrow U h \) that arises by functoriality from the inclusion \( h \hookrightarrow g \) sends a \( k \)-basis of \( U h \) to linearly independent elements of \( U h \), proving injectivity. The Poincaré-Birkhoff-Witt Theorem also implies that the standard monomials with factors from \( Z \) give a basis of \( U h \) as left \( U h \)-module. For the right \( U h \)-module structure, argue similarly declaring instead that \( y > z \) for \( y \in Y \) and \( z \in Z \). \( \square \)
5.4.4. More Structure

In analogy with the material for group algebras in Section 3.3, we will now construct additional structure maps for enveloping algebras. These maps will equip $U\mathfrak{g}$ with the structure of a Hopf algebra (Part IV) and they will allow us to perform interesting constructions with representations of Lie algebras. The material below does not depend on the Poincaré-Birkhoff-Witt Theorem, even though we will occasionally refer to this result and we will permit ourselves to suppress the canonical map $\iota_{\mathfrak{g}} : \mathfrak{g} \to U\mathfrak{g}$ and identify Lie algebra elements with their images in the enveloping algebra for notional ease.

Direct Products and Opposites. We begin by describing the enveloping algebras of direct products of Lie algebras and of opposite Lie algebras. The results will be the exact analogs of corresponding results for group algebras (Exercise 3.1.2) and they will be easy consequences of the universal property (5.15) of enveloping algebras.

Recall from §5.1.6 that, for given $\mathfrak{g}_1, \mathfrak{g}_2 \in \text{Lie}_k$, the direct product $\mathfrak{g}_1 \times \mathfrak{g}_2$ is a Lie algebra with the componentwise Lie bracket. Furthermore, analogous to the earlier notions of opposite algebras and groups, we may define a new Lie bracket for a given $\mathfrak{g} \in \text{Lie}_k$ by declaring $[x, y]^\text{op} := [y, x] = -[x, y]$ for $x, y \in \mathfrak{g}$. This yields the opposite Lie algebra $\mathfrak{g}^\text{op} \in \text{Lie}_k$. As with groups and algebras, it is often convenient to work with a $k$-linear isomorphism $\mathfrak{g} \cong \mathfrak{g}^\text{op}$, $x \leftrightarrow x^\text{op}$, and write the bracket of $\mathfrak{g}^\text{op}$ as $[x^\text{op}, y^\text{op}] = [y, x]^\text{op}$.

**Proposition 5.26.** Let $\mathfrak{g}, \mathfrak{g}_1, \mathfrak{g}_2 \in \text{Lie}_k$. Then:

(a) $U(\mathfrak{g}_1 \times \mathfrak{g}_2) \cong U\mathfrak{g}_1 \otimes U\mathfrak{g}_2$.
(b) $U(\mathfrak{g}^\text{op}) \cong (U\mathfrak{g})^\text{op}$.

**Proof.** (a) For brevity, let us put $\mathfrak{g} = \mathfrak{g}_1 \times \mathfrak{g}_2$, $U = U\mathfrak{g}$, and $U_i = U\mathfrak{g}_i$ ($i = 1, 2$). The map

$$
\begin{array}{ccc}
\mathfrak{g}_1 & \longrightarrow & \mathfrak{g} \\
\downarrow & & \downarrow \text{can.} \\
\mathfrak{g} & \longrightarrow & U \\
\downarrow & & \downarrow \\
x & \longmapsto & (x, 0)
\end{array}
$$

extends uniquely to an algebra map $f_1 : U_1 \to U$ by (5.15). Similarly, we have a map $f_2 : U_2 \to U$. The images of these two maps commute elementwise in $U$, because $[(x_1, 0), (0, x_2)] = 0$ for $x_j \in \mathfrak{g}_j$. So we obtain an algebra map $f : U_1 \otimes U_2 \to U$ with $f(u_1 \otimes u_2) = (f_1 u_1)(f_2 u_2)$ (Exercise 1.1.10). The Poincaré-Birkhoff-Witt Theorem implies that this map is an isomorphism, but one can also obtain an explicit inverse.
directly from (5.15) as follows. Consider the map

\[
\begin{array}{c}
g = g_1 \times g_2 \\
\downarrow \quad \downarrow \\
(x_1, x_2) \mapsto x_1 \otimes 1 + 1 \otimes x_2
\end{array}
\]

(5.19)

It is readily checked that (5.19) is a Lie homomorphism; so (5.15) yields a unique lift \( g : U \to U_1 \otimes U_2 \) in \( \text{Alg}_{k} \). To see that this is the desired inverse for \( f \), we compute, for \( x_i \in g_i \),

\[
(f \circ g)(x_1, x_2) = f(x_1 \otimes 1 + 1 \otimes x_2) = (f_1 x_1) 1 + 1 (f_2 x_2) = (x_1, x_2).
\]

Since the elements \( (x_1, x_2) \) generate the algebra \( U \), this shows that \( f \circ g = \text{Id}_U \).

Similarly one checks that \( g \circ f = \text{Id}_{U_1 \otimes U_2} \).

(b) We continue to write \( U = U_\mathfrak{g} \). Recall that \( U^{\text{op}} \cong U \), \( a^{\text{op}} \leftrightarrow a \), is a \( k \)-linear isomorphism with \( a^{\text{op}} b^{\text{op}} = (ba)^{\text{op}} \) for \( a, b \in U \). The map

\[
\begin{array}{c}
g^{\text{op}} \\
\downarrow \quad \downarrow \\
x^{\text{op}} \mapsto x^{\text{op}}
\end{array}
\]

is a Lie map: \([x^{\text{op}}, y^{\text{op}}] = [y, x]^{\text{op}} \mapsto (yx - xy)^{\text{op}} = x^{\text{op}} y^{\text{op}} - y^{\text{op}} x^{\text{op}} \in U^{\text{op}} \).

Therefore, there is a lift \( f : U(\mathfrak{g})^{\text{op}} \rightarrow U^{\text{op}} \) in \( \text{Alg}_{k} \) by (5.15). To see that \( f \) is an isomorphism, we may either invoke the Poincaré-Birkhoff-Witt Theorem or else produce the inverse directly by considering the algebra \( A = U(\mathfrak{g})^{\text{op}} \) and the map \( \mathfrak{g} \to A_{\text{Lie}}, x \mapsto (x^{\text{op}})^{\text{op}} \). Again, one checks easily that this is a map in \( \text{Lie}_k \), and so there is a lift \( g : U \to A \) in \( \text{Alg}_{k} \). The map \( g^{\text{op}} : U^{\text{op}} \to A^{\text{op}} = U(\mathfrak{g})^{\text{op}} \) is readily checked to be inverse to \( f \).

\[ \square \]

**Counit, Comultiplication, and Antipode.** The map \( \mathfrak{g} \to 0 \) is clearly a Lie homomorphism. Applying the functor \( U : \text{Lie}_k \to \text{Alg}_k \), we obtain a map \( U\mathfrak{g} \to U0 = k \) in \( \text{Alg}_k \) with the following description:

\[
\begin{array}{c}
\varepsilon : U\mathfrak{g} \\
\downarrow \quad \downarrow \\
x \mapsto 0 \quad (x \in \mathfrak{g})
\end{array}
\]

(5.20)

The images \( \varepsilon(a) \) for arbitrary \( a \in U\mathfrak{g} \) are then determined, because the algebra \( U\mathfrak{g} \) is generated by \( \mathfrak{g} \subseteq U\mathfrak{g} \). In analogy with group algebras, \( \varepsilon \) is called the **augmentation map or counit** of \( U\mathfrak{g} \) and the ideal

\[
(U\mathfrak{g})^+ \overset{\text{def}}{=} \text{Ker} \varepsilon
\]

is called the **augmentation ideal** of \( U\mathfrak{g} \). Evidently, \((U\mathfrak{g})^+ = \mathfrak{g}(U\mathfrak{g}) = (U\mathfrak{g})\mathfrak{g} \).
The “diagonal map” \( g \to g \times g, x \mapsto (x, x) \), is also a map in \( \text{Lie}_k \). In view of the isomorphism \( U(g \times g) \cong U_g \otimes U_g \) (Proposition 5.26; see especially formula (5.19)), the functor \( U : \text{Lie}_k \to \text{Alg}_k \) yields the algebra map

\[
\Delta: \quad U_g \longrightarrow U_g \otimes U_g
\]

Again, the images \( \Delta(a) \) for arbitrary \( a \in U_g \) are then determined, but \( \Delta(a) \) generally has a more complicated expression than \( \Delta(x) \) above. The map \( \Delta \) is called the \textbf{comultiplication} of \( U_g \). As in the case of group algebras, the comultiplication \( \Delta \) is \textbf{cocommutative}: for the switch automorphism \( \tau \in \text{Aut}_{\text{Alg}_k}(U_g \otimes U_g) \), we have

\[
\Delta = \tau \circ \Delta.
\]

Indeed, since \( g \) generates the algebra \( U_g \), it suffices to check (5.22) under evaluation on \( x \in g \). But then (5.22) becomes \( x \otimes 1 + 1 \otimes x = \tau(x \otimes 1 + 1 \otimes x) \), which is obviously true. In the same way, one easily verifies commutativity of the diagrams below, the first of which states the coassociativity property of \( \Delta \); the diagrams are the analogs of the group algebra diagrams (3.26) and justify the names “comultiplication” and “counit.”

Finally, the map \( g \cong g^\text{op}, x \mapsto -x^\text{op} \) is clearly an isomorphism in \( \text{Vect}_k \) and the calculation \([-x^\text{op}, -y^\text{op}] = [x^\text{op}, y^\text{op}] = [y, x]^\text{op} = -[x, y]^\text{op} \) shows that it is also a map in \( \text{Lie}_k \). As above, we derive from this and the isomorphism \( U(g^\text{op}) \cong (U_g)^\text{op} \) (Proposition 5.26) an isomorphism of \( k \)-algebras,

\[
S: \quad U_g \longrightarrow (U_g)^\text{op}
\]

We will generally think of \( S \) as a map from \( U_g \) to \( U_g \) satisfying \( Sx = -x \) for \( x \in g \). Then \( S \) becomes an involution, that is \( S(ab) = S(b)S(a) \) for all \( a, b \in U_g \) and \( S^2 = \text{Id} \). The map \( S \) is called the \textbf{standard involution} or the \textbf{antipode} of \( U_g \).
5.4. Enveloping Algebras

5.4.1 (Unit of adjunction). Assume that there is a bijection
\[ \sim: \text{Hom}_{\text{Lie}}(\mathfrak{g}, A_{\text{Lie}}) \sim \text{Hom}_{\text{Alg}}(U_{\mathfrak{g}}, A) \]
that is functorial in both \( \mathfrak{g} \in \text{Lie}_k \) and \( A \in \text{Alg}_k \) as in (5.15) and let \( \sim' \) denote the inverse bijection. Define \( \iota_\mathfrak{g} := \text{Id}_{U_{\mathfrak{g}}} : \mathfrak{g} \to (U_{\mathfrak{g}})_{\text{Lie}} \). Show that \( h' = h_{\text{Lie}} \circ \iota_\mathfrak{g} \) for \( h \in \text{Hom}_{\text{Alg}}(U_{\mathfrak{g}}, A) \). Conclude that \( f = f \circ \iota_\mathfrak{g} \) for \( f \in \text{Hom}_{\text{Lie}}(\mathfrak{g}, A_{\text{Lie}}) \) as in (5.16). The map \( \iota_\mathfrak{g} \) is called the unit of the adjunction (5.15); e.g., MacLane [140, IV.1].

5.4.2 (Some properties of filtered algebras). Let \( A \) be an algebra with a filtration \( (A_n) \) as in §5.4.3. For \( 0 \neq x \in A \), define \( \deg x = \min\{n \in \mathbb{Z} \mid x \in A_n\} \) and \( \hat{x} = x + A_{\deg x - 1} \in \text{gr} A \setminus \{0\} \). Assuming that \( \text{gr} A = \bigoplus_{n \geq 0} A_n/A_{n-1} \) is a domain (not necessarily commutative), show:

(a) If \( 0 \neq x, y \in A \), then \( xy \neq 0 \) and \( \deg xy = \deg x + \deg y, (xy)' = \hat{xy} \).
(b) Assume that \( 0 \neq x \in A \) is a normal element of \( A \), that is, \( xA = Ax = \langle x \rangle \).
(c) For a given normal element \( 0 \neq x \in A \), consider the algebra \( B = A/(x) \) and let \( \pi : A \to B \) be the canonical map. If \( B \) is filtered by \( B_n = \overline{A_n} \), then \( \text{gr} B \cong \text{gr} A/(\hat{x}) \).

5.4.3 (Filtrations by generators). Let \( A \) be a \( k \)-algebra. Assume that \( A \) is generated as \( k \)-algebra by \( X = (x_i)_{i \in I} \) and consider the \( X \)-filtration and the associated graded algebra \( \text{gr}_X A = \bigoplus_{n \geq 0} A_n/A_{n-1} \) as in Example 5.21. Show:

(a) The algebra \( \text{gr}_X A \) is generated by the elements \( \hat{x}_i = x_i + k \in \text{gr}^1 A \) (\( i \in I \)).
(b) \( \text{gr}_X A \) is commutative if and only if there is a \( \mathfrak{g} \in \text{Lie}_k \) and an epimorphism of associative \( k \)-algebras \( U \to A \) such that \( U_n \to A_n \) for all \( n \). Here, \( (U_n) \) denotes the standard filtration of the enveloping algebra \( U = U_{\mathfrak{g}} \) as in §5.4.3.

5.4.4 (Further properties of enveloping algebras). For any \( \mathfrak{g} \in \text{Lie}_k \), show:

(a) \( (U_\mathfrak{g})^* = k^* \).
(b) No element of \( U_{\mathfrak{g}} \setminus k \) is algebraic over \( k \).
(c) Every derivation \( d \in \text{Der} \mathfrak{g} \) extends uniquely to a derivation of the tensor algebra \( T_{\mathfrak{g}} \) and this extension maps the ideal \( L = (x \otimes y - y \otimes x - [x, y] \mid x, y \in \mathfrak{g}) \) to itself. Conclude that \( d \) extends uniquely to a derivation of the enveloping algebra \( U_{\mathfrak{g}} \) (viewing \( \mathfrak{g} \) as contained in \( U_{\mathfrak{g}} \) via the canonical map \( \mathfrak{g} \hookrightarrow U_{\mathfrak{g}} \)).

5.4.5 (Relative augmentation ideals for enveloping algebras). Let \( \mathfrak{g} \) be an ideal of the Lie algebra \( \mathfrak{g} \) and let \( \pi : \mathfrak{g} \to \mathfrak{g}/\mathfrak{a} \) be the canonical map. Put \( U = U_{\mathfrak{g}} \) and consider the map \( U\pi : U \to U_{\mathfrak{g}/\mathfrak{a}} \) in \( \text{Alg}_k \). Show that \( \text{Ker} U\pi = aU = U\mathfrak{a} \).

5.4.6 (Enveloping algebras in positive characteristic). Let \( \mathfrak{g} \in \text{Lie}_k \) be finite dimensional and assume that \( \text{char} k = p > 0 \). The goal of this exercise is to show that the
enveloping algebra $U = U_\mathfrak{g}$ is a finite module over its center; in particular, $U$ is a PI-algebra. This result is due to Jacobson [108].

(a) A polynomial of the form $\sum_{i=0}^n \lambda_i x^i \in \mathbb{k}[t]$ is called a $p$-polynomial. Show that every $0 \neq f(t) \in \mathbb{k}[t]$ is a factor of some nonzero $p$-polynomial. (Use the fact that the $t^p^i$ are linearly dependent modulo $f(t)$.)

(b) Let $x, y \in \mathfrak{g}$. Conclude from (a) that the endomorphism $\text{ad} x \in \text{End}_\mathbb{k} \mathfrak{g}$ satisfies a nonzero $p$-polynomial. Moreover, for all $i \geq 0$, show that $(\text{ad} x)^{p^i} \cdot y = x^{p^i} y - y x^{p^i}$ in $U$.

(c) Conclude from (b) that, for each $x \in \mathfrak{g}$, there is a polynomial $0 \neq f(t) \in \mathbb{k}[t]$ such that $f(x) \in \mathcal{Z} U$.

(d) Fix a $\mathbb{k}$-basis $(x_i)_{i=1}^d$ for $\mathfrak{g}$ and fix a nonzero polynomials $f_j(t)$ such that $f_j(x_i) \in \mathcal{Z} U$ as in (c). Use the Poincaré-Birkhoff-Witt Theorem (§5.4.3) to show that the monomials $x_1^{e_1} x_2^{e_2} \ldots x_d^{e_d}$ with $0 \leq e_i < \deg f_j(t)$ generate $U$ as module over $\mathcal{Z} U$. Thus, $U$ is a finite module over $\mathcal{Z} U$.

(e) Conclude from the Artin-Tate Lemma (Exercise 1.1.7) that $\mathcal{Z} U$ is an affine $\mathbb{k}$-algebra and from Exercise 1.2.7 that all irreducible representations of $\mathfrak{g}$ are finite dimensional.

5.4.7 (Enveloping algebras and skew polynomial algebras). This exercise assumes familiarity with skew polynomial algebras; see Exercise 1.1.6.

(a) Assume that $\mathfrak{g} \in \text{Lie}_\mathbb{k}$ has an ideal $\mathfrak{a}$ of codimension 1. Thus, $\mathfrak{g} \cong \mathfrak{a} \rtimes \mathbb{k} t$ is a semidirect product with $t$ acting on $\mathfrak{a}$ via $\delta = \text{ad} t|_\mathfrak{a} \in \text{Der} \mathfrak{a}$. Let $\delta$ also denote the unique extension of $\delta$ to a derivation of $U \mathfrak{a}$ (Exercise 5.4.4). Use the Poincaré-Birkhoff-Witt Theorem to show that $U \mathfrak{a}$ is a skew polynomial algebra: $U \mathfrak{a} \cong (U \mathfrak{a})[t; \delta]$.

(b) Let $\mathfrak{g} \in \text{Lie}_\mathbb{k}$ be finite dimensional and solvable. Recall from Exercise 5.3.2 that $\mathfrak{g}$ has a chain of Lie subalgebras $\mathfrak{g}_i$ such that $\mathfrak{g}_i \cong \mathfrak{g}_{i-1} \rtimes \mathbb{k} x_i$. Conclude from (a) that $U \mathfrak{g}_i \cong (U \mathfrak{g}_{i-1})[x_i; \delta_i]$ for some $\delta_i \in \text{Der} U \mathfrak{g}_{i-1}$. Thus, $U \mathfrak{g}$ is an iterated skew polynomial algebra.

(c) For the 2-dimensional non-abelian Lie algebra $\mathfrak{g} = \mathbb{k} x \oplus \mathbb{k} y$ with $[x, y] = y$, show that $U \mathfrak{g} \cong \mathbb{k}[y][x; \delta]$ with $\delta = y \frac{\partial}{\partial y} \in \text{Der} \mathbb{k}[y]$.

(d) Let $\mathfrak{h} = \mathbb{k} x \oplus \mathbb{k} y \oplus \mathbb{k} z$ be the Heisenberg Lie algebra, with $[x, y] = z$ and $[x, z] = [y, z] = 0$. Show that $U \mathfrak{h} \cong \mathbb{k}[y, z][x; \delta]$ with $\delta = z \frac{\partial}{\partial y} \in \text{Der} \mathbb{k}[y, z]$.

5.5. Generalities on Representations of Lie Algebras

Much of this section roughly follows the trajectory of §3.3.3, which covered similar material in the context of group representations, although we will now proceed at a slightly brisker pace and without any mention of characters. In addition, we introduce the representation ring of a Lie algebra and we discuss the important
“symmetrization” isomorphism between the enveloping algebra and the symmetric algebra of a Lie algebra over a base field $k$ with $\text{char } k = 0$.

Throughout this section, $\mathfrak{g} \in \text{Lie}_k$ is arbitrary. The base field $k$ continues to be arbitrary as well unless specified otherwise.

### 5.5.1. Invariants and the Trivial Representation

The counit $\varepsilon : U\mathfrak{g} \to k$ in (5.20) gives rise to a degree-1 representation, 

$$1_k = \varepsilon.$$

Explicitly, $1_k = k$ with $\mathfrak{g}$-action $x.\lambda = 0$ for all $x \in \mathfrak{g}$, $\lambda \in k$. This representation will be called the **trivial representation** of $\mathfrak{g}$; it plays a role analogous to that of the trivial representation of a group.

For an arbitrary $V \in \text{Rep} \mathfrak{g}$, the $1_k$-homogeneous component $V(1)$ is exactly the space of $\mathfrak{g}$-invariants in $V$ (§5.3.1):

$$V^1 = V(1) = \{ v \in V \mid a.v = \varepsilon(a)v \text{ for all } a \in U\mathfrak{g} \}$$
$$= \{ v \in V \mid x.v = 0 \text{ for all } x \in \mathfrak{g} \}.$$

As for groups (and in the proof of Lie’s Theorem), we will often consider more general **weight spaces** of the form

$$V_\lambda \overset{\text{def}}{=} \{ v \in V \mid x.v = \lambda(x)v \text{ for all } x \in \mathfrak{g} \}$$
$$= \{ v \in V \mid a.v = \lambda(a)v \text{ for all } a \in U\mathfrak{g} \}$$

for some $\lambda \in \left(\mathfrak{g}/[\mathfrak{g}, \mathfrak{g}]\right)^* = \text{Hom}_{\text{Lie}_k}(\mathfrak{g}, k) = \text{Hom}_{\text{Alg}_k}(U\mathfrak{g}, k)$. Any $0 \neq v \in V_\lambda$ is called a **weight vector** or semi-invariant of $\mathfrak{g}$ in $V$.

### 5.5.2. Homomorphisms

Let $V, W \in \text{Rep} \mathfrak{g}$. Then, as was explained in §3.3.3 for representations of arbitrary algebras, the $k$-vector space $\text{Hom}_k(V, W)$ becomes a representation of $U\mathfrak{g} \otimes (U\mathfrak{g})^{\text{op}}$ by $(a \otimes b^{\text{op}}).f = a_W \circ f \circ b_V$ for $a, b \in U\mathfrak{g}$ and $f \in \text{Hom}_k(V, W)$). Precomposing the resulting algebra map $U\mathfrak{g} \otimes (U\mathfrak{g})^{\text{op}} \to \text{End}_k(\text{Hom}_k(V, W))$ with the homomorphism

$$(\text{Id} \otimes S) \circ \Delta : U\mathfrak{g} \to U\mathfrak{g} \otimes U\mathfrak{g} \to U\mathfrak{g} \otimes (U\mathfrak{g})^{\text{op}},$$

where $\Delta$ and $S$ are the comultiplication (5.21) and the antipode (5.24), we obtain an algebra homomorphism $U\mathfrak{g} \to \text{End}_k(\text{Hom}_k(V, W))$. This makes $\text{Hom}_k(V, W)$ into a representation of $\mathfrak{g}$. Since $(\text{Id} \otimes S) \circ \Delta)x = x \otimes 1 - 1 \otimes x$ for $x \in \mathfrak{g}$, we have the following explicit formula:

(5.25) $x.f = x_W \circ f - f \circ x_V \quad (x \in \mathfrak{g}, f \in \text{Hom}_k(V, W))$
Exactly as in (3.30) and in Example 3.18, which established corresponding facts for group representations, one sees that

\[(5.26) \quad \text{Hom}_k(V, W)^g = \text{Hom}_{Ug}(V, W)\]

and that there is an isomorphism in Rep $g$,

\[(5.27) \quad \text{Hom}_k(\mathbb{1}, V) \xrightarrow{\sim} V\]

The Dual. The special case $W = \mathbb{1}$ of the foregoing makes the dual vector space $V^* = \text{Hom}_k(V, k)$ a representation of $g$. Formula (5.25) now becomes $(x.f)v = -f(x.v)$ for $x \in g$, $v \in V$ and $f \in V^*$ or, more generally, $a.f = f \circ S(a)_V$ for $a \in Ug$. The latter formula can also be written as follows:

\[(5.28) \quad a_{V^*} = (S(a)_V)^* \quad (a \in Ug)\]

Exactly as for groups, duality gives an exact contravariant functor

$.^* : \text{Rep} g \rightarrow \text{Rep} g$.

In particular, if $V^*$ is irreducible, then so is $V$, and the converse holds if $V$ is finite dimensional, because $V \cong V^{**}$ in this case.

5.5.3. Tensor Products

For given $V, W \in \text{Rep} g$, the tensor product $V \otimes W$ becomes a representation of $g$ via the algebra map

\[\begin{array}{cc}
Ug & \xrightarrow{\Delta} Ug \otimes Ug \xrightarrow{\text{End}_k(V) \otimes \text{End}_k(W)} \text{End}_k(V \otimes W)
\end{array}\]

The action of $g$ is explicitly given by

\[(5.29) \quad x_{V \otimes W} = x_V \otimes \text{Id}_W + \text{Id}_V \otimes x_W \quad (x \in g)\]

In particular, $V \otimes \mathbb{1} \cong V$ and $V \otimes W \cong W \otimes V$ via the switch map $\tau(v \otimes w) = w \otimes v$. It is also straightforward to check that the standard associativity isomorphism for tensor products respects the $g$-action (5.29); so the tensor product in Rep $g$ is associative.

Thus, as was also the case with group representations, Rep $g$ is a tensor category.

Finally, the reader is invited to ascertain that the canonical embedding

\[(5.30) \quad W \otimes V^* \hookrightarrow \text{Hom}_k(V, W)\]
5.5. Generalities on Representations of Lie Algebras

is a homomorphism in $\text{Rep}\mathfrak{g}$. Since the image of this embedding consists of all finite-rank $\mathbb{k}$-linear maps from $V$ to $W$, (5.30) is an isomorphism in $\text{Rep}\mathfrak{g}$ if at least one of $V$ or $W$ is finite dimensional.

5.5.4. Tensor, Exterior and Symmetric Powers

All tensor powers $V^{\otimes k}$ ($k \geq 1$) of a given $V \in \text{Rep}\mathfrak{g}$ become representations of $\mathfrak{g}$ by inductive application of (5.29). The action of an element $x \in \mathfrak{g}$ on $V^{\otimes k}$ ($k \geq 1$) is explicitly given by

\begin{equation}
(5.31) \quad x.(v_1 \otimes \cdots \otimes v_k) = \sum_i v_1 \otimes \cdots \otimes v_{i-1} \otimes x.v_i \otimes v_{i+1} \otimes \cdots \otimes v_k.
\end{equation}

Putting $V^{\otimes 0} = \mathbb{1}$, as we did for groups, the tensor algebra $TV = \bigoplus_{k \geq 0} V^{\otimes k}$ becomes a $\mathfrak{g}$-representation. All elements of $\mathfrak{g}$ act as (graded) derivations on $TV$, that is, the following Leibniz product rule holds:

\[ x.(ab) = (x.a)b + a(x.b) \quad (x \in \mathfrak{g}, \ a, b \in TV). \]

It suffices to check this equality for $a \in V^{\otimes k}, b \in V^{\otimes k'}$, in which case it follows readily from (5.31). Note also that $\mathfrak{g}$-action of $TV$ is completely determined by the given $\mathfrak{g}$-action on the degree-1 component, $V = V^{\otimes 1}$, together with the Leibniz product rule, because the tensor algebra $TV$ is generated by $V$.

Turning to the symmetric algebra, recall that $\text{Sym} V = (TV)/I$, where $I$ is the ideal of $TV$ that is generated by the Lie commutators $[v, v'] = v \otimes v' - v' \otimes v$ with $v, v' \in V$. Since $\mathfrak{g}$ acts by derivations on $TV \in \text{Alg}_\mathbb{k}$, it does so on $(TV)_{\text{Lie}} \in \text{Lie}_\mathbb{k}$ as well (Exercise 5.1.5); so $x.[v, v'] = [x.v, v'] + [v, x.v']$ for $x \in \mathfrak{g}$. Thus, the $\mathfrak{g}$-action on $TV$ stabilizes the $\mathbb{k}$-subspace that is spanned by the Lie commutators, and hence the entire ideal $I$ is $\mathfrak{g}$-stable by virtue of the Leibniz product rule. Therefore, (5.31) passes down to $\text{Sym} V$, giving $\text{Sym}^k V \in \text{Rep}\mathfrak{g}$ with $\mathfrak{g}$-action

\begin{equation}
(5.32) \quad x.(v_1 \cdots v_k) = \sum_i v_1 \cdots v_i x.v_{i+1} \cdots v_k.
\end{equation}

Again, all elements of $\mathfrak{g}$ act as graded derivations on $\text{Sym} V$ and the resulting $\mathfrak{g}$-action is the unique extension of the given action on $V = \text{Sym}^1 V$ to an action by derivations on $\text{Sym} V$.

Analogous remarks apply to the exterior algebra $\wedge V = (TV)/J$ with $J$ being generated by the squares $v^{\otimes 2} = v \otimes v$ for $v \in V$. Since

\[ x.(v^{\otimes 2}) = (x.v) \otimes v + v \otimes (x.v) = (x.v + v)^{\otimes 2} - (x.v)^{\otimes 2} - v^{\otimes 2} \in J, \]

we see as above that each $\wedge^k V$ becomes a $\mathfrak{g}$-representation, with $x \in \mathfrak{g}$ acting by

\begin{equation}
(5.33) \quad x.(v_1 \wedge \cdots \wedge v_k) = \sum_i v_1 \wedge \cdots \wedge v_{i-1} \wedge x.v_i \wedge v_{i+1} \wedge \cdots \wedge v_k.
\end{equation}

All elements of $\mathfrak{g}$ act as (graded) derivations on $\wedge V$ and this, together with the fact that the $\mathfrak{g}$-action on $\wedge V$ extends the given action on $V$, characterizes the action.
5.5.5. \( \mathfrak{g} \)-Algebras

As we have seen in §5.5.4, the tensor, symmetric and exterior algebras of a given representation of \( \mathfrak{g} \) all become representations of \( \mathfrak{g} \) in their own right, with \( \mathfrak{g} \) acting by derivations. Such algebras are called \( \mathfrak{g} \)-algebras; explicitly, these are algebras \( A \in \text{Alg}_k \) that are equipped with a Lie homomorphism,

\[
\mathfrak{g} \rightarrow \text{Der} A.
\]

Thus, \( A \in \text{Rep} \mathfrak{g} \) and the Leibniz product rule holds:

\[
x.(ab) = (x.a)b + a(x.b) \quad (x \in \mathfrak{g}, \ a, b \in A)
\]

Note that this rule states exactly that the multiplication \( A \otimes A \rightarrow A \) is a map in \( \text{Rep} \mathfrak{g} \). It also follows that \( x.1 = 0 \); so the unit map \( k = 1 \rightarrow A \) is a map in \( \text{Rep} \mathfrak{g} \) as well. Thus, \( \mathfrak{g} \)-algebras can equivalently be defined as “algebras in the category \( \text{Rep} \mathfrak{g} \).” Homomorphisms of \( \mathfrak{g} \)-algebras are defined to be algebra maps that are also maps in \( \text{Rep} \mathfrak{g} \), that is, \( \mathfrak{g} \)-equivariant algebra maps. In this way, we obtain a category,

\[
\mathfrak{g}\text{-Alg}.
\]

The tensor algebra construction for a representation of \( \mathfrak{g} \), endowed with its unique \( \mathfrak{g} \)-algebra structure extending the given representation, gives a functor \( T : \text{Rep} \mathfrak{g} \rightarrow \mathfrak{g}\text{-Alg} \). This functor is left adjoint to the forgetful functor \( \cdot : \mathfrak{g}\text{-Alg} \rightarrow \text{Rep} \mathfrak{g} \) that simply forgets the multiplication and unit of a given \( \mathfrak{g} \)-algebra as in Proposition 1.1. Similar things can be said for \( \text{Sym} \) and \( \Lambda \). The canonical epimorphisms \( TV \rightarrow \text{Sym} V \) and \( TV \rightarrow \Lambda V \) give rise to natural transformations of functors \( T \rightarrow \text{Sym}, T \rightarrow \Lambda : \text{Rep} \mathfrak{g} \rightarrow \mathfrak{g}\text{-Alg} \).

Viewing a given \( A \in \mathfrak{g}\text{-Alg} \) as a representation of \( U \mathfrak{g} \), arbitrary elements of \( U \mathfrak{g} \) do not annihilate \( 1 \in A \) but act via the counit:

\[
(5.34) \quad u.1 = \varepsilon(u)1 \quad (u \in U \mathfrak{g}).
\]

Also, writing \( \Delta u = \sum_i u_i' \otimes u_i'' \in U \mathfrak{g} \otimes U \mathfrak{g} \) for \( u \in U \mathfrak{g} \), the generalized version of the Leibniz product rule takes the form

\[
(5.35) \quad u.(ab) = \sum_i (u_i',a)(u_i'',b) \quad (u \in U \mathfrak{g}, \ a, b \in A).
\]

It is easy to see that the \( \mathfrak{g} \)-invariants in \( A \) form a subalgebra of \( A \):

\[
A^0 = \{ a \in A \mid x.a = 0 \text{ for all } x \in \mathfrak{g} \} = \{ a \in A \mid u.a = \varepsilon(u)a \text{ for all } u \in U \mathfrak{g} \}.
\]

Some of the foregoing closely mirrors our remarks about \( G \)-algebras in §3.3.3. In Section 10.4, we shall encounter a common generalization of \( \mathfrak{g} \)-algebras and \( G \)-algebras under the name of \( H \)-module algebras, where \( H \) is a general Hopf algebra; our special cases correspond to \( H = U \mathfrak{g} \) and \( H = kG \), respectively.
5.5.6. Adjoint Actions

Applying §§5.5.4, 5.5.5 to the adjoint representation $g = g_{\text{ad}}$ (Example 5.3), the tensor algebra $T_g$, the symmetric algebra $\text{Sym}_g$, and the exterior algebra $\Lambda g$ all become $g$-algebras. The $g$-action on these algebras will also be called adjoint.

It turns out that the enveloping algebra $U_g$ can be added to this list. Indeed, the Leibniz product rule together with the Jacobi identity for $g$ imply that the ideal $L$ of $T_g$ giving the relations for $U_g = T_g / L$ is stable under the adjoint $g$-action on $T_g$: for $x, y, z \in g$, we compute

$$x.(y \otimes z - z \otimes y - [y, z]) = ([x, y] \otimes z - z \otimes [x, y] - [[x, y], z])$$

$$+ (y \otimes [x, z] - [x, z] \otimes y - [y, [x, z]]) \in L.$$

Therefore, the adjoint $g$-action on $T_g$ passes down to an action on $U_g = T_g / L$, also called adjoint. Thus, $U_g \in g_{\text{Alg}}$ with corresponding Lie map

$$\text{ad} : g \rightarrow \text{Der} U_g.$$

The adjoint $g$-action on $U_g$ extends the original adjoint $g$-action on $g = g_{\text{ad}}$, viewed as contained in $U_g$ via the canonical embedding (Poincaré-Birkhoff-Witt Theorem).

As above, the map (5.36) is determined by this. Explicitly, the adjoint action of $x \in g$ on $a \in U_g$ is given by

$$x.a = [x, a] = xa - ax \quad (x \in g, \ a \in U_g),$$

because this formula holds for $a \in g$.

It follows from (5.37) that the adjoint action of $g$ on $U_g$ stabilizes every ideal $I$ of $U_g$, giving rise to another adjoint representation, $(U_g/I)_{\text{ad}}$. Since the image of $g$ generates the algebra $U_g/I$, the invariants of this representation are given by

$$(U_g/I)^g_\text{ad} = \mathcal{Z}(U_g/I).$$

All terms $U_n = U_n g$ of the standard filtration of $U_g$ are also stable under the adjoint action. Therefore, the associated graded algebra $\text{gr} U_g = \bigoplus_n U_n / U_{n-1}$ also becomes a $g$-algebra. The Poincaré-Birkhoff-Witt isomorphism (Corollary 5.25)

$$(5.39) \quad \phi : \text{Sym}_g \xrightarrow{\sim} \text{gr} U_g$$

is in fact an isomorphism of graded $g$-algebras. For $g$-equivariance, note that, in degree 1, the map $\phi$ is the canonical isomorphism $\text{Sym}^1 g \xrightarrow{\sim} g \xrightarrow{\sim} \text{gr}^1 U_g$, which is equivariant for the adjoint actions.

5.5.7. Symmetrization

We now take a closer look at the adjoint representation $(U_g)_{\text{ad}}$ in the case where $\text{char} \ k = 0$. Recall that the symmetric group $S_n$ acts on $g^{\otimes n} \subseteq T_g$ by place permutations and that the invariants of this action are isomorphic to the $n^{\text{th}}$ homogeneous
component \( \text{Sym}^n g \subseteq \text{Sym} g \) via symmetrization (Lemma 3.36):

\[
\sigma_n : \text{Sym}^n g \overset{\sim}{\longrightarrow} (g^\otimes n)^{S_n}
\]

\[
x_1 x_2 \cdots x_n \overset{\sim}{\mapsto} \frac{1}{n!} \sum_{s \in S_n} x_{s1} \otimes x_{s2} \otimes \cdots \otimes x_{sn}
\]

This map is \( g \)-equivariant for the adjoint actions on \( \text{Sym} g \) and \( T g \): for \( x, x_i \in g \),

\[
\sigma_n(x(x_1 \cdots x_n)) = \sum_i \sigma_n(x_1 \cdots x_{i-1}[x, x_i]x_{i+1} \cdots x_n)
\]

\[
= \sum_i \frac{1}{n!} \sum_{s \in S_n} x_{s1} \otimes \cdots \otimes [x, x_{si}] \otimes \cdots \otimes x_{sn}
\]

\[
= x.\sigma_n(x_1 \cdots x_n).
\]

Now let \( (U_n) \) be the standard filtration of \( U g \). Since the canonical map \( g^\otimes n \rightarrow U_n \), \( x_1 \otimes \cdots \otimes x_n \mapsto x_1 \cdots x_n \) is also \( g \)-equivariant, we obtain a \( g \)-equivariant linear map

\[
\omega_n : \text{Sym}^n g_{\text{ad}} \longrightarrow U_n
\]

\[
\omega_n : \text{Sym}^n g_{\text{ad}} \longrightarrow U_n
\]

\[
x_1 x_2 \cdots x_n \mapsto \frac{1}{n!} \sum_{s \in S_n} x_{s1} x_{s2} \cdots x_{sn}
\]

**Proposition 5.27.** Assume that \( \text{char} k = 0 \) and let \( U = (U g)_{\text{ad}} \) and \( U^n = \text{Im} \omega_n \).

Then \( U_n = U^n \oplus U_{n-1} \) and the following map is an isomorphism in \( \text{Rep} g \):

\[
\omega = \bigoplus_{n \geq 0} \omega_n : \text{Sym} g_{\text{ad}} = \bigoplus_{n \geq 0} \text{Sym}^n g_{\text{ad}} \overset{\sim}{\longrightarrow} \bigoplus_{n \geq 0} U^n = U.
\]

**Proof.** As we have seen, the canonical maps \( \pi_S : g^\otimes n \rightarrow \text{Sym}^n g \), \( \pi_U : g^\otimes n \rightarrow U_n \) and \( - : U_n \rightarrow \text{gr}^n U = U_n/U_{n-1} \) as well as the symmetrization maps \( \sigma_n : \text{Sym}^n g \rightarrow g^\otimes n \) and \( \omega_n : \text{Sym}^n g \rightarrow U_n \) are all morphisms in \( \text{Rep} g \) for the adjoint \( g \)-actions. Furthermore, \( \pi_S \circ \sigma_n = \text{Id}_{\text{Sym}^n g}, \pi_U \circ \sigma_n = \omega_n \) and \( \phi^n \circ \pi_S = - \circ \pi_U \), where \( \phi^n : \text{Sym}^n g \rightarrow \text{gr}^n U g \) is the restriction of the Poincaré-Birkhoff-Witt isomorphism (5.39). The following calculation now shows that \( \omega_n \) is injective:

\[
- \circ \omega_n = - \circ \pi_U \circ \sigma_n = \phi^n \circ \pi_S \circ \sigma_n = \phi^n.
\]

Consequently, \( \omega_n : \text{Sym}^n g \rightarrow U^n \) is an isomorphism in \( \text{Rep} g \) and the remaining assertions of the proposition also follow. \( \square \)

The map \( \omega \) in Proposition 5.27 will also be referred to as the **symmetrization map**. Note that \( \omega \) is not an algebra map. Here is an application to the faithfulness issue that was briefly raised earlier (§5.4.2).

**Proposition 5.28.** Let \( V \in \text{Rep} g \). If \( TV \) is faithful for \( U \), then \( V \) is \( g \)-faithful. The converse holds if \( \text{char} k = 0 \).
5.5. Generalities on Representations of Lie Algebras

Proof. Since \( x_{\varphi \otimes n} = \sum_{i=0}^{n-1} \text{Id}_V \otimes x_i \otimes \text{Id}_V \otimes x_i \otimes \cdots \otimes x_i \) for \( x \in g \), we certainly have the inclusion \( \text{Ker}_n V = \{ x \in g \mid x_V = 0 \} \subseteq \text{Ker}_U TV \), which implies the first assertion.

For the converse, assume that \( V \) is \( g \)-faithful and \( \text{char} \, k = 0 \). It will be advantageous to replace \( V \) by \( 1 \otimes V \); this changes neither \( \text{Ker}_n V \) nor \( \text{Ker}_U TV \), but it ensures that \( g_V \subseteq \text{End}_k(V) \) is linearly independent from \( \text{Id}_V \). Now let \( 0 \neq u \in U \) be given and let \( n \) be such that \( u \in U_n \setminus U_{n-1} \). We will show that \( u_{V \otimes n} \neq 0 \); this will prove faithfulness of \( TV \). Fix a basis \( (x_i)_{i \in I} \) of \( g \) and choose an ordering \((I, \leq)\). Then a \( k \)-basis of \( g \otimes n \) is given by the \( n \)-tensors \( x_i^\otimes = x_{i_1} \otimes x_{i_2} \otimes \cdots \otimes x_{i_n} \) for all \( n \)-term sequences of indices \( i = (i_1, i_2, \ldots, i_n) \) in \( I \). In view of the embedding \( g \otimes n \subseteq \text{End}_k(V)^{\otimes n} \), the elements \( (x_i^\otimes)_{V \otimes n} = (x_{i_1})_V \otimes (x_{i_2})_V \otimes \cdots \otimes (x_{i_n})_V \) are linearly independent and they are also linearly independent modulo the subspace \( F \subseteq \text{End}_k(V^{\otimes n}) \) that is generated by the \( n \)-tensors \( f_1 \otimes \cdots \otimes f_n \) with \( f_i \in \text{End}_k(V) \) and \( f_j = \text{Id}_V \) for at least one \( j \). Write the given \( u \in U \) as \( u = u_0 + u_- \) with \( 0 \neq u_0 \in U^n \) and \( u_- \in U_{n-1} \) (Proposition 5.27). By definition of the symmetrization map, \( u_0 \) has the form

\[
\hat{u} = \frac{1}{n!} \sum_i \alpha_i \sum_{s \in S_n} x_{i_1} x_{i_2} \cdots x_{i_n},
\]

where each \( i = (i_1 \leq i_2 \leq \cdots \leq i_n) \) is a weakly increasing sequence in \( I \) and \( \alpha_i \in k \). Observe that \((u_-)_{V \otimes n} \in F \) and, for any sequence \( i = (i_1, i_2, \ldots, i_n) \) in \( I \),

\[
(x_{i_1, i_2, \ldots, i_n})_{V \otimes n} \equiv \sum_{s \in S_n} (x_{i_1})_V \otimes (x_{i_2})_V \otimes \cdots \otimes (x_{i_n})_V \mod F.
\]

Hence

\[
\frac{1}{n!} \sum_{s \in S_n} (x_{i_1} x_{i_2} \cdots x_{i_n})_{V \otimes n} \equiv \sum_{s \in S_n} (x_{i_1})_V \otimes (x_{i_2})_V \otimes \cdots \otimes (x_{i_n})_V \mod F.
\]

The elements \( \sum_{s \in S_n} (x_{i_1})_V \otimes (x_{i_2})_V \cdots (x_{i_n})_V \) for different weakly increasing sequences \( i \) are linearly independent modulo \( F \) by what we said above. Therefore, \( \hat{u}_{V \otimes n} \) is nonzero modulo \( F \). Since \( u_{V \otimes n} = \hat{u}_{V \otimes n} + (u_-)_{V \otimes n} \equiv \hat{u}_{V \otimes n} \mod F \), it follows that \( u_{V \otimes n} \neq 0 \), as we wished to show. \( \square \)

5.5.8. The Representation Ring of a Lie Algebra

The Grothendieck group \( \mathcal{R}(A) \) of the category \( \text{Rep}_{\text{fin}} A \) of all finite-dimensional representations of a given \( k \)-algebra \( A \) was introduced in §1.5.5. Recall that \( \mathcal{R}(A) \) is a free abelian group: a basis is given by the classes \([S]\) with \( S \in \text{Irr}_{\text{fin}} A \), a full representative set of the isomorphism classes of all finite-dimensional irreducible representations of \( A \).

In the special case where \( A = U_q \), we have seen that \( \text{Rep} A = \text{Rep} g \) and \( \text{Rep}_{\text{fin}} A \equiv \text{Rep}_{\text{fin}} g \) (§5.4.2). In place of \( \mathcal{R}(U_q) \), we will also write \( \mathcal{R}(g) \).
The group $R(\mathfrak{g})$ is called the representation ring of $\mathfrak{g}$. Indeed, while $R(A)$ is generally merely an abelian group, the fact that $\text{Rep}_{\text{fin}} \mathfrak{g}$ is a tensor category (§5.5.3) allows us to endow $R(\mathfrak{g})$ with a multiplication: for any $V, W \in \text{Rep}_{\text{fin}} \mathfrak{g}$, define

$$[V] \cdot [W] \overset{\text{def}}{=} [V \otimes W].$$

Exactness of the functor $V \otimes - : \text{Rep}_{\text{fin}} \mathfrak{g} \to \text{Rep}_{\text{fin}} \mathfrak{g}$ implies that $[V] \cdot -$ is a well-defined group endomorphism of $R(\mathfrak{g})$; likewise for $- \cdot [W]$. Associativity of the product $\cdot$ follows from associativity of the tensor product in $\text{Rep} \mathfrak{g}$. Therefore, the multiplication $\cdot$ makes $R(\mathfrak{g})$ into a ring with identity element $1 = [1]$, the class of the trivial representation. In fact, $R(\mathfrak{g})$ is a commutative ring, because $V \otimes W \cong W \otimes V$ in $\text{Rep} \mathfrak{g}$.

In later sections, we will study the ring $R(\mathfrak{g})$ in more detail for certain Lie algebras $\mathfrak{g}$ and see that it has an intriguing structure (Sections 5.7.7 and 8.5). More generally, there will be an analogous ring $R(H)$, not necessarily commutative, for any Hopf algebra $H$ (Section 10.3).

**Exercises for Section 5.5**

In these exercises, $\mathfrak{g} \in \text{Lie}_k$ is arbitrary and $U = U\mathfrak{g}$.

5.5.1 (Representations and $\mathfrak{g}$-algebras). Let $V \in \text{Rep} \mathfrak{g}$ and let $\mathfrak{g}$ act on $\text{End}_k(V)$ by (5.25). Show that $\text{End}_\mathfrak{g}(V)$ is a $\mathfrak{g}$-algebra and that the representation map $U \to \text{End}_k(V)$ is a homomorphism of $\mathfrak{g}$-algebras for the adjoint $\mathfrak{g}$-action on $U$.

5.5.2 (Smash products). Let $A \in \mathfrak{g}\text{Alg}$. Show that the vector space $B = A \otimes U$ can be equipped with a $k$-algebra structure such that the maps $A \to B$, $a \mapsto a \otimes 1$, and $U \to B$, $u \mapsto 1 \otimes u$, are $k$-algebra homomorphisms and such that, identifying $A$ and $U$ with their images in $B$, the relations $xa = ax + x.a$ for $x \in \mathfrak{g}$ and $a \in A$ hold in $B$. The algebra $B$ is written as $B = A \# U$ and it is called the smash product of $A$ and $U$. When $\mathfrak{g} = k\mathfrak{g}$ is 1-dimensional, then $B$ is a skew polynomial algebra (Exercise 1.1.6): $A \# kx \cong A[x; \delta]$ with $\delta \in \text{Der} A$.

5.5.3 (Semidirect products and smash products). Let $\mathfrak{g} = a \ltimes b$ be a semidirect product Lie algebra. View $U \in \mathfrak{g}\text{Alg}$ via the adjoint action (§5.5.6) and observe that $Ua$ is a $\mathfrak{g}$-subalgebra of $U$. By restriction, $Ua$ becomes a $b$-algebra. Use the Poincaré-Birkhoff-Witt Theorem to show that $U$ is a smash product, $U \cong (Ua)\#(Ub)$. This generalizes Exercise 5.4.7.

5.5.4 (Bimodules and adjoint actions). Let $M$ be a $(U, U)$-bimodule. Using a dot-less notation for the left and right $U$-actions on $M$, define $M_{\text{ad}} := M$, equipped with the $\mathfrak{g}$-action $x.m := xm - mx$ for $x \in \mathfrak{g}$ and $m \in M$. Show that $M_{\text{ad}} \in \text{Rep} \mathfrak{g}$. Furthermore, show that the invariants are given by $M_{\text{ad}}^\mathfrak{g} = \{m \in M \mid um = um \}$.
mu for all \( u \in U \) and that \( g.M_{\text{ad}} \) is the \( \mathbb{k} \)-subspace of \( M \) that is spanned by the elements \( um - mu \) with \( u \in U, m \in M \).

5.5.5 (1-dimensional representations). Put \( g^{ab} = g/[g, g] \) and view \((g^{ab})^* \subseteq g^*\) via the canonical map \( g \to g^{ab} \). For \( \alpha \in (g^{ab})^* \), let \( \mathbb{k}_\alpha \in \text{Irr}_\mathbb{k} g \) be given by the Lie algebra map \( \alpha : g \to \mathbb{k}_\text{Lie} \). Show that \( \alpha \mapsto [\mathbb{k}_\alpha] \) gives a monomorphism of groups, \((g^{ab}, +) \hookrightarrow \mathcal{H}(g)\).

5.6. The Nullstellensatz for Enveloping Algebras

The focus in this section is on the collection \( \text{Prim} U \) of all primitive ideals of the enveloping algebra \( U = U_g \), where \( g \in \text{Lie}_\kappa \) is finite dimensional. The material presented here is mainly of interest for \( \text{char} \mathbb{k} = 0 \), although the principal results are valid for any \( \mathbb{k} \). Indeed, if \( \text{char} \mathbb{k} = p > 0 \), then all irreducible representations of \( g \) are known to be finite dimensional (Exercise 5.4.6). Thus, all primitive ideals of \( U \) have finite codimension in this case and, in particular, they are all maximal (Theorem 1.38). If \( \text{char} \mathbb{k} = 0 \), then the equality \( \text{Prim} U = \text{MaxSpec} U \) only holds if \( g \) is nilpotent, as we shall see.

This section is exclusively concerned with finite-dimensional Lie algebras. All hypotheses on \( g \in \text{Lie}_\mathbb{k} \) and the base field \( \mathbb{k} \) will be stated as needed.

5.6.1. The Statement

The ring theoretic formulation of the classical Nullstellensatz consists of the following two assertions concerning an arbitrary affine commutative \( \mathbb{k} \)-algebra \( A \) (see Section C.1):

- If \( P \in \text{MaxSpec} A \), then \( A/P \) is algebraic over \( \mathbb{k} \) (and so \( \dim_\mathbb{k} A/P < \infty \)).
- Every semiprime ideal of \( A \) is an intersection of maximal ideals of \( A \).

The first statement is the (ring theoretic version of the) weak Nullstellensatz; the second expresses the Jacobson property of \( A \). Note that affine commutative algebras \( A \) are exactly the homomorphic images of polynomial algebras in finitely many variables, that is, enveloping algebras of finite-dimensional abelian Lie algebras. Our first goal is to show that both parts of the Nullstellensatz, when properly interpreted, remain true for the enveloping algebra \( U = U_g \) of any finite-dimensional Lie algebra \( g \) and for all homomorphic images of \( U \).

For the Jacobson property, we will have to consider intersections of primitive ideals rather than just maximal ideals. While primitive and maximal ideals are identical for commutative algebras by (1.37), this generally fails for enveloping algebras: \( \text{MaxSpec} U \) is usually a proper subset of \( \text{Prim} U \); see Example 5.29 below. The noncommutative version of the weak Nullstellensatz replaces factors modulo

---

*In terms of Hochschild (co)homology, this exercise states that \( M_{\text{ad}}^0 = H^0(U, M) \) and \( M_{\text{ad}}/\text{ad}^0 M_{\text{ad}} = H_0(U, M) \); see Cartan-Eilenberg [37, Prop. IX.4.1].
maximal ideals by the Schur division algebras $D(V) = \text{End}_U(V)$ of irreducible representations $V$ of $U$. Again, for any commutative algebra $A$, all irreducible representations are equivalent to $A/P$ for some maximal ideal $P$ of $A$, and $D(A/P) \cong A/P$ (Lemma 1.5). Thus, the following result may rightfully be regarded as a version of the Nullstellensatz for enveloping algebras; note also that both (a) and (b) below pass verbatim to all homomorphic images of $U$.

**Nullstellensatz for Enveloping Algebras.** Let $\mathfrak{g} \in \text{Lie}_k$ be finite dimensional and let $U = U_\mathfrak{g}$. Then:

(a) **Weak Nullstellensatz:** For every $V \in \text{Irr} U$, the Schur division algebra $D(V) = \text{End}_U(V)$ is algebraic over $k$.

(b) **Jacobson Property:** Every semiprime ideal of $U$ is an intersection of primitive ideals of $U$.

The Jacobson property (b) was first established by Duflo [62]. The weak Nullstellensatz for $U$, originally due to Quillen [169], is of course trivial for finite-dimensional $V$ (see Schur’s Lemma) and Exercise 1.2.12 outlines a very short proof over large enough base fields $k$. The crux of (a) is that it also works for irreducible representations of $U$ that are infinite dimensional, which is typically the case if $\text{char} k = 0$, and for arbitrary base fields. Finally, we remark that both (a) and (b) also hold, and are in fact much easier, for all finite-dimensional algebras. Indeed, (a) was stated as part of Schur’s Lemma and (b) is a consequence of the fact that semiprime ideals are always intersections of primes, because prime ideals of finite-dimensional algebras are maximal (Theorem 1.38).

### 5.6.2. The Proof

The proof uses some results and arguments from commutative algebra; the link to commutative algebras is provided by the fact that the associated graded algebra $\text{gr} U$ for the standard filtration of $U = U_\mathfrak{g}$ is isomorphic to a polynomial algebra by the Poincaré-Birkhoff-Witt Theorem (Corollary 5.25). In fact, in place of $U$, we will consider an associative algebra $A \in \text{Alg}_k$ having the following property:

There exists a finite set of algebra generators $X \subseteq A$ such that the associated graded algebra $\text{gr}_X A$ (Example 5.21) is commutative.

Following Duflo [62], such algebras are called **almost commutative**. Recall that $\text{gr}_X A = \bigoplus_{n \geq 0} A_n/A_{n-1}$, where

$$A_n = \langle x_1 x_2 \ldots x_m \mid x_i \in X, m \leq n \rangle_k.$$  

It is easy to see that the associated graded algebra $\text{gr}_X A$ of any almost commutative algebra $A$ is automatically **affine** commutative and that the class of all almost commutative algebras coincides with the class of homomorphic images of enveloping algebras of finite-dimensional Lie algebras (Exercise 5.4.3). Thus, working in the
context of almost commutative algebras does not in fact yield a more general result
than the Nullstellensatz for enveloping algebras.

**Proof of (a).** The proof of the weak Nullstellensatz depends on the following result
from commutative algebra; see Section C.2 for a proof. For any commutative ring
$R$ and any $0 \neq f \in R$, we let $R_f$ denote the localization $R[1/f]$.

**Generic Flatness Lemma.** Let $R$ be a commutative domain and let $S$ be a finitely
generated commutative $R$-algebra. Then, for any finitely generated $M \in \text{Mod}_R$, there exists $0 \neq f \in R$ such that $M_f = R_f \otimes_R M$ is a free as $R_f$-module.

Now let $A$ be an almost commutative $k$-algebra, let $V \in \text{Irr} A$, and let $\phi \in D(V) = \text{End}_A(V)$ be given. Assume, by way of contradiction, that $\phi$ is transcendental over $k$; so the subalgebra $R := k[\phi] \subseteq D(V)$ is a polynomial algebra. Viewing $V$ as $R$-module in the obvious way, it suffices to show that, for some $0 \neq f \in R$, the localization $V_f = R_f \otimes_R V$ is free over $R_f$. For, then we may choose any $0 \neq g \in R$
that is not invertible in $R_f$ (e.g., take $g$ to be any irreducible polynomial not dividing $f$) to conclude that $g.V_f \nsubseteq V_f$, and hence $g.V \nsubseteq V$. But this contradicts the fact
that the action of $g$ on $V$ is invertible, because $g$ is a nonzero element of the Schur
division algebra $D(V)$.

To find $f$, we will use the Generic Flatness Lemma. View $V$ as a left module
over the algebra $A' := R \otimes A = A[\phi]$ via $(\phi \otimes a).v = a.\phi(v) = \phi(a.v)$. Filtering $A'$ by the subspaces $A'_n = R \otimes A_n$, the associated graded algebra has the form
$\text{gr } A' \cong \text{gr } X A$; this is a finitely generated commutative $R$-algebra. Fixing
$0 \neq v \in V$ and putting $V_n = A'_n.v$, we obtain a filtration $0 = V_{-1} \subseteq V_0 = R.v \subseteq \cdots \subseteq V_n \subseteq V_{n+1} \subseteq \cdots \subseteq V = \bigcup_n V_n$ such that $A'_n.V_n \subseteq V_{n+m}$. The $A'$-module structure on $V$ therefore yields an $S$-module structure on $M := \text{gr } V = \bigoplus_{n \geq 0} V_n/\text{gr } (V_{n+1})$, and this $S$-module is generated by $v \in V_0/\text{gr } (V_0) = V_0$. By the Generic Flatness Lemma, there exists $0 \neq f \in R$ such that $M_f \cong \bigoplus_n (V_n/\text{gr } (V_{n+1}))_f$ is free over $R_f$. Since $R_f$ is a PID, it follows that each $(V_n/\text{gr } (V_{n+1}))_f$ is free over $R_f$. Consequently, $V_f$ is free over $R_f$, being a successive extension of free $R_f$-modules. This proves (a).

**Proof of (b).** Replacing $A$ by $A/P$, where $P$ is a given semiprime ideal of $A$, we may assume that $A$ is a semiprime almost commutative algebra, and we need to show that the intersection of all primitive ideals of $A$ is $0$, that is, $\text{rad } A = 0$. In fact, it suffices to show that $\text{rad } A$ is nil: every element of $\text{rad } A$ is nilpotent. For, $A$ is noetherian by Lemma 5.23 and a classical theorem of Levitzki states that nil ideals of noetherian rings are in fact nilpotent (e.g., [125, 10.30]). Inasmuch as $A$ is semiprime, nilpotency of $\text{rad } A$ will force $\text{rad } A = 0$ as desired.

It remains to show that every $a \in \text{rad } A$ is nilpotent. Borrowing an argument
from commutative algebra (“Rabinowitsch trick”; see also Section C.1), we consider
the polynomial algebra $A[x]$ and prove the following

**Claim.** $A[x](1 - ax) = A[x]$. 

This will allow us to write $1 = f(1 - ax)$ for some $f \in A[x]$. Necessarily, $f = \sum_n a^n x^n$, whence $a^n = 0$ for $n > \deg f(x)$ and so $a$ is nilpotent.

Suppose the Claim is false. Then we may choose a maximal left ideal $L$ of $A[x]$ containing $A[x](1 - ax)$ and consider the representation $V := A[x]/L \in \mathrm{Irr} \ A[x]$. Putting $\nu = 1 + L \in V$, we have $\nu \neq 0$ and $\nu = ax \cdot \nu$. Thus the action of $x$ on $V$ gives a nonzero element $x_\nu \in D(V) = \mathrm{End}_{A[x]}(V)$, and hence $x_\nu$ is invertible. The equality $\nu = ax \cdot \nu$ shows that $x_\nu^{-1} \cdot \nu = a \cdot \nu$. By the weak Nullstellensatz (a) applied to the almost commutative algebra $A[x]$, we also know that $x_\nu$ is algebraic over $k$. Thus, $x_\nu = p(x_\nu^{-1})$ for some polynomial $p(x) \in k[x]$. Hence $\nu = ax \cdot \nu = ap(x_\nu^{-1}) \cdot \nu = ap(a) \cdot \nu$ and so $(1 - ap(a)) \cdot \nu = 0$. However, since $a \in \mathrm{rad} A$, the element $1 - ap(a)$ is invertible in $A$ (Exercise 1.3.3). This contradicts the fact that $\nu \neq 0$, proving the claim and thereby completing the proof of the Nullstellensatz for enveloping algebras.

### 5.6.3. Locally Closed Primes

The Nullstellensatz for enveloping algebras, specifically the Jacobson property, often enables us to verify that a certain given prime ideal $P$ of $U = U_\mathfrak{g}$ is primitive without having to produce an irreducible representation of $U$ whose kernel is $P$. Indeed, suppose we are able to show that $P \in \mathrm{Spec} \ U$ satisfies

\begin{equation}
P \not\subseteq \bigcap_{Q \in \mathrm{Spec} \ U \atop P \not\subseteq Q} Q.
\end{equation}

Then the Jacobson property tells us that $P$ must be primitive. It is easy to see that (5.42) is equivalent to the condition that the one-point set $\{P\}$ is locally closed subset of the topological space $\mathrm{Spec} \ U$, endowed with the Jacobson-Zariski topology (Exercise 5.6.3). Therefore, prime ideals $P$ satisfying (5.42) are called **locally closed**. To summarize, the Jacobson property of $U$ yields the following implication, for all $P \in \mathrm{Spec} \ U$,

\begin{equation}
P \text{ is locally closed} \implies P \text{ is primitive}
\end{equation}

We mention that the Jacobson property can be stated in purely topological terms; see Exercise 5.6.4.

**Example 5.29** ($\mathrm{Spec} \ U_\mathfrak{g}$ for the 2-dimensional non-abelian Lie algebra $\mathfrak{g}$). Consider the 2-dimensional non-abelian Lie algebra $\mathfrak{g} = \mathbb{k} x \oplus \mathbb{k} y$, $[x, y] = y$, and assume that $\text{char} \mathbb{k} = 0$. In $U = U_\mathfrak{g}$, we may write the defining relation of $\mathfrak{g}$ as $xy = y(x + 1)$. Thus, the element $y \in U$ is a normal in the sense that $Uy = yU$, and hence $(y^n) = Uy^n = y^nU$ for all $n \geq 0$. We leave it to the reader to check that if $I$ is any nonzero ideal of $U$, then $I \cap \mathbb{k}[y] = (y^n)$ for some $n \geq 0$—this is not hard to deduce from the structure of $U$ as a skew polynomial algebra over the subalgebra $\mathbb{k}[y]$ (Exercise 5.6.2). In particular, every $0 \neq P \in \mathrm{Spec} \ U$ must contain $y$, and
and hence $P$ corresponds to a prime of the polynomial algebra $U/yU \cong \mathbb{k}[x]$. Consequently,
\[
\text{Spec } U = \{(0)\} \sqcup \{P \in \text{Spec } U \mid y \in P\} \underset{\sim}{\hookrightarrow} \{(0)\} \sqcup \text{Spec } \mathbb{k}[x].
\]
By (5.43) we conclude that the zero ideal must be primitive, since it is clearly locally closed. (See also Exercise 5.6.2 for a direct argument.) The remaining primitive ideals of $U$ correspond to the maximal ideals of $\mathbb{k}[x]$. Thus,
\[
\text{Prim } U \hookrightarrow \{(0)\} \sqcup \text{MaxSpec } \mathbb{k}[x].
\]
The part of Prim $U$ corresponding to MaxSpec $\mathbb{k}[x]$ consists of the maximal ideals of $U$; they are exactly the ideals of the form $(y, f(x))$ with $f(x) \in \mathbb{k}[x]$ monic irreducible. Figure 5.1 gives two different renderings of Spec $U$ in the style of Examples 1.22–1.24. In both cases primitive ideals are marked in red. In the picture on the left, lines represent inclusions. The large read area in the second picture represents the zero ideal $(0)$; it is a generic point for the topological space Spec $U$ in the sense that the closure of this point is all of Spec $U$. The smaller red dots, corresponding to the ideals $(y, f(x))$ with $f(x) \in \mathbb{k}[x]$ monic irreducible, are closed points of Spec $U$; they all belong to the closure of the prime $(y)$, which is represented by the black region. Finally, $(y)$ is determined by the ideals $(y, f(x))$, being equal to their intersection.

![Figure 5.1. Spec $U$ for the Lie algebra $\mathfrak{g} = \mathbb{k}x \oplus \mathbb{k}y, [x, y] = y$](image)

### 5.6.4. Central Characters

For any $A \in \text{Alg}_k$, there is a well-defined map

\[
\begin{align*}
\text{Spec } A & \longrightarrow \text{Spec } \mathcal{Z} A \\
\psi & \psi \\
P & \hookrightarrow P \cap \mathcal{Z} A
\end{align*}
\]

where $\mathcal{Z}A$ denotes the center of $A$. This map is continuous for the Jacobson-Zariski topologies on Spec $A$ and Spec $\mathcal{Z}A$; see Exercise 1.3.5 for a somewhat more general fact. If the weak Nullstellensatz holds for $A$, then (5.44) sends Prim $A$ to Prim $\mathcal{Z}A = \text{MaxSpec } \mathcal{Z}A$. Indeed, choosing $V \in \text{Irr } A$ with $P = \text{Ker } V$, the
image of \( \mathcal{Z}(A/P) \) in \( \text{End}_k(V) \) is contained in \( D(V) = \text{End}_A(V) \), which is algebraic over \( k \) by assumption. Therefore, \( \mathcal{Z}(A/P) \) is an (algebraic) extension field of \( k \), and hence so is the subalgebra \( \mathcal{Z}A/P \cap \mathcal{Z}A \) of \( \mathcal{Z}(A/P) \).

By the weak Nullstellensatz for enveloping algebras, all this holds for the enveloping algebra \( U = U_0 \) of any finite-dimensional Lie algebra \( g \). Consequently, \( \mathcal{Z}(U/P) \) is an algebraic extension field of \( k \) for each \( P \in \text{Prim} \mathcal{Z}U \) and we have a (continuous) map

\[
\begin{array}{ccc}
\text{Prim } U & \longrightarrow & \text{MaxSpec } \mathcal{Z}U \\
P & \longmapsto & P \cap \mathcal{Z}U
\end{array}
\]

We also know that the field \( \mathcal{Z}U/P \cap \mathcal{Z}U \) is algebraic over \( k \). Therefore, if \( k \) is algebraically closed and \( \rho: U \to \text{End}_k(V) \) is an irreducible representation, then \( \rho(\mathcal{Z}U) \) consists of scalar operators. The resulting algebra map

\[
(5.46) \quad \rho|_{\mathcal{Z}U}: \mathcal{Z}U \to k
\]

is called the central character of \( V \).

5.6.5. \( \text{Prim } U_0 \) for Nilpotent \( g \)

As an application of the foregoing, we present an intrinsic characterization of primitive ideals of enveloping algebras of finite-dimensional nilpotent Lie algebras. The following proposition, which is of independent interest, states some crucial observations needed for the proof of this result.

**Proposition 5.30.** Let \( g \in \text{Lie}_k \) be finite dimensional and nilpotent and let \( I \) be an ideal of the enveloping algebra \( U = U_0 \). Then:

(a) \( I \) is prime if and only if \( \mathcal{Z}(U/I) \) is a domain;

(b) \( I \) is maximal if and only if \( \mathcal{Z}(U/I) \) is a field.

**Proof.** The conditions on \( \mathcal{Z}(U/I) \) in (a) and (b) are clearly necessary, even for any algebra in place of \( U \). The reverse implications are consequences of the following Claim. If \( I, J \) are ideals of \( U \) such that \( I \nsubseteq J \), then \( (J/I) \cap \mathcal{Z}(U/I) \neq 0 \).

To prove this, consider the standard filtration \( (U_n) \) of \( U \) and choose \( n \) so that \( V := J \cap U_n/I \cap U_n \neq 0 \). Then \( V \in \text{Rep}_{\text{fin}} g \) via the adjoint action of \( g \) on \( U \) and \( V^g \subseteq (J/I) \cap \mathcal{Z}(U/I) \) by (5.38). Since \( g \) is nilpotent, \( \text{ad } x \) is a nilpotent operator on \( g \) for each \( x \in g \) (Corollary 5.16). Since \( x \) acts by derivations on \( U_n \), this action is nilpotent as well and hence so is the operator \( x_V \). Engel’s Theorem therefore implies that \( V^g \neq 0 \), proving the claim.

In particular, if \( \mathcal{Z}(U/I) \) is a field, then \( I \) must be maximal by the Claim. This proves (b). Finally assume that \( \mathcal{Z}(U/I) \) is a domain and consider ideals \( J_1, J_2 \nsubseteq I \).
The claim allows us to pick elements $u_i \in J_i \setminus I$ that are central modulo $I$. Then $u_1u_2 \notin I$ and so $J_1J_2 \notin I$, proving that $I$ is prime. \hfill\Box

The following result is due to Dixmier [56]. With $\mathfrak{g}$ abelian and $k$ algebraically closed, Dixmier’s Theorem gives the classical equivalences (1.38).

**Theorem 5.31.** Let $\mathfrak{g} \in \mathfrak{Lie}_k$ be finite dimensional and nilpotent and let $U = U_{\mathfrak{g}}$. The following are equivalent for $P \in \text{Spec } U$:

(i) $P$ is primitive;
(ii) $\mathcal{Z}(U/P)$ is an algebraic extension field of $k$;
(iii) $\mathcal{Z}(U/P)$ is a field;
(iv) $P$ is maximal.

**Proof.** We have already pointed out above that (i) $\Rightarrow$ (ii), for any finite-dimensional $\mathfrak{g} \in \mathfrak{Lie}_k$ (§5.6.4). The implication (ii) $\Rightarrow$ (iii) is trivial and Proposition 5.30(b) gives (iii) $\Rightarrow$ (iv). Finally, (iv) $\Rightarrow$ (i) holds for any algebra. \hfill\Box

Assuming $\text{char } k = 0$, the equality $\text{Prim } U = \text{MaxSpec } U$ characterizes nilpotent Lie algebras among all finite-dimensional Lie algebras: if $\mathfrak{g} \in \mathfrak{Lie}_k$ is finite dimensional and not nilpotent, then $U = U_{\mathfrak{g}}$ contains a non-maximal primitive ideal. Indeed, if $\mathfrak{g}$ is solvable, then $\mathfrak{g}$ maps onto the 2-dimensional non-abelian Lie algebra (Exercise 5.3.3). Pulling back the zero ideal from Example 5.29, we obtain the desired primitive ideal. If $\mathfrak{g}$ is not solvable, then $\mathfrak{g}/\text{rad } \mathfrak{g}$ is nonzero semisimple. The enveloping algebra of any nonzero semisimple Lie algebra $\mathfrak{g}$ is known to have non-maximal primitive ideals. For example, the ideal $(U^+ \cap \mathcal{Z})U$ is primitive and strictly contained in the augmentation ideal $U^+$ of $U$; see Jantzen [112, 7.3].

**Example 5.32** (Spec $U_{\mathfrak{h}}$ for the Heisenberg Lie algebra $\mathfrak{h}$). Let $\mathfrak{h} = kx \oplus ky \oplus k\mathfrak{z}$ with $[x, y] = z$ and central $z$, and assume that $k$ is algebraically closed with $\text{char } k = 0$. By the Poincaré-Birkhoff-Witt Theorem, we may view the enveloping algebra $U = U_{\mathfrak{h}}$ as an iterated skew polynomial algebra (Exercise 5.4.7):

$$U \cong k[y, z][x; \delta] \quad \text{with } \delta = z \frac{\partial}{\partial y} \in \text{Der } k[y, z]$$

Using the relations $[x, y^i] = izy^{i-1}$ and $[y^i, y] = izx^{i-1}$ in $U$, it is easy to see that $\mathcal{Z} U = k[z]$. Thus, if $0 \neq P \in \text{Spec } U$, then it follows from the Claim in the proof of Proposition 5.30 that $P \cap k[z] = (z - \lambda)$ for some $\lambda \in k$. The ideals $P_{\lambda} := (z - \lambda)U$ are all prime. Indeed, $U/P_{\lambda} = k[x, y]$ is a polynomial algebra in two variables, and $U/P_{\lambda}$ for $\lambda \neq 0$ is isomorphic to the Weyl algebra $A_1(k)$, because $[y^{-1}x, y] = y^{-1}z \equiv 1 \text{ mod } P_{\lambda}$. Therefore, in terms of the Jacobson-Zariski topology (§1.3.4), the closed set $\mathcal{V}(z)$ of Spec $U$ can be identified with Spec $k[x, y]$, while the complement $\mathcal{V}(z)^C$ consists of the zero ideal and the closed
sets \( \mathcal{V}(z - \lambda) = \{ P_\lambda \} \) with \( \lambda \in k^\times \). Thus, the open set \( \mathcal{V}(z)^\complement \) can be identified with the spectrum of the Laurent polynomial algebra \( k[z^{\pm 1}] \). To summarize,

\[
\text{Spec } U \overset{\sim}{\longrightarrow} \text{Spec } k[\overline{x}, \overline{y}] \sqcup \text{Spec } k[z^{\pm 1}]
\]

and

\[
\text{Prim } U = \text{MaxSpec } U \overset{\sim}{\longrightarrow} \text{MaxSpec } k[\overline{x}, \overline{y}] \sqcup \text{MaxSpec } k[z^{\pm 1}]
\]

\[
\overset{\sim}{\longrightarrow} k^2 \sqcup k^x
\]

Figure 5.2 is our attempt at a visual rendering of Spec \( U \), with red points marking primitive (maximal) ideals. The large gray oval represents the generic point \((0)\) of Spec \( U \); the diagonal red line is MaxSpec \( k[z^{\pm 1}] = \{(z - \lambda) \mid \lambda \in k^\times\} \), with a gap for \( \lambda = 0 \); the plane in the picture is the closed set \( \mathcal{V}(z) = \text{Spec } k[\overline{x}, \overline{y}] \), depicted as earlier (Figure 1.2).

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure52.png}
\caption{Spec \( U_h \) for the Heisenberg Lie algebra \( \mathfrak{h} \)}
\end{figure}

### 5.6.6. Rational Ideals and the Dixmier-Mœglin Equivalence

The value of the characterization of primitivity of a prime \( P \in \text{Spec } U \) in terms of the center \( Z(U/P) \) in Theorem 5.31 is at least twofold:

- It allows us to check primitivity of \( P \) without the need to produce an irreducible representation of \( U \) whose kernel is \( P \); and
- it shows that primitivity is a left-right symmetric condition for primes of \( U \)—this is not the case for arbitrary algebras (Section 1.3).

Unfortunately, Theorem 5.31 fails for non-nilpotent Lie algebras. However, it turns out that, replacing \( Z(U/P) \) by the center of a suitable quotient ring of \( U/P \), it is possible to give an analogous characterization of primitivity that is valid for
5.6. The Nullstellensatz for Enveloping Algebras

arbitrary finite-dimensional Lie algebras and offers the same benefits. The complete
proof this result, which is known as the Dixmier-Mœglin equivalence, is outside
the scope of this book. In this subsection, we do at least give the full statement
of the Dixmier-Mœglin equivalence and the proof in the important special case of
solvable Lie algebras.

The Symmetric Ring of Quotients and the Extended Center. First, we need to
explain some ring theoretic background. In this paragraph, \( A \) be an arbitrary ring
(associative, with 1). We will give a brief description of the symmetric ring of
quotients, \( QA \), originally due to Kharchenko \[121\], \[122\], and its center \( Z(QA) \).
The reader wishing to see the details of the construction of \( QA \) and proofs of
the various assertions made in the following may indulge his or her curiosity by
perusing Appendix E.

Let \( \mathcal{E} = \mathcal{E}(A) \) denote the collection of all ideals \( I \) of \( A \) having zero right and
left annihilator in \( A \). The symmetric ring of quotients, \( QA \), is a ring having the
following properties and being determined by them up to isomorphism:

(i) \( QA \) contains \( A \) as a subring;

(ii) for each \( q \in QA \), there exists \( I \in \mathcal{E} \) with \( qI \subseteq A \) and \( Iq \subseteq A \);

(iii) each \( I \in \mathcal{E} \) has zero right and left annihilator in \( QA \);

(iv) given \( I \in \mathcal{E} \) and maps \( f : A \rightarrow A, a \mapsto af \), in \( \text{AMod} \) and \( g : I \rightarrow A \),
\( b \mapsto gb \), in \( \text{Mod}_A \) such that \( (af)b = a(gb) \) for all \( a, b \in I \), there exists
\( q \in QA \) with \( aq = af \) and \( qb = gb \) for all \( a, b \in I \).

We will mostly be concerned with the center of \( QA \); it is called the extended center
of \( A \) and will be denoted by

\[
CA \overset{\text{def}}{=} Z(QA)
\]

The extended center coincides with the centralizer of \( A \) in \( QA \):

\[
CA = \{ q \in QA \mid qa = aq \forall a \in A \} = \{ q \in QA \mid \exists I \in \mathcal{E} : qa = aq \forall a \in I \}.
\]

In particular, \( CA \) contains the ordinary center \( Z(A) \). The following proposition
summarizes the operative facts about the extended center; see Appendix E for the
proof.

Proposition 5.33. (a) If \( A \) is prime, then the extended center \( CA \) is a field.
Furthermore, if \( A \) is semiprime and \( CA \) is a domain, then \( A \) is prime.

(b) If \( V \in \text{Irr} A \) is faithful, then the embedding \( Z(A) \hookrightarrow Z(D(V)), a \mapsto a_V, \)
extends to an embedding of fields, \( CA \hookrightarrow Z(D(V)) \).
We will apply Proposition 5.33 to quotients of the form \( A/P \) with \( P \in \text{Spec } A \). Thus, \( C(A/P) \) is a field by (a), called the heart (or the core) of the prime ideal \( P \); see, e.g., Borho-Gabriel-Rentschler [21] and Dixmier [60]. We remark that the original literature on enveloping algebras employs the so-called classical ring of quotients rather than \( QA \). However, the centers of both quotient rings coincide in all situations that will be of concern to us (Appendix E).

**Example 5.34** (The heart of primes of \( U_g \) for \( g \) nilpotent). Let \( g \) be finite dimensional and nilpotent. Then the heart of any prime ideal \( P \) of \( U = U_g \) is the same as the field of fractions of \( \mathcal{Z}(U/P) \):

\[
C(U/P) = \text{Fract } \mathcal{Z}(U/P).
\]

Indeed, putting \( A = U/P \), we know that \( C(A) \) is always a field containing \( \mathcal{Z}(A) \), and so certainly \( C(A) \supseteq \text{Fract } \mathcal{Z} A \), even when \( g \) is not nilpotent or finite dimensional. For the reverse inclusion, let \( q \in C(A) \) and consider the ideal \( I = \{ a \in A \mid \text{aq} \in A \} \) of \( A \). Since \( I \neq 0 \) by property (ii) of \( QA \), we know by the Claim in the proof of Proposition 5.30 that there exists an element \( 0 \neq a \in I \cap \mathcal{Z} A \). Thus, \( b = \text{aq} \in \mathcal{Z} A \) and \( q = a^{-1}b \in \text{Fract } \mathcal{Z} A \), proving the claimed quality.

**Rational Ideals.** The material on symmetric rings of quotients and extended centers developed so far was purely ring theoretic and made no reference to a base field. Now let us return to our standard setting and assume that \( A \) is a \( k \)-algebra. Then it follows from (5.47) that \( QA \) and \( CA \) are \( k \)-algebras as well, with \( A \) being a subalgebra of \( QA \) and \( \mathcal{Z} A \) a subalgebra of \( CA \). Moreover, Proposition 5.33 tells us that the heart \( C(A/P) \) of any prime \( P \in \text{Spec } A \) is a \( k \)-field. The prime \( P \) is called rational if \( C(A/P) \) is algebraic over \( k \).

If \( P \) is primitive, say \( P = \text{Ker } V \) with \( V \in \text{Irr } A \), then we know by Proposition 5.33 that \( C(A/P) \) is a \( k \)-subfield of \( \mathcal{Z}(D(V)) \). Thus, if \( A \) satisfies the weak Nullstellensatz, then all primitive ideals of \( A \) are rational. In conjunction with our earlier observation (5.43) about locally closed primes and the Jacobson property, we obtain the following implications, valid for any prime \( P \in \text{Spec } A \) as long as the algebra \( A \) satisfies the weak Nullstellensatz and has the Jacobson property:

<table>
<thead>
<tr>
<th>Property</th>
<th>Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P ) is locally closed</td>
<td>( P ) is primitive</td>
</tr>
<tr>
<td>Jacobson property</td>
<td>weak Nullstellensatz</td>
</tr>
</tbody>
</table>

**The Dixmier-Mœglin Equivalence.** By the Nullstellensatz for enveloping algebras, the above implications certainly hold for primes of enveloping algebras of finite-dimensional Lie algebras. In fact, more is true:

**Dixmier-Mœglin Equivalence for Enveloping Algebras.** Let \( g \in \text{Lie}_k \) be finite dimensional and let \( U = U_g \). Then, for any \( P \in \text{Spec } U \),

<table>
<thead>
<tr>
<th>Property</th>
<th>Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td>( P ) is locally closed</td>
<td>( P ) is primitive</td>
</tr>
</tbody>
</table>

\[ P \text{ is locally closed } \iff P \text{ is primitive } \iff P \text{ is rational } \]
In light of the foregoing, it suffices to prove the implication “$P$ rational $\Rightarrow P$ locally closed” in order to establish the Dixmier-Mœglin equivalence for enveloping algebras. We will do this for the special case of solvable Lie algebras below. For nilpotent Lie algebras, we already know the Dixmier-Mœglin equivalence:

**Example 5.35 (The Dixmier-Mœglin equivalence for $\mathfrak{g}$ nilpotent).** Let $\mathfrak{g}$ be finite dimensional and nilpotent. Then $\mathcal{C}(U/P) = \text{Fract } \mathcal{Z}(U/P)$ (Example 5.34), and so

\[ P \text{ is rational } \iff \mathcal{Z}(U/P) \text{ is an algebraic extension field of } \mathbb{k}. \]

By Theorem 5.31, the right hand side of this equivalence implies that $P$ is maximal; in other words, $P$ is closed in $\text{Spec } U$ and so $P$ is certainly locally closed. Thus, Theorem 5.31 gives the Dixmier-Mœglin equivalence for finite-dimensional nilpotent Lie algebras, and we may even add “$P$ is maximal” to the list of equivalent properties in this case.

The Dixmier-Mœglin equivalence for enveloping algebras holds for arbitrary base fields $\mathbb{k}$. The main difficulty lies in the case where $\text{char } \mathbb{k} = 0$; in positive characteristics, more can be said rather easily:

**Example 5.36 (The Dixmier-Mœglin equivalence for $\text{char } \mathbb{k} = p > 0$).** If $\text{char } \mathbb{k} = p > 0$, then the enveloping algebra $U = U_{\mathfrak{g}}$ of any finite-dimensional $\mathfrak{g} \in \text{Lie } \mathbb{k}$ is a finite module over its center, $\mathcal{Z} = \mathcal{Z} U$, and $\mathcal{Z}$ is an affine commutative $\mathbb{k}$-algebra (Exercise 5.4.6). Thus, if $P \in \text{Spec } U$ is rational, then the inclusion $\mathcal{C}(U/P) \supseteq \mathcal{Z}/P \cap \mathcal{Z}$ yields that $\mathcal{Z}/P \cap \mathcal{Z}$ is finite dimensional over $\mathbb{k}$, and hence $\dim_{\mathbb{k}} U/P < \infty$. But then Theorem 1.38 tells us that $P$ is maximal, which trivially implies that $P$ is locally closed as in Example 5.35 above. Thus, the Dixmier-Mœglin equivalence holds in positive characteristics, and we may also add “$P$ is maximal” and “$P$ has finite codimension” to the equivalence.

**Proof of the Dixmier-Mœglin Equivalence for Solvable $\mathfrak{g}$**. This result is also due to Dixmier [57]. The basic strategy of the proof below, which roughly follows Borho-Gabriel-Rentschler [21], elaborates on the proof of the Dixmier-Mœglin equivalence for nilpotent Lie algebras (Example 5.35). The crucial technical point in this case was the Claim in the proof of Proposition 5.30, which in turn ultimately relied on Engel’s Theorem. A similar role in the current proof will be played by Lie’s Theorem. This requires char $\mathbb{k} = 0$—we may certainly assume this by Example 5.36—and we also need $\mathbb{k}$ to be algebraically closed. We allow ourselves to assume that as well; the Dixmier-Mœglin equivalence for general $\mathbb{k}$ then follows without much difficulty by a scalar extension argument. See Dixmier [60, 4.5.7].

Let $\mathfrak{g} \in \text{Lie } \mathbb{k}$ be finite dimensional and solvable and let $P$ be a rational ideal of $U = U_{\mathfrak{g}}$. Put $A = U/P$; so $CA = \mathbb{k}$. Our goal is to show that the intersection of all nonzero primes of $A$ is nonzero. For this, we consider the adjoint action of $\mathfrak{g}$ on $A$ and write the corresponding representation as $A_{\text{ad}}$ as in §5.5.6. While the Claim in the proof of Proposition 5.30 was concerned with the $\mathfrak{g}$-invariants, $A_{\text{ad}}^{\mathfrak{g}} = \mathcal{Z} A$, ...
we now need to consider weight spaces $A_A = \{ a \in A_{ad} \mid x.a = \lambda(x)a \text{ for all } x \in g \}$ with $\lambda \in \Hom_{\text{Lie}}(g, k_{\text{Lie}}) \subseteq g^*$. 

**Claim.** $\dim_k A_A \leq 1$ for all $\lambda$.

To prove this, assume that there is some $0 \neq a \in A_A$. By (5.37), we may rewrite the equation $x.a = \lambda(x)a$ for $x \in g$ in the form $x.a = a(x + (x)\lambda)$ with $x = x + P \in A$. Thus, $a$ is a normal element of $A$, that is, $Aa = aA = (a)$ is a nonzero ideal of $A$.

Since the algebra $A$ is prime, it follows that $(a) \in \mathcal{E} = \mathcal{E}(A)$ and that $a$ is a non-zero divisor of $A$. Thus, we obtain an automorphism $\tau_a \in \Aut_{\text{Alg}}(A)$ by $ba = a\tau_a(b)$ for $b \in A$. The formula $\tau_a(\lambda) = \lambda + \lambda(x)$ for the generators $\lambda$ of $A$ shows that $\tau_a$ only depends on $\lambda$ and not on the specific chosen $0 \neq a \in A_A$. Given any other $a' \in A_A$, we may consider the map $f : (a) \rightarrow A$, $a'(x) = \tau_{a'}(x)$ determines $\tau_{a'}(x)$. Since $f$ is plainly an $(A, A)$-bimodule map, we may use property (iv) of QA (with $f = g$) to deduce the existence of an element $q \in QA$ with $f = f(b) = qb$ for all $b \in (a)$. Therefore, $q \in CA$ by (5.47) and $a' = f(a) = qa$. Our rationality hypothesis $CA = \k$ therefore implies that $a' \in \mathcal{E}$. Hence, $\dim_k A_A = 1$ as we have claimed.

Now consider the set of weights $\Lambda := \{ \lambda \in g^* \mid A_A \neq 0 \}$; this is an additive subsemigroup of $g^*$. Indeed, since $g$ acts by derivations on the algebra $A$, it follows that $A_A A_{\lambda'} \subseteq A_{\lambda + \lambda'}$, and since nonzero weight vectors are non-zero divisors in $A$, we have $A_A A_{\lambda'} \neq 0$ for $\lambda, \lambda' \in \Lambda$. Thus, by the above Claim, we have

$$0 \neq A_A A_{\lambda'} = A_{\lambda + \lambda'}$$

for $\lambda, \lambda' \in \Lambda$. Let $\Lambda_\pm$ denote the additive subgroup of $g^*$ that is generated by $\Lambda$; so $\Lambda_\pm$ consists of all differences $\lambda - \lambda'$ with $\lambda, \lambda' \in \Lambda$.

**Claim.** $\Lambda_\pm$ is finitely generated.

To see this, let $(U_n)$ be the standard filtration of $U$ and observe that each $A_A (\lambda \in \Lambda)$ is contained in some $A_A = U_n/P \cap U_n$, a finite-dimensional subrepresentation $A_{ad}$. Hence $A_A$ is a composition factor of $A_n$. Since $A_n$ is an image of $T_n g = \bigoplus_{i \leq n} g_{\otimes i}$, also equipped with the adjoint $g$-action, all composition factors of $A_n$ must occur among the composition factors of the various $g_{\otimes i}$ by the Jordan-Hölder Theorem. By Lie’s Theorem, the composition factors of $g_{\otimes i}$ are 1-dimensional; so they have the form $k_{\gamma_j}$ for suitable $\{\gamma_j\}_1 \subseteq \Hom_{\text{Lie}}(g, k_{\text{Lie}})$, where $k_{\gamma_j} = k$ and $x \in g$ acts by $x.1 = \gamma_j(x)$. Therefore, the composition factors of $g_{\otimes i}$ are equivalent to the $i$-fold tensor products $k_{\gamma_1} \otimes \cdots \otimes k_{\gamma_i} \cong k_{\sum_i \gamma_i}$. Since our given $\lambda \in \Lambda$ occurs among them, we conclude that $\Lambda \subseteq \sum_i \mathbb{Z}_+ \gamma_i$ and so $\Lambda_\pm \subseteq \sum_i \mathbb{Z}_+ \gamma_i$, proving the claim.

Fix finitely many weights $\lambda_1, \ldots, \lambda_s \in \Lambda$ that generate the group $\Lambda_\pm$ and put $\lambda_0 = \sum_i \lambda_i \in \Lambda$. Since $A_{\lambda_0}$ is a non-zero, the following claim will finish the proof.

**Claim.** Every nonzero prime ideal of $A$ contains $A_{\lambda_0}$. 

Indeed, if $0 \neq P \in \text{Spec } A$, then $P \cap A_n$ is a nonzero finite-dimensional subrepresentation of $A_{\text{ad}}$ for some $n$, and so $A_\lambda \subseteq P$ for some $\lambda \in \Lambda$ by Lie’s Theorem. Write $\lambda = \sum n_i \lambda_i$ with $n_i \in \mathbb{Z}$. Replacing $\lambda$ by $m\lambda_0 + \lambda$ and $A_\lambda$ by $A_{\lambda + m\lambda_0} = A_\lambda A_{\lambda_0}^m$ for some $m \in \mathbb{Z}_+$ if necessary, we may assume that all $n_i \geq 0$. Thus, $A_\lambda = A_{\lambda_1}^{n_1} A_{\lambda_2}^{n_2} \cdots A_{\lambda_q}^{n_q} \subseteq P$. Since each $A_{\lambda_i}$ is generated by a normal element of $A$, the prime $P$ must contain one of them. Therefore, $P$ also contains $A_{\lambda_0} = A_{\lambda_1} A_{\lambda_2} \cdots A_{\lambda_q}$, which proves the claim and hence the theorem. \hfill \Box

Outlook. Things become substantially more difficult for non-solvable Lie algebras. Over an uncountable algebraically closed field $k$ of characteristic 0, the general equivalence “$P$ is primitive $\iff$ $P$ is rational” is due to Dixmier [59], while Mœglin [151] proved “$P$ is primitive $\iff$ $P$ is locally closed.” The restrictions on $k$ were removed by Irving and Small [105]. The reader may consult Rentschler [174] for a survey on primitive ideals of enveloping algebras for general finite-dimensional Lie algebras. More recently, the Dixmier-Mœglin equivalence has been proven for other algebras besides enveloping algebras. The monograph Brown-Goodearl [34] offers a panoramic overview of the Dixmier-Mœglin equivalence in the context of algebraic quantum groups.

Exercises for Section 5.6

5.6.1 (Ideals in skew polynomial algebras). Let $A \in \text{Alg}_k$ and let $B = A[x; \delta]$ be a skew polynomial algebra with $\delta \in \text{Der } A$ (Exercise 1.1.6).

(a) For $a \in A$ and $n \in \mathbb{Z}_+$, show that $x^n a = \sum_{i=0}^n \binom{n}{i} \delta^i(a) x^{n-i}$ in $B$.

(b) Assume that $A$ is a commutative domain, $\text{char } k = 0$ and $\delta \neq 0$. Show that $I \cap A$ is a nonzero $\delta$-stable ideal of $A$ for every nonzero ideal $I$ of $A$.

5.6.2 (The 2-dimensional non-abelian Lie algebra). Let $g = kx \oplus ky$ with $[x, y] = y$ and let $U = U_g$. Assume that char $k = 0$.

(a) Recall from Exercise 5.4.7 that $U \cong k[y][x; \delta]$ with $\delta = y \frac{d}{dy} \in \text{Der } k[y]$. Use Exercise 5.6.1 to show that if $I$ is a nonzero ideal of $U$, then $I \cap k[y] = (y^n)$ for some $n \geq 0$.

(b) Show that $V = U/U(y - 1)$ is a faithful irreducible representation of $U$. Thus, $U$ is primitive.

5.6.3 (Locally closed primes). (a) A subset $S$ of an arbitrary topological space $X$ is said to be locally closed if $S$ is open in its closure $\overline{S}$, that is, $S = \overline{S} \cap S^\circ$ is a closed subset of $X$. Show that this is equivalent to $S$ being closed in some neighborhood of $S$ in $X$ and also to $S$ being an intersection of a closed subset and an open subset of $X$.

(b) Let $P \in X = \text{Spec } A$ for $A \in \text{Alg}_k$, with the Jacobson-Zariski topology. Show that $\{P\} \subseteq X$ is locally closed if and only if $P \subseteq \cap \{Q \in \text{Spec } A \mid P \subseteq Q\}$. 

5.6.4 (Jacobson property). Let $X$ and $Y$ be topological spaces. A map $f : X \to Y$ is said to be a **quasi-homeomorphism** if $C \mapsto f^{-1}(C)$ gives a bijection between the collections of closed subsets of $Y$ and $X$ [92, 2.7.2].

Show that the Jacobson property for $A \in \text{Alg}_k$ is equivalent to the inclusion $\text{Prim} A \hookrightarrow \text{Spec} A$ being a quasi-homeomorphism.

5.6.5 (Nullstellensatz and the quantum plane). The Nullstellensatz is known to hold for any $A \in \text{Alg}_k$ having a sequence of subalgebras $k = A_0 \subseteq A_1 \subseteq \cdots \subseteq A_t = A$ such that, for all $i > 0$, either $A_i$ is finitely generated as $A_{i-1}$-module on each side, or $A_i$ is generated as $k$-algebra by $A_{i-1}$ together with an element $x_i$ such that $A_{i-1}x_i + A_{i-1} = x_iA_{i-1} + A_{i-1}$. See [34, II.7.17] or [149, 9.4.21].

Show that the quantum plane $A = O_q(k^2)$ has such a sequence of subalgebras (with $t = 2$). Conclude from the description of Spec $A$ in Example 1.24 that $(0)$ is a locally closed prime ideal of $A$ if $q \in k^\times$ is not a root of unity. Deduce further that $(0)$ is a primitive ideal of $A$ in this case. (This was already proved more directly in Exercise 1.3.4.)

5.7. Representations of $\mathfrak{sl}_2$

The main goal of this section is to determine the finite-dimensional representations of the Lie algebra $\mathfrak{sl}_2 = \mathfrak{sl}_2(k)$ from scratch using little more than basic linear algebra (Theorem 5.39). We will also describe the representation ring of $\mathfrak{sl}_2$ as well as the prime and primitive ideals of the enveloping algebra $U(\mathfrak{sl}_2)$. Much of the material in this section will soon be superseded by more general developments for arbitrary semisimple Lie algebras (Chapters 6–8). However, the representation theory of $\mathfrak{sl}_2$ serves as a model for the general case. In fact, as we will see in due course, the structural analysis of arbitrary semisimple Lie algebras makes use of certain Lie subalgebras isomorphic to $\mathfrak{sl}_2$, the so-called $\mathfrak{sl}_2$-**triples**.

Throughout this section, the base field $k$ is assumed to have characteristic 0 and to be algebraically closed. Also, $\mathfrak{g} = \mathfrak{sl}_2$ and $U = U(\mathfrak{sl}_2)$.

5.7.1. The Adjoint Representation of $\mathfrak{sl}_2$

With $f = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$, $h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ and $e = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, we have $\mathfrak{sl}_2 = kf \oplus kh \oplus ke$ and the Lie bracket is given by

\begin{equation}
[h, f] = -2f, \quad [h, e] = 2e \quad \text{and} \quad [e, f] = h.
\end{equation}

Thus, the adjoint action of $h$ on $\mathfrak{g} = \mathfrak{sl}_2$ is diagonalizable, with eigenvalues $-2, 0$ and $2$ and $1$-dimensional eigenspaces $\mathfrak{g}_{-2} = kf$, $\mathfrak{g}_0 = kh$ and $\mathfrak{g}_2 = ke$. The adjoint
representation \( g_{ad} \) can be visualized by the diagram

\[
\begin{array}{c}
0 \xrightarrow{f} \mathbb{K}f = \mathfrak{sl}_2 \xrightarrow{f} \mathbb{K}0 = \mathfrak{h} \xrightarrow{e} \mathbb{K}e = \mathfrak{g}_2 \xrightarrow{f} 0
\end{array}
\]  

As we have already pointed out (Example 5.14), it follows that \( g_{ad} \) is irreducible. We will see that a version of (5.49) is replicated in all finite-dimensional irreducible representations of \( sl_2 \).

### 5.7.2. The Representations \( V(m) \)

Some finite-dimensional representations of \( sl_2 \) are easy to come by. Indeed, besides the adjoint representation \( g_{ad} \) and the trivial representation \( 1 \), we certainly also have the defining representation of \( sl_2 \): it comes from the inclusion \( sl_2 \hookrightarrow gl_2 = gl(\mathbb{K}^2) \).

This representation will simply be denoted by \( \mathbb{K}^2 \). The \( sl_2 \)-action on the standard basis vectors \( x = (1, 0) \) and \( y = (0, 1) \) of \( \mathbb{K}^2 \) is given by

\[
(5.50) \quad f.x = y, \quad f.y = 0, \quad e.x = 0, \quad e.y = x \quad \text{and} \quad h.x = x, \quad h.y = y.
\]

We obtain further representations by forming the symmetric powers,

\[
V(m) := \text{Sym}^m(\mathbb{K}^2) \quad (m \geq 0).
\]

So \( V(0) = 1 \), \( V(1) = \mathbb{K}^2 \) and, in general, we know by (1.10) that \( \text{dim}_\mathbb{K} V(m) = m + 1 \).

Identifying \( \text{Sym}(\mathbb{K}^2) \) with the polynomial algebra \( \mathbb{K}[x, y] \) as usual and keeping in mind that \( sl_2 \) acts by derivations on \( \text{Sym}(\mathbb{K}^2) \), we see that \( e, f \) and \( h \) act as the differential operators \( x \frac{\partial}{\partial y}, y \frac{\partial}{\partial x} \) and \( x \frac{\partial}{\partial x} - y \frac{\partial}{\partial y} \), respectively. The representation \( V(m) \) is the space of homogeneous polynomials of total degree \( m \) in \( \mathbb{K}[x, y] \), with standard basis given by the monomials \( b_i = x^{m-i} y^i \) \( (i = 0, 1, \ldots, m) \). Thus, \( h.b_i = (x \frac{\partial}{\partial x} - y \frac{\partial}{\partial y}) x^{m-i} y^i = (m - 2i) b_i \).

We record this fact and the results of similar computations for \( e \) and \( f \) in the following formulae, with the understanding that \( b_i := 0 \) for \( i < 0 \) or \( i > m \):

\[
(5.51) \quad f.b_i = (m - i) b_{i+1}, \quad h.b_i = (m - 2i) b_i \quad \text{and} \quad e.b_i = ib_{i-1}.
\]

Comparison with (5.48) shows that \( V(2) \cong g_{ad} \) via \( b_0 \leftrightarrow e, b_1 \leftrightarrow h \) and \( b_2 \leftrightarrow f \).

Exactly as in the case of \( g_{ad} \), we see from (5.51) that \( V = V(m) \) is irreducible. In detail, each of the basis vectors \( b_i \) of \( V \) is an \( h \)-eigenvector with eigenvalue \( m - 2i \). Thus, \( V \) is completely reducible as representation of the algebra \( \mathbb{K}[h_{\mathbb{K}}] \), with homogeneous components the 1-dimensional spaces \( \mathbb{K} b_i \). It follows that every subrepresentation \( 0 \neq W \subseteq V \) must contain one of the \( b_i \) (Proposition 1.31). Now successive application of \( e_W \) or \( f_W \) using (5.51) shows that all \( b_i \) belong to \( W \), whence \( W = V \).

The following proposition summarizes the foregoing for future reference. For any \( V \in \text{Rep} sl_2 \), we will refer to the eigenvalues of the operator \( h_{\mathbb{K}} \) as the weights
of $V$ and the corresponding eigenspaces will be called **weight spaces**:

$$V_\lambda = \{ v \in V \mid h.v = \lambda v \}.$$ 

The dimension of $V_\lambda$ will be called the **multiplicity** of the weight $\lambda$ in $V$.

**Proposition 5.37.** $V = V(m) \in \text{Rep } \mathfrak{sl}_2$ is irreducible of dimension $m + 1$. The operator $h_v$ is diagonalizable with multiplicity-1 weights $-m, -m + 2, \ldots, m - 2, m$:

$$V = \bigoplus_{i=0}^{m} V_{m-2i} \quad \text{with} \quad \dim V_{m-2i} = 1.$$ 

The action of $\mathfrak{sl}_2$ on $V$ can be visualized by the following ladder diagram, with all maps $\overset{e}{\rightarrow}$ being isomorphisms (except for the zero maps at the ends):

0 \overset{f}{\rightarrow} V_{-m} \overset{e}{\rightarrow} V_{-m+2} \overset{f}{\rightarrow} V_{-m+4} \overset{e}{\rightarrow} \cdots \overset{f}{\rightarrow} V_{m-2} \overset{e}{\rightarrow} V_m \overset{f}{\rightarrow} 0

**Corollary 5.38.** If $M \subseteq \mathbb{Z}_+$ is infinite, then $\bigoplus_{m \in M} V(m)$ is a faithful representation of $U = U(\mathfrak{sl}_2)$.

**Proof.** Let $0 \neq u \in U$ be given. We need to show that $u.V(m) \neq 0$ for some $m \in M$. By the Poincaré-Birkhoff-Witt Theorem, we may write $u = \sum_{(i,j) \in S} f^i e^j p_{i,j}$ for some finite subset $S \subseteq \mathbb{Z}_+^2$ and $0 \neq p_{i,j} \in \mathbb{k}[h]$. Choose $m \in M$ so that, for all $(i,j) \in S$, we have $m \geq \max\{i, j\}$ and $p_{i,j}$ does not evaluate to 0 at $m - 2j$. Put $V = V(m)$ and let $j_0$ be the smallest $j$ with $(i, j) \in S$. Then $f^i e^j p_{i,j} V_{m-2j_0} \subseteq f^i V_{m+2(j-j_0)} = 0$ for all $(i, j) \in S$ with $j \neq j_0$. Therefore, letting $\pi_i \in \mathbb{k}^\times$ denote the value of $p_{i,j_0}$ at $m - 2j_0$, we obtain

$$u.V_{m-2j_0} = \left( \sum_{i,(j,j_0) \in S} f^i e^{j_0} p_{i,j_0} \right) V_{m-2j_0} = \left( \sum_{i,(j,j_0) \in S} \pi_i f^i \right) V_m.$$ 

Finally, $f^i V_m = V_{m-2i} \neq 0$ for all $i$ in the sum on the right, and hence $0 \neq u.V_{m-2j_0} \subseteq u.V$, as desired. \hfill $\square$

### 5.7.3. Finite-Dimensional Representations of $\mathfrak{sl}_2$

We now consider general finite-dimensional representations of $\mathfrak{sl}_2$. It turns out that any each $V \in \text{Rep}_{\text{fd}} \mathfrak{sl}_2$ can be built from the irreducible representations $V(m)$:

**Theorem 5.39.** (a) The symmetric powers $V(m) = \text{Sym}^m(\mathbb{k}^2)$ ($m \geq 0$) of the defining $\mathfrak{sl}_2$-representation are a full set of non-isomorphic finite-dimensional irreducible representations of $\mathfrak{sl}_2$. 

(b) Each \( V \in \text{Rep}_{\text{fin}} \mathfrak{sl}_2 \) is completely reducible. The endomorphism \( h_V \) is diagonalizable with integer weights, each occurring along with its negative with equal multiplicity:

\[
V = \bigoplus_{\lambda \in \mathbb{Z}} V_\lambda \quad \text{and} \quad \dim_k V_\lambda = \dim_k V_{-\lambda}.
\]

Finally, length \( V = \dim_k \ker e_V = \dim_k \ker f_V = \dim_k V_0 + \dim_k V_1 \).

The essence of Theorem 5.39 consists of the following two assertions, for any \( V \in \text{Rep}_{\text{fin}} \mathfrak{sl}_2 \):

- if \( V \) is irreducible, then \( V = V(m) \) with \( m = \dim_k V - 1 \), and
- in general, \( V \) is a (direct) sum of irreducible representations.

The assertions about diagonalizability of \( h_V \), integrality of weights and their \( \pm \)-symmetric distribution then follow from what we already know from Proposition 5.37 about the representations \( V(m) \). The formulae for the composition length are also clear from the fact that, for \( V = V(m) \), we have \( \dim_k \ker e_V = \dim_k \ker f_V = \dim_k V_0 + \dim_k V_1 = 1 \). The proof of the above two key statements will be given in 5.7.6 after some preparations in the next couple of subsections.

**Locally Finite Representations.** A representation \( V \in \text{Rep} A \), for an arbitrary \( A \in \text{Alg}_k \), is said to be **locally finite** if every \( v \in V \) is contained in some finite-dimensional subrepresentation of \( V \) or, equivalently, \( V \) is the sum of finite-dimensional subrepresentations. Complete reducibility, as will be established in Theorem 5.39 for finite-dimensional representations of \( A = \mathbb{U}(\mathfrak{sl}_2) \), generalizes directly to locally finite \( V \in \text{Rep} \mathfrak{sl}_2 \): each finite-dimensional subrepresentation of \( V \) is a sum of irreducible subrepresentations, and hence so too is \( V \).

For example, the adjoint representation \( (\mathbb{U}g)_{\text{ad}} \) of any \( g \in \text{Lie}_k \), while infinite dimensional if \( g \neq 0 \), is at least locally finite for any finite-dimensional \( g \), being the union of the finite-dimensional subrepresentations \( U_n \) that are given by the standard filtration (§5.5.6). For \( g = \mathfrak{sl}_2 \), Theorem 5.39 therefore implies that \( \mathbb{U}(\mathfrak{sl}_2)_{\text{ad}} \) is completely reducible. We shall determine the precise structure of \( \mathbb{U}(\mathfrak{sl}_2)_{\text{ad}} \) later in this section (Example 5.45).

### 5.7.4. Weight Ladders

Let \( V \in \text{Rep} \mathfrak{sl}_2 \) be arbitrary, not necessarily finite dimensional. Then we may consider the generalized weight spaces for \( \lambda \in k \) (see §5.3.3):

\[
V^\lambda = \{ v \in V \mid (h - \lambda)^t(v) = 0 \text{ for some } t \geq 0 \}.
\]

The ordinary weight space \( V_\lambda \) is evidently contained in \( V^\lambda \) and \( V^\lambda \neq 0 \) if and only if \( V_\lambda \neq 0 \). The sum of the various \( V^\lambda \) is direct, because the polynomials \( h - \lambda \in \mathbb{Z}[h] \) are relatively prime. Moreover, \( e V^\lambda \subseteq V^{\lambda+2} \); this is a consequence of the formula \( (h - \lambda - 2)e = e(h - \lambda) \) in the enveloping algebra \( U = \mathbb{U}(\mathfrak{sl}_2) \), which
in turn follows directly from the relation \( he - eh = [h, e] = 2e \). Similarly, the inclusion \( f.V^4 \subseteq V^{4-2} \) follows from \([h, f] = -2f\). Thus we have the following ladder diagram for generalized weight spaces:

\[
\cdots \xrightarrow{e} V^4 \xrightarrow{f} e \xrightarrow{f} V^{4+2} \xrightarrow{e} \cdots
\]

For the same reasons, we have an analogous ladder diagram for the ordinary weight spaces \( V_A \).

The following lemma takes care of most of the technicalities needed for the proof of Theorem 5.39. Recall that, for \( \lambda \in \mathbb{Z}_+ \), we have the irreducible representation \( V(\lambda) \) of degree \( \lambda + 1 \) as in Proposition 5.37.

**Lemma 5.40.** Let \( V \in \text{Rep}\mathfrak{sl}_2 \) and let \( 0 \neq v \in V^4 \) be such that \( e.v = 0 \) and \( f^n.v = 0 \) for some \( n \in \mathbb{Z}_+ \). Then \( \lambda \in \mathbb{Z}_+ \) and \( v \in V_\lambda \). Moreover, \( \mathcal{U}(\mathfrak{sl}_2).v \equiv V(\lambda) \).

**Proof.** We work in the enveloping algebra \( \mathcal{U} = \mathcal{U}(\mathfrak{sl}_2) \) and begin by trotting out two formulae whose verification is left to the reader (Exercise 5.7.1): for any \( v \in V \) with \( e.v = 0 \) and any \( k \in \mathbb{Z}_+ \),

\[
e^k f^k.v = k f^{k-1} (h - k + 1).v \quad \text{and} \quad e^k f^k.v = k! \prod_{i=0}^{k-1} (h - i).v.
\]

Turning now to the specific situation in the lemma, let \( 0 \neq v \in V^4 \) with \( e.v = 0 \) and put \( N := \min\{n \in \mathbb{Z}_+ \mid f^n.v = 0\} \). Then \( 0 = e^N f^N.v \) and so (5.53) implies that \( \prod_{i=0}^{N-1} (h - i).v = 0 \). On the other hand, \( (h - \lambda)^t.v = 0 \) for some \( t \geq 1 \). It follows that \( \lambda \in \{0, 1, \ldots, N - 1\} \), because otherwise \( \prod_{i=0}^{N-1} (h - i) \) and \( (h - \lambda)^t \) would be relatively prime and could not both annihilate a nonzero vector. For the same reason, we must have \( (h - \lambda).v = 0 \), because this vector is annihilated by \( \prod_{i=0}^{N-1} (h - i) \) and \( (h - \lambda)^{t-1} \). We have thus shown that \( \lambda \in \mathbb{Z}_+ \) and \( v \in V_\lambda \).

It remains to show that \( U.v \equiv V(\lambda) \). First, we determine \( \lambda \in \{0, 1, \ldots, N - 1\} \):

**Claim.** \( \lambda = N - 1 \).

Indeed, \( e f^{k+1}.v = (\lambda + 1) f^{k} (h - \lambda).v = 0 \) by (5.53). Thus, we may apply (5.53) again, with \( v' := f^{k+1}.v \) playing the role of \( v \). Taking \( k = N - \lambda - 1 \geq 0 \) using the fact that \( v' \in V_{\lambda-2} \) by (5.52), we obtain

\[
0 = e^k f^N.v = e^k f^k.v' = k! \prod_{i=0}^{k-1} (h - i).v' = k! \prod_{i=0}^{k-1} (-\lambda - 2 - i)v'.
\]

Since all scalar factors in the last expression are nonzero, we must have \( v' = 0 \). Therefore, \( \lambda + 1 \geq N \) by minimality of \( N \), proving the claim.

Now put \( v_i := f^i.v \) (\( i = 0, 1, \ldots, \lambda \)) for brevity; so \( f.v_\lambda = 0 \) while \( 0 \neq v_i \in V_{\lambda-2i} \). In particular, the \( v_i \) are linearly independent. Note also that \( \bigoplus_{i=0}^{\lambda} k v_i \) is
5.7. Representations of \( \mathfrak{sl}_2 \)

stable under the action of \( \mathfrak{sl}_2 \):

\[
h.v_i = (\lambda - 2i)v_i, \quad f.v_i = v_{i+1} \quad \text{and} \quad e.v_i = i(\lambda - i + 1)v_{i-1}.
\]

It follows that \( U.v = \bigoplus_{i=0}^{\lambda} k v_i \). Moreover, comparison of the above action formulae with (5.51) shows that the map \( b_i \mapsto \frac{1}{\lambda - i + 1} v_i \), where \( (\lambda)_i = \lambda(\lambda - 1) \cdots (\lambda - i + 1) \), gives an isomorphism \( V(\lambda) \cong U.v \). The lemma is thus proved. \( \square \)

We are primarily interested in the case where \( V \) is finite dimensional. Once Theorem 5.39 is proved, we will know that \( h_v \) is diagonalizable or, equivalently, \( V_\lambda = V^A \) for all \( \lambda \), but this is not yet clear at this point.

5.7.5. The Casimir Element

While the analysis of representations of \( \mathfrak{sl}_2 \) heretofore was focused on the element \( h \in \mathfrak{sl}_2 \), the proof of Theorem 5.39 will also make use of a certain element in the center of the enveloping algebra \( U = U(\mathfrak{sl}_2) \), the so-called Casimir element. Before discussing this particular element, let us make the following general observation.

**Lemma 5.41.** Let \( A \in \text{Alg}_K \) be arbitrary and let \( V \in \text{Rep}_{\text{fin}} A \) be indecomposable. Then, for any \( c \in \mathcal{Z} A \), the operator \( c_\psi \) has only one eigenvalue.

**Proof.** Consider the decomposition (5.11) of \( V \) as the direct sum of the generalized eigenspaces \( V^A \) for the operator \( \phi = c_\psi \in \text{End}_K(V) \). Since generalized eigenspaces of \( \phi \) are evidently stable under any \( \psi \in \text{End}_K(V) \) that commutes with \( \phi \), all \( V^A \) are in fact \( A \)-subrepresentations of \( V \). Therefore, our hypothesis that \( V \) is indecomposable implies that only one \( V^A \) can be nonzero. \( \square \)

Now let us turn specifically to the Casimir element of \( U = U(\mathfrak{sl}_2) \); this element is defined by

\[
(5.54) \quad c := 2ef + 2fe + h^2 = 4fe + h(h + 2) = 4ef + h(h - 2)
\]

It is a simple matter to check that \([h, c] = [e, c] = [f, c] = 0\) holds in \( U \); for example, \([h, e] = 4h, f e] = 4[h, f] = e + 4f[h, e] = -8fe + 8fe = 0\). So \( c \) belongs to the center of \( U \). In fact, we will see below that \( c \) generates \( \mathcal{Z} U \) (Proposition 5.47), but this will not be needed for the proof of Theorem 5.39. We will however need to know the action of \( c \) on the irreducible representations \( V(m) \).

**Lemma 5.42.** The Casimir element \( c \) acts as the scalar \( m(m + 2) \) on \( V(m) \).

**Proof.** Since \( e.V(m)_m = 0 \), the second expression in (5.54) shows that \( c \) acts as the scalar \( m(m + 2) \) on \( V(m)_m \). Since \( c \in \mathcal{Z} U \) and \( V(m) = U.V(m)_m \), it follows that \( c \) acts as \( m(m + 2) \) on all of \( V(m) \). \( \square \)
5.7.6. Proof of Theorem 5.39

**Proof of part (a).** Let \( 0 \neq V \in \text{Rep}_{\text{fin}} \mathfrak{sl}_2 \). Then (5.11) gives the decomposition

(5.55) \[ V = \bigoplus_{\lambda \in \mathbb{C}} V^\lambda \]

and the ladder diagram (5.52) allows us to choose \( \lambda \) and \( 0 \neq \mathbf{v} \in V^\lambda \) as in Lemma 5.40. Therefore, if \( V \) is irreducible, then \( V = U.\mathbf{v} \cong V(\lambda) \). \( \square \)

**Proof of part (b).** In light of the remarks in §5.7.3, it remains to show that an arbitrary \( V \in \text{Rep}_{\text{fin}} \mathfrak{sl}_2 \) is completely reducible. For this, we may clearly assume that \( V \) is indecomposable. It follows that the endomorphism \( c_V \) for the Casimir element \( c \in \mathcal{Z} U \) has only one eigenvalue (Lemma 5.41).

First, let us determine the eigenvalue of \( c_V \) and the weights of \( V \). For this, consider a composition series,

\[ 0 = W_0 \subset W_1 \subset W_2 \subset \cdots \subset W_l = V. \]

By part (a), we know that \( \overline{W}_i := W_i/W_{i-1} \cong V(m_i) \) for suitable \( m_i \in \mathbb{Z}_+ \) and Lemma 5.42 further tells us that \( c \) acts the scalar \( m_i(m_i + 2) \) on \( \overline{W}_i \). Thus, building a \( \mathbb{C} \)-basis of \( V \) successively from bases for \( W_1, W_2 \) and so on, the matrix of \( c_V \) is upper triangular with \( l \) blocks consisting of the scalar matrices \( m_i(m_i + 2) I_{(m_i+1) \times (m_i+1)} \) along the diagonal. Since the diagonal entries \( m_i(m_i + 2) \) must all be equal to the unique eigenvalue of \( c_V \), it follows that \( m_1 = m_2 = \cdots = m \). Using the images of the standard basis \( (b_i) \) of \( V(m) \) as in (5.51) as basis for each composition factor \( \overline{W}_i \), the matrix of \( h_V \) is also upper triangular with \( l \) blocks along the diagonal, all of which are equal to the diagonal matrix \( \text{diag}(m, m - 2, \ldots, -m) \). Thus, the scalars \( \lambda \in \{m, m - 2, \ldots, -m\} \) are the weights of \( V \) and the multiplicity of each \( \lambda \) as a root of the characteristic polynomial of \( h_V \) is equal to \( l \); so \( \text{dim}_\mathbb{C} V^\lambda = l \) for all \( \lambda \).

To finish the proof note that \( V^m \neq 0 \) but \( e.V^m \subseteq V^{m+2} = 0 \) by (5.52). Thus, Lemma 5.40 tells us that the generalized weight space \( V^m \) coincides with the ordinary weight space \( V_m \). Furthermore, we know that \( U.\mathbf{v} \cong V(m) \) holds for each \( 0 \neq \mathbf{v} \in V_m \). Therefore, the subrepresentation \( V' := U.V_m \) of \( V \) is the sum of \( l = \text{dim}_\mathbb{C} V_m \) many subrepresentations, each isomorphic to \( V(m) \); so \( V' \cong V(m)^{\otimes d} \) with \( d \leq l \). In fact, we must have \( d = l \), because the isomorphism \( V' \overset{\cong}{\to} V(m)^{\otimes d} \) embeds \( V_m \subseteq V' \) into \( V(m)^{\otimes d} \), and so \( l = \text{dim}_\mathbb{C} V_m \leq \text{dim}_\mathbb{C} V(m)^{\otimes d} = d \). Thus \( V' \cong V(m)^{\otimes d} \). It follows that \( \text{dim}_\mathbb{C} V'_\lambda = l = \text{dim}_\mathbb{C} V^\lambda \) for all \( \lambda \in \{m, m - 2, \ldots, -m\} \). Since \( V'_\lambda \subseteq V^\lambda \subseteq V^\lambda \), we conclude that \( V'_\lambda = V^\lambda \) for all \( \lambda \). Hence (5.55) gives \( V = V' \), proving that \( V \cong V(m)^{\otimes d} \) is completely reducible and finishing the proof of Theorem 5.39. \( \square \)
5.7. Representations of $\mathfrak{sl}_2$

5.7.7. Formal Characters

For any $V$ in $\text{Rep}_{\text{fin}} \mathfrak{sl}_2$, we know that $V_\lambda = \{ v \in V \mid h.v = \lambda v \}$ is nonzero only for certain (at most finitely many) integer values of $\lambda$, the weights of $V$ (Theorem 5.39). Therefore, we may define the following Laurent polynomial, which is called the formal character of $V$:

$$ \text{ch} V \overset{\text{def}}{=} \sum_{\lambda \in \mathbb{Z}} (\dim_k V_\lambda) t^\lambda \in \mathbb{Z}[t^{\pm 1}] $$

Example 5.43. For the irreducible representation $V(m)$ of degree $m + 1$, we have $\dim_k V_\lambda = 1$ for $\lambda = -m, -m + 2, \ldots, m - 2, m$ and $\dim_k V_\lambda = 0$ for all other $\lambda$ (Proposition 5.37). Thus,

$$ \text{ch} V(m) = \sum_{i=0}^{m} t^{m-2i} = \frac{t^{m+1} - t^{-m-1}}{t - t^{-1}}. $$

Lemma 5.44. Let $V, V' \in \text{Rep}_{\text{fin}} \mathfrak{sl}_2$. Then:

(a) $\text{ch}(V \oplus V') = \text{ch} V + \text{ch} V'$ and $\text{ch}(V \otimes V') = \text{ch} V \cdot \text{ch} V'$.

(b) $V \cong V'$ if and only if $\text{ch} V = \text{ch} V'$.

Proof. (a) For any $\lambda$, we have

$$ (V \oplus V')_\lambda = V_\lambda \oplus V'_\lambda \quad \text{and} \quad (V \otimes V')_\lambda = \bigoplus_{\mu + \mu' = \lambda} (V_\mu \otimes V'_\mu). $$

Both equalities follow from the fact that $V$ and $V'$ are the direct sums of their various weight spaces (Theorem 5.39); the second formula also uses the tensor product action (5.29), $h_{V \otimes V'} = h_V \otimes \text{Id}_{V'} + \text{Id}_V \otimes h_{V'}$, which implies $V_\mu \otimes V'_\mu \subseteq (V \otimes V')_{\mu + \mu'}$. Taking dimensions gives the expressions for formal characters in (a).

(b) Clearly, the formal character $\text{ch}(V)$ depends only on the isomorphism type of $V$. For the converse, we invoke Theorem 5.39 to write $V \cong \bigoplus_{m \in \mathbb{Z}_+} V(m)^{\oplus l(m)}$ for suitable $l(m) \in \mathbb{Z}_+$. In view of (a), this gives $\text{ch}(V) = \sum_{m} l(m) \text{ch}(V(m))$. Since the formal characters $\text{ch} V(m)$ are linearly independent in $\mathbb{Z}[t^{\pm 1}]$ by Example 5.43, it follows that the coefficients $l(m)$ are determined by $\text{ch}(V)$. Hence $\text{ch}(V)$ determines the isomorphism type of $V$. □

Formal characters are useful tools in finding the decomposition into irreducible constituents of representations that are derived from other representations by forming tensor products, symmetric powers etc. Here is one example; see the exercises for others.
Example 5.45 (The adjoint representation \(U_{\text{ad}}\)). By Proposition 5.27, we know that \(U_{\text{ad}} \cong \bigoplus_{n \geq 0} \text{Sym}^n g_{\text{ad}}\). Hence, it suffices to determine the structure of \(\text{Sym}^n g_{\text{ad}}\). With the understanding that \(V(m) = 0\) for \(m < 0\), we claim that

\[
\text{Sym}^n g_{\text{ad}} \cong \bigoplus_{i \geq 0} V(2n - 4i).
\]

In view of Corollary 5.38, this will in particular show that \(U_{\text{ad}}\) is a faithful representation of \(U\). To prove (5.56), we compare formal characters. Put \(V := \bigoplus_{i \geq 0} V(2n - 4i)\) and recall that the multiplicity of the weight \(\lambda \in \mathbb{Z}\) in \(V(2n - 4i)\) equals 1 if \(\lambda\) is even and \(|\lambda| \leq 2n - 4i\); otherwise the multiplicity is 0. Thus,

\[
\dim_k V_{\lambda} = \begin{cases} 
\#\{i \in \mathbb{Z} \mid 0 \leq i \leq \frac{2n - |\lambda|}{4}\} & \text{for } \lambda \text{ even}, \\
0 & \text{for } \lambda \text{ odd}.
\end{cases}
\]

For \(\text{Sym}^n g_{\text{ad}}\), we may use the \(k\)-basis consisting of the monomials \(e^i f^j h^k\) with \(i, j, k \in \mathbb{Z}_+\), \(i + j + k = n\). Since \(e^i f^j h^k\) is a weight vector of weight \(2(i - j)\), we obtain, for each given \(\lambda \in \mathbb{Z}\),

\[
\dim_k (\text{Sym}^n g_{\text{ad}})_{\lambda} = \#\{(i, j, k) \in \mathbb{Z}_+^3 \mid i + j + k = n, 2(i - j) = \lambda\}
\]

\[
= \#\{(i, j) \in \mathbb{Z}_+^2 \mid i + j \leq n, 2(i - j) = \lambda\}
\]

\[
= \begin{cases} 
\#\{i \in \mathbb{Z} \mid 0 \leq i \leq \frac{2n - |\lambda|}{4}\} & \text{for } \lambda \text{ even}, \\
0 & \text{for } \lambda \text{ odd}.
\end{cases}
\]

Since the multiplicities agree, the isomorphism (5.56) follows.

5.7.8. The Representation Ring of \(sl_2\)

Formal characters also provide a means for getting our hands on the structure of the representation ring \(\mathcal{R}(sl_2)\) as introduced in §5.5.8 for arbitrary Lie algebras. Specifically, note that formal characters yield a ring homomorphism

\[
\text{ch}: \mathcal{R}(sl_2) \rightarrow \mathbb{Z}[t^{\pm 1}]
\]

\[
[V] \longmapsto \text{ch} V
\]

Indeed, if \(0 \rightarrow U \rightarrow V \rightarrow W \rightarrow 0\) is a short exact sequence in \(\text{Rep}_{\text{fin}} sl_2\), then \(V \cong U \oplus W\) by complete reducibility, and so and the direct sum formula in Lemma 5.44 implies that \(\text{ch} V = \text{ch} U + \text{ch} V\). Thus, \(\text{ch}\) is a well-defined group homomorphism. Multiplicativity is a consequence of the tensor product formula in Lemma 5.44. In order to describe the image of \(\text{ch}\) in \(\mathbb{Z}[t^{\pm 1}]\), we let the symmetric group \(S_2\) act on \(\mathbb{Z}[t^{\pm 1}]\) by ring automorphisms:

\[
S_2 = \langle s \rangle \subset \mathbb{Z}[t^{\pm 1}] \quad \text{with} \quad s.t = t^{-1}.
\]
The assertion in Theorem 5.39 that each weight of any $V \in \text{Rep}_{\text{fin}} \mathfrak{sl}_2$ occurs along with its negative, with equal multiplicity, states exactly that $\text{ch} V$ is contained in the invariant subring for this action.\(^9\)

\[
\mathbb{Z}[t^{\pm 1}]^{S_2} \overset{\text{def}}{=} \left\{ f \in \mathbb{Z}[t^{\pm 1}] \mid s.f = f \right\}.
\]

**Proposition 5.46.** The formal character map gives a ring isomorphism

\[
\text{ch}: \mathcal{R}(\mathfrak{sl}_2) \xrightarrow{\sim} \mathbb{Z}[t^{\pm 1}]^{S_2} = \mathbb{Z}[t + t^{-1}].
\]

In particular, the class of the defining representation $\text{Rep}(1) = \mathbb{Z}^2$ freely generates the representation ring $\mathcal{R}(\mathfrak{sl}_2)$.

**Proof.** First recall from Proposition 1.46 that, for any $A \in \text{Alg}_{\mathbb{Z}}$, the classes $[S]$ with $S \in \text{Ir}_{\text{fin}} A$ form a $\mathbb{Z}$-basis of the group $\mathcal{R}(A)$. In the present context, these are the classes $[V(m)]$ with $m \in \mathbb{Z}_+$ (Theorem 5.39). Inasmuch as the formal characters $\text{ch} V(m)$ are $\mathbb{Z}$-linearly independent in $\mathbb{Z}[t^{\pm 1}]$, as we have noted earlier (Example 5.43), the formal character map $\text{ch}$ is a monomorphism. Thus, $\mathcal{R}(\mathfrak{sl}_2)$ is isomorphic to $\text{Im}(\text{ch})$, and we have already observed that $\text{Im}(\text{ch})$ is contained in the ring of invariants $\mathbb{Z}[t^{\pm 1}]^{S_2}$. Since $t + t^{-1} = \text{ch} V(1) \in \text{Im}(\text{ch})$, it suffices to show that $\mathbb{Z}[t^{\pm 1}]^{S_2} \subseteq \mathbb{Z}[t + t^{-1}]$. For this, note that

\[
\mathbb{Z}[t^{\pm 1}] = \mathbb{Z}[t + t^{-1}] + t \mathbb{Z}[t + t^{-1}].
\]

Indeed, the right hand side is a $\mathbb{Z}$-submodule of $\mathbb{Z}[t^{\pm 1}]$ containing $t$ and $t^{-1}$ and it is easily seen to be closed under multiplication. Thus, any $c \in \mathbb{Z}[t^{\pm 1}]$ has the form $c = a + tb$ with $a, b \in \mathbb{Z}[t + t^{-1}]$. Since $a$ and $b$ are both $S_2$-invariant, the condition $c = s.c$ is equivalent to $(t - t^{-1})b = 0$, or $b = 0$. This shows that $\mathbb{Z}[t^{\pm 1}]^{S_2} \subseteq \mathbb{Z}[t + t^{-1}]$, which completes the proof. \(\square\)

**5.7.9. The Center of $U(\mathfrak{sl}_2)$**

The remainder of this chapter is concerned with ring theoretic properties of the enveloping algebra of $\mathfrak{sl}_2$. We will write $U = U(\mathfrak{sl}_2)$ throughout. Our first goal is to show that the center $\mathcal{Z} U$ is generated by the Casimir element $c$ from (5.54), as was announced earlier.

**Proposition 5.47.**

(a) $\mathcal{Z} U = \mathbb{Z}[c]$.

(b) If $I$ is a nonzero ideal of $U$, then $I \cap \mathcal{Z} U \neq 0$.

**Proof.** (a) Recall from (5.38) that $\mathcal{Z} U = U^3$, the invariants of the adjoint representation $U_{\text{ad}}$ of $\mathfrak{g} = \mathfrak{sl}_2$. By Proposition 5.27 and Example 5.45, we also know that $U_{\text{ad}} \cong \bigoplus_{n \geq 0} (U_n/U_{n-1})_{\text{ad}}$ and $(U_n/U_{n-1})_{\text{ad}} \cong \text{Sym}^n g_{\text{ad}} \cong \bigoplus_{i \geq 0} V(2n - 4i)$ in

---

\(^9\)The invariant ring $\mathbb{Z}[t^{\pm 1}]^{S_2}$ is an example of a *multiplicative invariant algebra* (over $\mathbb{Z}$). We will revisit this type of invariant algebra in §7.4.4.
Rep \mathfrak{g}. Thus, the trivial representation \mathbb{1} = V(0) occurs exactly once in \text{Sym}^n \mathfrak{g}_{\text{ad}} if 
 n is even and not at all if \( n \) is odd. Therefore,

\[ \mathcal{Z} U \cong \bigoplus_{k \geq 0} (U_{2k}/U_{2k-1})^0 \quad \text{and} \quad \dim_k (U_{2k}/U_{2k-1})^0 = 1. \]

Since \( c^k \) gives a nonzero invariant in \( U_{2k}/U_{2k-1} \), we must have \( k c^k \cong (U_{2k}/U_{2k-1})^0 \)
under the above isomorphism. Therefore, \( \mathcal{Z} U = k[c] \).

(b) By Corollary 5.38 we may choose \( V = V(m) \) so that \( I.V \neq 0 \). Since the homomorphism \( \rho: U \to \text{End}_k(V), u \mapsto u \cdot V \), is surjective by Burnside’s Theorem (§1.4.6) and the algebra \( \text{End}_k(V) \) is simple, it follows that \( \rho(I) = \text{End}_k(V) \). Moreover, the map \( \rho \) is a homomorphism of \( \mathfrak{g} \)-representations when \( U \) is viewed as \( U_{\text{ad}} \)
(Exercise 5.5.1). Since \( U_{\text{ad}} \) is completely reducible, the subrepresentation \( I \subseteq U_{\text{ad}} \)
is completely reducible as well. Hence, the epimorphism \( I \to \text{End}_k(V) \) splits (Theorem 1.28). The identity map \( \text{Id}_V \in \text{End}_k(V)^0 = D(V) \) therefore corresponds to a nonzero \( \mathfrak{g} \)-invariant in \( I \), which is the desired nonzero element of \( I \cap \mathcal{Z} U \). \( \square \)

5.7.10. Prime and Primitive Ideals of \( U(\mathfrak{sl}_2) \)

We shall now endeavor to describe the prime and the primitive ideals of the enveloping algebra \( U = U(\mathfrak{sl}_2) \).\(^{10}\) The theorem below was originally stated, without proof, in Nouazé and Gabriel [159]. As before, \( c \in \mathcal{Z} U \) will denote the Casimir element. An ideal \( I \) of an arbitrary \( A \in \text{Alg}_k \) is said to be \textbf{completely prime} if the factor \( A/I \) is a domain.

**Theorem 5.48.**

(a) The zero ideal of \( U \) is completely prime but not primitive.

(b) For each \( \lambda \in \mathbb{Z} \), the ideal \( P(\lambda) = (c - \lambda) U \) of \( U \) is completely prime and primitive. The ideals \( P(\lambda) \) are the minimal nonzero primes of \( U \).

(c) If \( \lambda \notin \{ m^2 + 2m \mid m \in \mathbb{Z}_+ \} \), then \( P(\lambda) \) is also a maximal ideal of \( U \). For \( \lambda = m^2 + m \), the kernel of the irreducible representation \( V(m) \) is the only proper ideal of \( U \) strictly containing \( P(\lambda) \).

---

\(^{10}\)Arbitrary ideals of \( U \) are described in Kirillov [123, 18.3], Bavula [11], [12] and Catoiu [38].
Proof of Theorem 5.48. (a) Since $U$ is a domain, the zero ideal is certainly completely prime. However, since $\mathcal{Z}U$ is a polynomial algebra (Proposition 5.47), the zero ideal cannot be primitive by the Nullstellensatz for enveloping algebras; see (5.45).

(b) Now let $P$ be a nonzero prime ideal of $U$. Then $P \cap \mathcal{Z}U$ is a nonzero prime ideal of $\mathcal{Z}U = \mathbb{k}[c]$ by Proposition 5.47. Thus, $P \cap \mathcal{Z}U = (c - \lambda)$ for some $\lambda \in \mathbb{k}$, and so $P \supseteq P(\lambda)$. Next, we show that each $P(\lambda)$ is completely prime, that is, the algebra

$$B = B(\lambda) := U/P(\lambda)$$

has no zero divisors. Let $\overline{-} : U \rightarrow B$ denote the canonical map and filter $B$ by the subspaces $B_n = \overline{U_n}$, where $\{U_n\}$ is the standard filtration of $U$. It is not hard to check that

$$\text{gr } B \equiv \text{gr } U/(c - \lambda)^\circ \text{ gr } U,$$

where $(c - \lambda)^\circ \in \text{gr } U$ is the symbol of $c - \lambda$ (Exercise 5.4.2). By Corollary 5.25(a), $\text{gr } U$ is the polynomial algebra $\mathbb{k}[\hat{e}, \hat{f}, \hat{h}]$ generated by the symbols of $e, f$ and $h$. Since $(c - \lambda)^\circ = 4\hat{e}\hat{f} - \hat{h}^2$ is an irreducible polynomial in $\mathbb{k}[\hat{e}, \hat{f}, \hat{h}]$, we conclude that $\text{gr } B$ is a domain, and hence so is $B$ (Lemma 5.23). This proves that the ideal $P(\lambda)$ is completely prime. Finally, in view of (5.43), primitivity of $P(\lambda)$ will be a consequence of part (c), which shows that $P(\lambda)$ is maximal or at least locally closed, depending on the value of $\lambda$.

(c) In order to describe the ideals of the algebra $B$, we first analyze the adjoint representation $B_{\text{ad}}$ (§5.5.6). The filtration $B_n$ of $B$ from the proof of (b) is stable under the adjoint action of $\mathfrak{sl}_2$ and all $B_n$ are completely reducible (Theorem 5.39). Thus, $B_{\text{ad}} \equiv \bigoplus_{n \geq 0} B_n/B_{n-1}$ in $\text{Rep } \mathfrak{sl}_2$. Put $x = \overline{\tau}, y = \overline{f}, z = \overline{h} \in B$.

Claim. The monomials $x^i z^j$ and $y^i z^j$ with $i + j \leq n$ form a $\mathbb{k}$-basis of $B_n$.

Proof of the Claim. The monomials $x^i y^j z^k$ with $i + j + k \leq n$ span $B_n$ by the Poincaré-Birkhoff-Witt Theorem. Furthermore, the equations $c = 4ef + h(h - 2)$ and $hf = f(h - 2)$ in $U$ give $4x y = \lambda - z(z - 2)$ and $zy = y(z - 2)$ in $B$. It readily follows that $B_n$ is spanned by the monomials $x^i z^j$ and $y^i z^j$ with $i + j \leq n$. Thus, $B_n = \sum_{i=-n}^{n} B_{n,i}$ with $B_{n,i} = \left\{ \begin{array}{ll} \sum_{j=0}^{n-i} \mathbb{k} x^j z^j & \text{for } i \geq 0, \\ \sum_{j=0}^{n+i} \mathbb{k} y^{-i} z^j & \text{for } i < 0. \end{array} \right.$

The above sum is in fact direct, because each $B_{n,i}$ consist of eigenvectors with eigenvalue $2i$ for the operator $\text{ad}_B h$. Moreover, any linear relation among $x^i, x^j z, x^j z^2, \ldots$ amounts to an equation $x^i p(z) = 0$ for some nonzero polynomial $p$, and hence $x^i (z - \mu_1)(z - \mu_2) \cdots (z - \mu_d) = 0$ for suitable $\mu_j \in \mathbb{k}$. Since it is easy to see that $x$ and all $z - \mu_j$ are nonzero elements of $B$ (use Exercise 5.4.2(a)), we obtain a contradiction to the fact that $B$ is a domain as was shown in (b). The same argument
also proves linear independence of the elements of the form $y^i z^j$ with fixed $i$. The claim follows.

Thus, the residue classes of $y^n, y^{n-1}z, \ldots, z^n, z^{n-1}, \ldots, z^n$ form a $\k$-basis of $B_n/\mathcal{B}_{n-1}$. Since these monomials are $\text{ad}_B h$-eigenvectors with respective weights $-2n, -2(n-1), \ldots, 2n$, it follows that $(B_n/\mathcal{B}_{n-1})_{\text{ad}} \cong \k V(2n)$ (Theorem 5.39). Therefore, $B_{\text{ad}} \cong \bigoplus_{n \geq 0} \k V(2n)$. Writing the $V(2n)$-homogeneous component of $B_{\text{ad}}$ simply as $V(2n)$, we have $B_n = V(2n) \oplus B_{n-1}$ and $x^n \in V(2n)$, because $x^n \in B_n$ has weight $2n$.

Now let $I$ be a nonzero proper ideal of $B$. Then $I$ is a subrepresentation of $B_{\text{ad}}$ and so $I = \bigoplus_{n \geq 0} I \cap \k V(2n)$ (Proposition 1.31). Let $n_0$ be minimal with $I \cap \k V(2n_0) \neq 0$ and note that $n_0 \geq 1$, because $V(0) = \k$. Then $\k V(2n_0) \subseteq I$ and in particular $x^{n_0} \in I$. It follows that $x^{n_0} \in I$ and so $\k V(2n) \subseteq I$ for all $n \geq n_0$. Therefore, $I$ is determined by $n_0$: $I = \bigoplus_{n \geq n_0} \k V(2n)$. The next claim shows that $n_0$ in turn is determined by $\lambda$; so $I$ is in fact unique.

**Claim.** $\lambda = n_0^2 - 1$.

The claim also shows that that $\lambda = m^2 + 2m$ with $m = n_0 - 1 \in \mathbb{Z}_+$. Finally, since $P(m^2 + 2m)$ is (properly) contained in the kernel of $V(m)$ by Lemma 5.42, it will follow that the unique nontrivial ideal of the algebra $B = B(m^2 + 2m)$ must be the image of this kernel, proving all assertions of (c).

It remains to justify the claim. Exercise 5.7.1 gives the relation $[x^{n_0}, y] = n_0x^{n_0}z(n_0 - 1)$. Since $x^{n_0} \equiv 0 \mod I$, we obtain $x^{n_0-1}z \equiv x^{n_0-1}(1 - n_0) \mod I$. Using this and the fact that $4xy + z(z - 2) \equiv \lambda$, we compute modulo $I$:

\[
x^{n_0-1}\lambda \equiv x^{n_0-1}(4xy + z(z - 2)) \\
\equiv 0 + x^{n_0-1}z^2 - 2x^{n_0-1}z \\
\equiv x^{n_0-1}(1 - n_0)^2 - 2x^{n_0-1}(1 - n_0) \\
\equiv x^{n_0-1}(n_0^2 - 1).
\]

Since $x^{n_0-1} \not\in I$, this proves the claim and finishes the proof of the theorem. \(\square\)

The analysis of the prime and primitive spectra of $U = U(sl_2)$ in the proof of Theorem 5.48 relied strongly on the maps $\text{Spec } U \rightarrow \text{Spec } \mathbb{k}[c]$ and $\text{Prim } U \rightarrow \text{MaxSpec } \mathbb{k}[c]$ that are given by $P \mapsto P \cap \mathbb{k}[c]$; these maps were already discussed more generally in (5.44) and (5.45). In the present setting, both maps are surjective, and they are very nearly injective as well. Indeed, only the fibres over the ideals $(c - m^2 - 2m) \in \text{MaxSpec } \mathbb{k}[c]$ with $m \in \mathbb{Z}_+$ consist of two points: $P(m^2 + 2m)$ and $\text{Ker } V(m)$. Figures 5.3 and 5.4 depict $\text{Spec } U$, with red indicating primitivity as before. In Figure 5.4, red points fading into white are primitive ideals that are not maximal, that is, they are not closed in $\text{Spec } U$. 

5. Lie Algebras and Enveloping Algebras
5.7. Representations of $\mathfrak{sl}_2$

\[ \text{Spec } U: \]
\[ \text{Spec } k[c]: \]

Figure 5.4. Spec $U(\mathfrak{sl}_2)$ and the map to Spec $k[c]$

Exercises for Section 5.7

In these exercises, the field $\mathbb{k}$ is assumed to be algebraically closed and, except when mentioned otherwise, of characteristic 0. The elements $f, h, e \in \mathfrak{sl}_2 = \mathfrak{sl}_2(\mathbb{k})$ have their usual meaning and we continue to write $U = \mathcal{U}(\mathfrak{sl}_2)$.

5.7.1 (Some computations). In this problem, $\text{char } \mathbb{k}$ can be arbitrary.

(a) Verify that the following commutator relations hold in $U$ for all $i \in \mathbb{Z}_+$:

\[
[f^i, e] = i f^{i-1}(i - h - 1), \quad [e^i, f] = i e^{i-1}(h + i - 1),
\]
\[
[h, e] = e((h + 2)^i - h^i), \quad [h, f] = f((h - 2)^i - h^i).
\]

(b) For $p(h) \in \mathbb{k}[h]$, show that $[p(h), e] = e\bar{p}(h)$ with $\bar{p}(h) = p(h + 2) - p(h) \in \mathbb{k}[h]$. Conclude that $\mathbb{e}_k[h] = \mathbb{k}[h]e$. Similarly for $f$.

(c) Let $V \in \text{Rep } \mathfrak{sl}_2$ and let $v \in V$ be such that $e.v = 0$. Show that $ef^k.v = k f^{k-1}(h - k + 1).v$ and $e^k f^k.v = k! \prod_{i=0}^{k-1}(h - i).v$

5.7.2 ($\mathfrak{sl}_2$-invariants). Let $V$ be a locally finite representation of $\mathfrak{sl}_2$. Show that $b = \mathbb{k}h \oplus \mathbb{k}e$ is a Lie subalgebra of $\mathfrak{sl}_2$ and that $V^{\mathfrak{sl}_2} = V^b$.

5.7.3 (Infinite-dimensional irreducible $\mathfrak{sl}_2$-modules). Put $V = \bigoplus_{i \in \mathbb{Z}_+} \mathbb{k}v_i$ and fix $\lambda \in \mathbb{k}$. Define

\[
h.v_i = (\lambda - 2i)v_i, \quad f.v_i = (i + 1)v_{i+1} \quad \text{and} \quad e.v_i = (\lambda - i + 1)v_{i-1}
\]

(a) Show that these formulae make $V$ a representation of $\mathfrak{sl}_2$.

(b) Put $b = \mathbb{k}h \oplus \mathbb{k}e$ as in Exercise 5.7.2 and let $\mathbb{k}_1 = \mathbb{k}$ with $b$-action given by $h.1 = \lambda$ and $e.1 = 0$. Show that $\mathbb{k}_1 \in \text{Rep } b$ and that $V \cong \text{Ind}^{U}_{b} \mathbb{k}_1$ in $\text{Rep } \mathfrak{sl}_2$.

(c) Prove that $V$ is irreducible for $\lambda \notin \mathbb{Z}_+$.

5.7.4 (Clebsch-Gordan formula). For $m \geq n$, prove the following isomorphism in $\text{Rep } \mathfrak{sl}_2$: $V(m) \otimes V(n) \cong V(m + n) \oplus V(m + n - 2) \oplus \cdots \oplus V(m - n)$.

5.7.5 (Hermite reciprocity). Prove that $\text{Sym}^n V(m) \cong \text{Sym}^m V(n)$ in $\text{Rep } \mathfrak{sl}_2$. 
5.7.6 (The Hessian). (a) Deduce the isomorphism \( \text{Sym}^2 V(3) \cong V(6) \oplus V(2) \) from equation (5.56) and Exercise 5.7.5. Note that the images of \( V(6) \) and \( V(2) \) in \( \text{Sym}^2 V(3) \) are uniquely determined, being the homogeneous components of \( \text{Sym}^2 V(3) \).

(b) View \( V(2) \) and \( V(3) \) as the spaces of all polynomials of degree 2 and 3 in the commuting variables \( x \) and \( y \) and let \( f : V(3) \rightarrow \text{Sym}^2 V(3) \rightarrow V(2) \) be the map that is given by squaring followed by projection along \( V(6) \) in (a)—the last map is determined up to a nonzero scalar multiple. Show that, up to a scalar factor, \( f(p) \) is the **Hessian** of the polynomial \( p \):

\[
f(p) = \frac{\partial^2}{\partial x^2} p \frac{\partial^2}{\partial y^2} p - \left( \frac{\partial}{\partial x} \frac{\partial}{\partial y} p \right)^2.
\]

5.7.7 (Jacobson-Morozov Lemma). Let \( V \in \text{Vect}_k \) be finite dimensional and let \( \phi \in \text{End}_k(V) \) be nilpotent. Show that, up to isomorphism, there is a unique \( sl_2 \)-module structure on \( V \) such that \( e_V = \phi \). (If \( V \) is indecomposable for \( \phi \), then \( V \) has a \( k \)-basis \( (b_i)_{i=0}^m \) with \( \phi(b_i) = ib_{i-1} \). Mimic (5.51) to define \( f_V \) and \( h_V \).)

5.7.8 (The algebras \( B(\lambda) \) and the Weyl algebra). For a given \( \lambda \in k \), let \( P(\lambda) \) denote the corresponding minimal primitive ideal of \( U \) as in Theorem 5.48, and let \( B(\lambda) = U / P(\lambda) \).

(a) Show that \( B(\lambda) \) embeds into the Weyl algebra \( A_1 = k\langle x, y \rangle / (yx - xy - 1) \) via \( e \mapsto -\mu - yx^2, f \mapsto y \) and \( h \mapsto \mu + 2yx \). Here, \( \mu \in k \) is chosen so that \( \mu^2 + 2\mu = \lambda \).

(b) Show that \( 1 \notin [B(\lambda), B(\lambda)] \). Hence \( B(\lambda) \) is not isomorphic to \( A_1 \).\(^{11}\)

5.7.9 (The restricted enveloping algebra of \( sl_2 \)). Assume that \( \text{char} k = p > 0 \) and put \( A = U / \langle e^p, f^p, h^p - h \rangle \); the algebra \( A \) is called the **restricted enveloping algebra** of \( sl_2 \) in characteristic \( p \). Let \( k^2 \) denote the defining representation of \( sl_2 \). Show:

(a) \( \dim_k A = p^3 \).

(b) The representations \( V(m) = \text{Sym}^m(k^2) \in \text{Rep} U \) with \( 0 \leq m < p \) give a full set of non-isomorphic irreducible representations of \( A \).\(^{12}\) (Adapt the proof of Lemma 5.40.)

---

\(^{11}\) Dixmier [58] has also shown that \( B(\lambda) \neq B(\lambda') \) if \( \lambda \neq \lambda' \).

\(^{12}\) See also Exercise 3.6.1.
Semisimple Lie Algebras

Among all finite-dimensional Lie algebras, semisimple Lie algebras have the richest structure and are the ones most frequently encountered in applications. Therefore, the remainder of Part III is entirely focused on the semisimple case. The current chapter provides the foundations, while Chapters 7 and 8 will dig deeper into the structure, classification and representation theory of semisimple Lie algebras.

Recall (§5.2.2) that a finite-dimensional Lie algebra \( g \) is called

**semisimple** if \( g \) has no nonzero abelian ideals or, equivalently, \( \text{rad} \ g = 0 \)

and

**simple** if \( g \) is non-abelian and has no ideals other than 0 and \( g \).

Equivalently, a finite-dimensional \( g \in \text{Lie}_k \) is simple if and only if the adjoint representation \( g_{\text{ad}} \) is irreducible and \( \neq 1 \). As we have seen (Example 5.14), this certainly holds for \( g = \mathfrak{sl}_2(k) \) if \( \text{char} \ k \neq 2 \). The Lie algebra \( \mathfrak{sl}_2 \) will in fact play a crucial role in our analysis of general semisimple Lie algebras.

Simple Lie algebras are evidently semisimple. As one of several characterizations of semisimplicity proved in this chapter for a base field \( k \) with \( \text{char} \ k = 0 \), we shall show that semisimple Lie algebras are the same as finite direct products of simple Lie algebras. We then proceed to describe the so-called root space decomposition of semisimple Lie algebras. Finally, we will discuss in detail four “classical” infinite families of simple Lie algebras that are closely related to symmetries of Euclidean spaces; they are commonly referred to as the Lie algebras of types A, B, C and D.
6.1. Characterizations of Semisimplicity

The following theorem summarizes various characterizations of semisimplicity for Lie algebras. The reader may wish to compare the theorem with corresponding results for associative algebras in §1.4.4, especially the direct product decomposition in Wedderburn’s Structure Theorem.

**Theorem 6.1 (char k = 0).** The following are equivalent for a finite-dimensional \( g \in \text{Lie}_k \):

(i) \( g \) is semisimple;

(ii) \( g \) is isomorphic to a finite direct products of simple Lie algebras;

(iii) the Killing form of \( g \) is non-degenerate;

(iv) all finite-dimensional representations of \( g \) are completely reducible.

The Killing form, a bilinear form \( g \times g \rightarrow k \), will be introduced below. In this section, we also prove the equivalence of (i), (ii) and (iii) and show that these conditions follow from (iv). The converse is a celebrated result due to Hermann Weyl; the proof requires further preparations and will be given in the next section.

6.1.1. The Killing Form

Let \( g \in \text{Lie}_k \) and \( V \in \text{Rep}_{\text{fin}} g \). Then we may define a bilinear form \( B_V : g \times g \rightarrow k \) by

\[
B_V(x, y) \overset{\text{def}}{=} \text{trace}(x_V y_V) \quad (x, y \in g),
\]

where \( x_V y_V \) is the product (composite) of the operators \( x_V, y_V \in \text{End}_k(V) \). Besides obviously being symmetric, the form \( B_V \) is also associative by (5.13):

\[
B_V([x, y], z) = B_V([x, y], z) \quad (x, y, z \in g).
\]

Associativity is equivalent to the fact that the following map is a morphism in \( \text{Rep}_g \):

\[
x \mapsto B_V(x, .)
\]

Therefore, \( \text{rad} B_V := \{ x \in g \mid B_V(x, .) = 0 \} \) is an ideal of \( g \) containing \( \text{Ker}_g V \). The form \( B_V \) is said to be non-degenerate if \( \text{rad} B_V = 0 \).

Now assume that \( g \) is finite dimensional. Then the **Killing form** of \( g \) is defined to be the form \( B_V \) for \( V = g_{\text{ad}} \), the adjoint representation of \( g \). We will write this form simply as \( B \) or occasionally as \( B_g \) if \( g \) needs to be made explicit. So

\[
B(x, y) = \text{trace}(\text{ad} x \text{ ad} y)
\]
The Killing form of an ideal. Let \( \mathfrak{a} \) be an ideal of \( \mathfrak{g} \), viewed as a Lie algebra in its own right. The Killing form of \( \mathfrak{a} \) is the restriction of the Killing form of \( \mathfrak{g} \) to \( \mathfrak{a} \times \mathfrak{a} \):

\[
B_{\mathfrak{a}} = B_{\mathfrak{g}}|_{\mathfrak{a} \times \mathfrak{a}}.
\]

Indeed, \((\text{ad}_\mathfrak{g} \times \mathfrak{a}) \mathfrak{g} = [\mathfrak{x}, \mathfrak{g}] \subseteq \mathfrak{a} \) for \( \mathfrak{x} \in \mathfrak{a} \). Hence, extending a basis of \( \mathfrak{a} \) to a basis of \( \mathfrak{g} \), the matrix of \( \text{ad}_\mathfrak{g} \mathfrak{x} \in \text{End}_k(\mathfrak{g}) \) has the form \(
\begin{pmatrix}
  a_{x} & * \\
  0 & 0
\end{pmatrix}
\)

where \( a_x \) is the matrix of \( \text{ad}_\mathfrak{g} \mathfrak{x} \). If \( \mathfrak{x}, \mathfrak{y} \in \mathfrak{a} \), then the matrix of the product \( \text{ad}_\mathfrak{g} \mathfrak{x} \text{ad}_\mathfrak{g} \mathfrak{y} \) has the same form, with \( a_x a_y \) in place of \( a_x \). So \( \text{trace}(\text{ad}_\mathfrak{g} \mathfrak{x} \text{ad}_\mathfrak{g} \mathfrak{y}) = \text{trace}(a_x a_y) = \text{trace}(\text{ad}_\mathfrak{g} \mathfrak{x} \times \text{ad}_\mathfrak{g} \mathfrak{y}) \).

**Example 6.2** (Killing form of \( \mathfrak{sl}_2 \)). Recall that \( \mathfrak{sl}_2 = k f \oplus k h \oplus k e \) with \([h, f] = -2f, [h, e] = 2e \) and \([e, f] = h \). The matrices of \( \text{ad} f, \text{ad} h \) and \( \text{ad} e \) for the basis \((f, h, e)\) of \( \mathfrak{sl}_2 \) are respectively \(
\begin{pmatrix}
  0 & 2 & 0 \\
  0 & 0 & -1 \\
  0 & 0 & 0
\end{pmatrix}
\) and \(
\begin{pmatrix}
  0 & 0 & 0 \\
  0 & 0 & 0 \\
  0 & 0 & 2
\end{pmatrix}
\). Taking traces of all products of these matrices, one obtains the matrix of the Killing form:

\[
\begin{pmatrix}
  0 & 0 & 8 \\
  0 & 8 & 0 \\
  4 & 0 & 0
\end{pmatrix}
\]

The following lemma assumes \( \text{char } k = 0 \), because the proof relies on Cartan’s Criterion (§5.3.4).

**Lemma 6.3** (char \( k = 0 \)). Let \( \mathfrak{g} \in \text{Lie}_k \) be finite dimensional and let \( V \in \text{Rep}_{\text{fin}} \mathfrak{g} \) be such that \( \text{Ker} \mathfrak{g} \) is solvable. Then \( B_V \subseteq \text{rad} \mathfrak{g} \). In particular, \( \text{rad} B \subseteq \text{rad} \mathfrak{g} \).

**Proof.** Since \( \text{trace}(x v y v) = 0 \) for all \( x, y \in \text{rad} B \), Cartan’s Criterion implies that the image of \( \text{rad} B \) in \( \mathfrak{g}(V) \) is solvable. Our hypothesis on \( \text{Ker} \mathfrak{g} \) therefore implies that \( \text{rad} B \) is a solvable ideal of \( \mathfrak{g} \) (Proposition 5.12). Therefore, \( \text{rad} B \subseteq \text{rad} \mathfrak{g} \).

For the Killing form \( B \), note that \( \text{Ker} \mathfrak{g} \text{ad} \mathfrak{g} = \mathfrak{z} \mathfrak{g} \) is certainly solvable. \( \square \)

### 6.1.2. Start of the Proof of Theorem 6.1

We will now prove the equivalence of (i), (ii) and (iii) in Theorem 6.1 as well as the implication (iv) \( \Rightarrow \) (ii). The proof of (i) \( \Rightarrow \) (iv) is postponed until §6.2.2.

(i) \( \Leftrightarrow \) (iii). If \( \mathfrak{g} \) is semisimple, then \( \text{rad} \mathfrak{g} = 0 \). In particular, we must have \( \text{rad} B = 0 \) by Lemma 6.3; so \( B \) is non-degenerate.

Conversely, assume that \( B \) is non-degenerate and let \( \mathfrak{a} \) be an abelian ideal of \( \mathfrak{g} \). It suffices to show that \( B(\mathfrak{x}, \mathfrak{y}) = 0 \) for all \( \mathfrak{x} \in \mathfrak{a} \) and \( \mathfrak{y} \in \mathfrak{g} \), for then \( \mathfrak{a} \subseteq \text{rad} \mathfrak{g} = 0 \). But \( \text{ad} \mathfrak{x} \text{ad} \mathfrak{y}: \mathfrak{g} \to \mathfrak{g} \to \mathfrak{a} \) and \((\text{ad} \mathfrak{x} \text{ad} \mathfrak{y})^2: \mathfrak{g} \to \mathfrak{a} \to [\mathfrak{a}, \mathfrak{a}] = 0 \). Thus, \( \text{ad} \mathfrak{x} \text{ad} \mathfrak{y} \) is a nilpotent operator, and hence it must have trace zero, whence \( B(\mathfrak{x}, \mathfrak{y}) = 0 \) as desired.

(i),(iii) \( \Rightarrow \) (ii). We may assume that \( \mathfrak{g} \neq 0 \). By (6.3), non-degeneracy of \( B \) says that the map \( x \mapsto B(\mathfrak{x}, \cdot) \) gives an isomorphism \( \mathfrak{g}_{\text{ad}} \xrightarrow{\sim} \mathfrak{g}_{\text{ad}}^* \) in \( \text{Rep} \mathfrak{g} \). If \( \mathfrak{g} \) has no proper nonzero ideals then \( \mathfrak{g} \) is already simple, because \( \mathfrak{g} \) is certainly non-abelian: otherwise, \( B = 0 \). So assume that \( \mathfrak{a} \) is a proper nonzero ideal of \( \mathfrak{g} \) and consider the epimorphism \( \mathfrak{g}_{\text{ad}} \xrightarrow{\sim} \mathfrak{g}_{\text{ad}}^* \to \mathfrak{a}^* \) in \( \text{Rep} \mathfrak{g} \) that is given by \( x \mapsto B(\mathfrak{x}, \cdot)|_{\mathfrak{a}} \). The kernel of this map, \( \mathfrak{a}^* = \{ \mathfrak{x} \in \mathfrak{g} \mid B(\mathfrak{x}, \mathfrak{a}) = 0 \} \), is an ideal of \( \mathfrak{g} \) with \( \mathfrak{g}/\mathfrak{a}^* \cong \mathfrak{a}^* \).
Claim. \( g \cong a \times a^\perp \).

To see this, note that the restriction of \( B \) to the ideal \( a \cap a^\perp \) of \( g \) vanishes. Thus, \( B_{a \cap a^\perp} = 0 \) by (6.5) and Lemma 6.3 further implies that \( a \cap a^\perp \) is solvable. Therefore, \( a \cap a^\perp \subseteq \text{rad } g \), giving \( a \cap a^\perp = 0 \) by (i). Since \( g/a^\perp \cong a^\perp \) and we also obtain \([a, a^\perp] = 0\), because \([a, a^\perp] \subseteq a \cap a^\perp\). This proves the claim.

Since \( B \) is non-degenerate, \( B_{a} \) and \( B_{a^\perp} \) are both non-degenerate as well by (6.5). Therefore, by induction on the dimension, both \( a \) and \( a^\perp \) are finite direct products of simple Lie algebras, which yields the desired decomposition of \( g \).

(ii) \( \Rightarrow \) (i). Now assume that \( g \cong g_1 \times \cdots \times g_t \) for simple Lie algebras \( g_i \). Equivalently, the adjoint representation \( g_{\text{ad}} \) is completely reducible with no \( 1 \)-components. Since ideals are the same as subrepresentations of \( g_{\text{ad}} \), any ideal \( a \) of \( g \) has a complement, say \( g = a \oplus b \), and \( a \) is also completely reducible with no \( 1 \)-components (Corollary 1.29). It follows that \([b, a] = 0\) and so \([g, a] = [a, a]\). Thus, if \( a \) is abelian, then \( a \) is \( 1 \)-homogeneous, which forces \( a = 0 \). Therefore, \( g \) is semisimple. This completes the proof of the equivalence of (i), (ii) and (iii).

(iv) \( \Rightarrow \) (i). Complete reducibility of \( g_{\text{ad}} \) says that \( g \) is a (finite) direct sum of simple ideals, say \( g = \bigoplus_i g_i \). Since \( 0 = g_i \cap g_j \supseteq [g_i, g_j] \) for \( i \neq j \), the Lie algebra \( g \) is isomorphic to the direct product of the Lie algebras \( g_i \). Also, the only ideals of each \( g_i \) are \( 0 \) and \( g_i \). It remains to show that no \( g_i \) is abelian or, equivalently, \( 1 \)-dimensional. But if \( g_i = kx \), say, then there are \( g_i \)-representations that are not completely reducible; for example, we can take \( k^2 \) with \( x \) acting via \(
\begin{pmatrix}
0 & 1 \\
0 & 0
\end{pmatrix}
\). Pulling this back along the projection \( g \rightarrow g_i \) yields a representation of \( g \) that is not completely reducible, contrary to our hypothesis. Therefore all components \( g_i \) must be non-abelian.

6.1.3. Some Consequences of Theorem 6.1

Even though the implication (i) \( \Rightarrow \) (iv) of Theorem 6.1 is still unproven, we proceed to list several properties of a semisimple Lie algebra \( g \) that follow from the equivalent characterizations (i)–(iii) in the theorem. The only part that is not immediate is the assertion that \( g \) has no outer derivations, that is, \( \text{ad } g = \text{Der } g \).

Corollary 6.4 (char \( k = 0 \)). Let \( g \in \text{Lie}_k \) be semisimple. Then all ideals and all homomorphic images of \( g \) are semisimple. Moreover, \( g = [g, g] \) and \( \text{ad } : g \rightarrow \text{Der } g \).

Proof. Any epimorphism \( g \rightarrow \overline{g} \) in \( \text{Lie}_k \) is also an epimorphism in \( \text{Rep } g \) for the adjoint \( g \)-actions on \( g \) and on \( \overline{g} \) and the latter action descends to a \( g \)-action. Since \( g_{\text{ad}} \) is completely reducible with no \( 1 \)-components by Theorem 6.1(ii), the same holds for \( \overline{g}_{\text{ad}} \) (Corollary 1.29). Hence \( \overline{g} \) is semisimple.
Next, let \( a \) be an ideal of \( \mathfrak{g} \). Since the adjoint representation \( \text{ad} \mathfrak{g} \) is completely reducible, there exists an ideal \( \mathfrak{b} \) of \( \mathfrak{g} \) such that \( \mathfrak{g} = \mathfrak{a} \oplus \mathfrak{b} \). Therefore, \( \mathfrak{a} \cong \mathfrak{g}/\mathfrak{b} \) is a homomorphic image of \( \mathfrak{g} \) and as such, \( \mathfrak{a} \) is semisimple.

As for the equality \( \mathfrak{g} = [\mathfrak{g}, \mathfrak{g}] \), note that \( \mathfrak{g}/[\mathfrak{g}, \mathfrak{g}] \) is semisimple and abelian, and hence it must vanish.

It remains to prove to establish the isomorphism \( \text{ad} : \mathfrak{g} \to \text{Der} \mathfrak{g} \). Certainly \( \text{Ker} \text{ad} = \{0\} \); so the crux is the equality \( \text{ad} \mathfrak{g} = \text{Der} \mathfrak{g} \). Put \( a = \text{ad} \mathfrak{g} \) and \( b = \text{Der} \mathfrak{g} \) and recall that \( a \) is an ideal of \( b \) by virtue of the identity \( [d, \text{ad} x] = \text{ad} d(x) \) for \( d \in b \) and \( x \in \mathfrak{g} \) (Exercise 5.1.5). Since \( \mathfrak{a} \cong \mathfrak{g} \) is semisimple, the Killing form of \( \mathfrak{a} \) is non-degenerate. In view of (6.5), this says that \( a \cap a^\perp = 0 \), where we have put \( a^\perp = \{ d \in b | B_b(d, a) = 0 \} \). Recall that \( a^\perp \) is an ideal of \( b \) by associativity of \( B_b \). If \( d \in a^\perp \) and \( x \in \mathfrak{g} \), then \( \text{ad} d(x) = [d, \text{ad} x] \in [a^\perp, a] \subseteq a^\perp \cap a = 0 \) and so \( d(x) = 0 \), because \( \text{ad} \) is injective. This shows that \( a^\perp = 0 \), giving an embedding \( \mathfrak{b} \hookrightarrow \mathfrak{a}^* \). Therefore, \( \mathfrak{b} = \mathfrak{a} \) for dimension reasons, which finishes the proof. \( \square \)

**Exercises for Section 6.1**

**6.1.1** (Killing form of nilpotent and solvable Lie algebras). Show that the Killing form of any finite-dimensional nilpotent Lie algebra is identically 0, but this need not be so for solvable Lie algebras.

**6.1.2** (Bilinear forms). (a) Let \( V \in \text{Vect}_k \) with \( \dim_k V < \infty \) and let \( b : V \times V \to k \) be a bilinear form. Show that \( \dim_k \{ v \in V | b(\cdot, v) = 0 \} = \dim_k \{ v \in V | b(\cdot, \cdot) = 0 \} \).

(b) Let \( \mathfrak{g} \in \text{Lie}_k \) be such that \( \mathfrak{g} = [\mathfrak{g}, \mathfrak{g}] \). Show that every associative \( k \)-bilinear form \( \beta : \mathfrak{g} \times \mathfrak{g} \to k \) is symmetric.

(c) Let \( \mathfrak{g} \in \text{Lie}_k \) be simple and let \( \beta, \gamma : \mathfrak{g} \times \mathfrak{g} \to k \) be two nonzero associative \( k \)-bilinear forms. Assuming \( k \) to be algebraically closed, show that \( \gamma = \lambda \beta \) for some \( \lambda \in k^\times \).

**6.1.3** (Ideals in direct products). Assume that \( \mathfrak{g} = \mathfrak{g}_1 \times \cdots \times \mathfrak{g}_t \) for simple \( \mathfrak{g}_i \in \text{Lie}_k \), with each \( \mathfrak{g}_i \) either \( \mathfrak{g}_i \) or 0.

**6.1.4** (Semisimplicity and field extensions). Let \( \mathfrak{g} \in \text{Lie}_k \) and let \( K/k \) be a field extension. Consider the Lie \( K \)-algebra \( \mathfrak{g} \otimes K \) (Exercise 5.2.5) and show:

(a) If \( \mathfrak{g} \otimes K \) is semisimple, then \( \mathfrak{g} \).

(b) The converse holds if \( k \) is perfect. (Use Exercise 1.4.7.)

**6.2. Complete Reducibility**

In this section, we complete the proof of Theorem 6.1 by proving the following theorem which is a fundamental result in the representation theory of Lie algebras.
Weyl’s Theorem \((\text{char } k = 0)\). All finite-dimensional representations of semisimple Lie algebras are completely reducible.

The special case of \(\mathfrak{sl}_2\) was treated earlier, working over a field \(k\) of characteristic 0 that is algebraically closed (Theorem 5.39). However, for any semisimple \(g \in \mathfrak{Lie}_k\), Weyl’s Theorem easily reduces to the case where \(k\) is algebraically closed. Indeed, letting \(\bar{k}\) denote an algebraic closure of \(k\), the Lie \(\bar{k}\)-algebra \(g \otimes \bar{k}\) is also semisimple (Exercise 6.1.4). Therefore, for any \(V \in \text{Rep}_{\text{fin}} g\), Weyl’s Theorem for algebraically closed fields states that \(V \otimes \bar{k} \in \text{Rep}(g \otimes \bar{k})\) is completely reducible, and this in turn easily implies complete reducibility of \(V\) (Exercise 1.4.7).

As we have already pointed out for \(\mathfrak{sl}_2\), Weyl’s Theorem immediately extends to locally finite representations of a semisimple Lie algebra \(g\). Thus, for example, the adjoint representation \((U g)_{\text{ad}}\) is completely reducible.

The proof of Weyl’s Theorem will be given in 6.2.2 after discussing some preliminaries about Casimir elements, which are interesting in their own right. We will end this section with an application of Weyl’s Theorem to Jordan canonical form.

### 6.2.1. Casimir Elements

To start, we explain some general ideas that have already been used to similar effect in the setting of Frobenius algebras (§2.2.3). Let \(g\) be an arbitrary finite-dimensional Lie algebra that is equipped with an associative and non-degenerate bilinear form, \(b: g \times g \to k\). Associativity of \(b\) is equivalent to the fact that the following map is a morphism in \(\text{Rep} g\) and non-degeneracy says it is an isomorphism:

\[
\begin{array}{ccc}
  d: & g_{\text{ad}} & \sim \rightarrow g^*_{\text{ad}} \\
  & w & w \\
  & x & \mapsto b(x, \cdot)
\end{array}
\]

Thus, we obtain the following isomorphism in \(\text{Rep} g\):

\[
\delta: \text{End}_k(g_{\text{ad}}) \xrightarrow{(5.30)} g_{\text{ad}} \otimes g^*_{\text{ad}} \xrightarrow{\sim} g_{\text{ad}} \otimes g_{\text{ad}}.
\]

Now let \(f: g_{\text{ad}} \to A\) in \(\text{Rep} g\) be given, where \(A\) be a \(g\)-algebra (§5.5.5); so the multiplication \(m: A \otimes A \to A\) is a map in \(\text{Rep} g\). Then we obtain the map \(m_f := m \circ (f \otimes f) \circ \delta: \text{End}_k(g_{\text{ad}}) \to g_{\text{ad}} \otimes g_{\text{ad}} \to A \otimes A \to A\) in \(\text{Rep} g\). Since \(\text{Id}_g \in \text{End}_k(g_{\text{ad}})\), we further obtain a \(g\)-invariant,

\[
c(f) := m_f(\text{Id}_g) \in A^g.
\]

Explicitly, writing \(\delta(\text{Id}_g) = \sum_i x_i \otimes y_i\), we have \(x = \sum_i x_i b(y_i, x)\) for all \(x \in g\). If the \(x_i\) are chosen linearly independent, as we may, then this condition says that
(\(x_i, y_i\)) are dual bases of \(\mathfrak{g}\) for the form \(b\), in the sense that \(b(y_i, x_j) = \delta_{i,j}\). The invariant \(c(f)\) is given by

\[
c(f) = \sum_i f(x_i) f(y_i).
\]

**Universal Casimir Element.** Now let us focus on the case where \(\mathfrak{g}\) is semisimple and \(\text{char } \mathbb{k} = 0\). Then, in the foregoing, we may choose the Killing form \(b = B\), the enveloping algebra \(A = \mathfrak{U}_\mathfrak{g}\) with the adjoint \(\mathfrak{g}\)-action, and the canonical embedding \(f: \mathfrak{g} \hookrightarrow U\) (§5.5.6). In this way, we obtain the **universal Casimir element** of \(\mathfrak{g}\):

\[
(6.7) \quad c = c(\mathfrak{g}) \in (\mathfrak{U}_\mathfrak{g})^0 = \mathfrak{Z}(\mathfrak{U}_\mathfrak{g}).
\]

**Example 6.5** (Universal Casimir element of \(\mathfrak{sl}_2\)). Recall from Example 6.2 that the Killing form of \(\mathfrak{sl}_2\) has matrix \[
\begin{pmatrix}
0 & 0 & 0 \\
0 & 8 & 0 \\
0 & 0 & 4
\end{pmatrix}
\]
for the standard basis \(f, h, e\) of \(\mathfrak{sl}_2\). Thus, the dual basis of \(\mathfrak{sl}_2\) is \(\frac{1}{2}e, \frac{1}{8}h, \frac{1}{2}f\). Therefore, the universal Casimir element of \(\mathfrak{sl}_2\) is

\[
c(\mathfrak{sl}_2) = \frac{1}{8}(2fe + h^2 + 2ef).
\]

This is identical to the Casimir element (5.54) up to the factor \(\frac{1}{8}\).

**Casimir Element of a Representation.** For any semisimple \(\mathfrak{g} \in \text{Lie}_\mathbb{k}\) with \(\text{char } \mathbb{k} = 0\), the form \(B_V\) that is associated to a given \(V \in \text{Rep}_\text{fin} \mathfrak{g}\) is also associative (and symmetric) and it is non-degenerate provided \(V\) is \(\mathfrak{g}\)-faithful (Lemma 6.3). Taking \(b = B_V\), \(A = \text{End}_\mathbb{k}(V)\) and letting the map \(\mathfrak{g} \rightarrow \text{End}_\mathbb{k}(V), x \mapsto x_V\), play the role of \(f\) (Exercise 5.5.1), we obtain the **Casimir element of the representation** \(V\):

\[
(6.8) \quad c(V) \in \text{End}_\mathbb{k}(V)^0 = \text{End}_{\mathbb{k}^0}(V).
\]

Part (b) of the following lemma explains the name “universal” Casimir element.

**Lemma 6.6** (\(\text{char } \mathbb{k} = 0\)). Let \(\mathfrak{g} \in \text{Lie}_\mathbb{k}\) be semisimple, let \(V \in \text{Rep}_\text{fin} \mathfrak{g}\) be \(\mathfrak{g}\)-faithful and let \(c = c(\mathfrak{g})\) be the universal Casimir element. Then:

(a) \(\text{trace } c(V) = \text{dim}_\mathbb{k} \mathfrak{g}\).

(b) If \(\mathfrak{g}\) is simple and \(\mathbb{k}\) is algebraically closed, then the Killing form \(B = B_\mathfrak{g}\) is a scalar multiple of \(B_V\) and \(c_V\) is a nonzero scalar multiple of \(c(V)\).

**Proof.** (a) If \((x_i, y_i)\) are dual bases of \(\mathfrak{g}\), with \(B_V(x_i, y_j) = \delta_{i,j}\), then \(c(V) = \sum_i (x_i)_V(y_i)_V\). Thus,

\[
\text{trace } c(V) = \sum_i \text{trace}((x_i)_V(y_i)_V) = \sum_i B_V(x_i, y_i) = \text{dim}_\mathbb{k} \mathfrak{g}.
\]

(b) If \(\mathfrak{g}\) is simple, then \(\mathfrak{g}_{\text{ad}} \in \text{Rep} \mathfrak{g}\) is irreducible. Thus, by Schur’s Lemma, any two isomorphisms \(\mathfrak{g}_{\text{ad}} \cong \mathfrak{g}_{\text{ad}}^\ast\) in \(\text{Rep} \mathfrak{g}\) are scalar multiples of each other. In particular, in view of (6.3), we must have \(B = \lambda B_V\) for some \(\lambda \in \mathbb{k}^\times\). If \((x_i, y_i)\) are
bases of \( \mathfrak{g} \) that are dual for \( B_V \) as in the proof of (a), then \((x_i, \lambda^{-1} y_i)_i\) are dual for \( B \). Thus, \( c = \sum_i x_i \lambda^{-1} y_i \) and \( \lambda c_V = \sum_i (x_i)_V(y_i)_V = c(V) \).

If \( V \) is irreducible and \( k \) is algebraically closed, then \( \text{End}_{U_0}(V) = k \text{ Id}_V \) and part (a) of the lemma gives \( c(V) = \frac{\dim_k b}{\dim_k V} \text{ Id}_V \).

### 6.2.2. Proof of Weyl’s Theorem

Let \( \mathfrak{g} \in \text{Lie}_k \) be semisimple, with char \( k = 0 \), and let \( V \in \text{Rep}_{\text{fin}} \mathfrak{g} \). We need to show that each subrepresentation \( W \subseteq V \) has a complement or, equivalently, there is a projection \( \phi : V \to W \) in \( \text{Rep} \mathfrak{g} \) with \( \phi|_W = \text{Id}_W \). Since homomorphic images of \( \mathfrak{g} \) are semisimple (Corollary 6.4), we may replace \( \mathfrak{g} \) by its image in \( \mathfrak{sl}(V) \) and hence assume that \( V \) is \( \mathfrak{g} \)-faithful. As we have remarked earlier, we may also assume \( k \) to be algebraically closed. We will write \( U = U_0 \).

**Step 1:** Reduction to the case where \( V/W \cong \mathbb{I} \). Suppose we can find complements for subrepresentations with factor \( \mathbb{I} \). Then, for an arbitrary subrepresentation \( 0 \neq W \subseteq V \), consider \( \text{Hom}_k(V, W) \in \text{Rep} \mathfrak{g} \) and define

\[
\mathcal{V} := \{ \phi \in \text{Hom}_k(V, W) : |\phi|_W \in k \text{ Id}_W \} \supseteq \mathcal{W} := \{ \phi \in \text{Hom}_k(V, W) : \phi|_W = 0 \}.
\]

Then \( \dim_k \mathcal{V}/\mathcal{W} = 1 \), because \( \mathcal{W} \) is the kernel of the map \( \mathcal{V} \to k \text{ Id}_W, \phi \mapsto \phi|_W \). Moreover, given \( x \in \mathfrak{g}, w \in W \) and \( \phi \in \mathcal{V} \), with \( \phi|_W = \lambda \text{ Id}_W \) say, we compute \( (x, \phi)(w) = x \cdot \phi(w) - \phi(xw) = x \lambda w - \lambda xw = 0 \). Thus, \( \mathfrak{g}, \mathcal{V} \subseteq \mathcal{W} \) and so \( \mathcal{V}, \mathcal{W} \in \text{Rep}_{\text{fin}} \mathfrak{g} \) with \( \mathcal{V}/\mathcal{W} \cong \mathbb{I} \). By assumption, \( \mathcal{V} \cong \mathcal{W} \oplus k \phi \) for some \( \phi \in \mathcal{V}^0 \), so \( \phi \in \text{Hom}_L(V, W) \) and \( \phi|_W = \lambda \text{ Id}_W \) with \( \lambda \in k^X \). Replacing \( \phi \) by a \( \lambda^{-1} \phi \), we obtain the desired projection of \( V \) onto \( W \).

**Step 2:** Reduction to the case where \( V/W \cong \mathbb{I} \) and \( W \) is irreducible. Assume that \( V/W \cong \mathbb{I} \) and proceed by induction on the dimension. If \( U \) is a nonzero proper subrepresentation of \( W \) then \((V/U)/(W/U) \cong V/W \cong \mathbb{I} \). By induction we may write \( V/U = W/U \oplus C/U \) for some subrepresentation \( C \) with \( U \subseteq C \subseteq V \). Since \( C/U \cong \mathbb{I} \), we may use induction again to conclude that \( C = U \oplus k v \) for some \( v \in \mathcal{V}^0 \). Since \( v \notin W \), we obtain \( V = W \oplus k v \) as desired.

**Step 3:** End of proof. We may assume that \( V/W \cong \mathbb{I} \) and \( W \) is irreducible. Consider the Casimir element \( c(V) \in \text{End}_U(V) \). Writing \( c(V) = \sum_j (x_j)_V(y_j)_V \) as in the proof of Lemma 6.6 and using the fact that \( V/W \cong \mathbb{I} \), we see that \( c(V) \) maps \( V \) into \( W \). Since \( W \) is irreducible, Schur’s Lemma further gives that \( c(V)|_W = \lambda \text{ Id}_W \) for some \( \lambda \in k \). Thus, trace \( c(V) = \lambda \dim_k W \) and we also know by Lemma 6.6(a) that \( \lambda \neq 0 \). Therefore, \( \lambda^{-1} c_V \) is the desired projection of \( V \) onto \( W \). This completes the proof of Weyl’s Theorem and hence of Theorem 6.1 as well. \( \square \)
6.2. Complete Reducibility

6.2.3. Reductive Lie Algebras

A finite-dimensional \( g \in \text{Lie}_k \) with \( \text{char} \, k = 0 \) is called reductive if \( \text{rad} \, g = Z \, g \). Any reductive Lie algebra \( g \) has the form

\[
(6.9) \quad g = \mathcal{Z} \, g \oplus [g, g]
\]

and \([g, g]\) is semisimple. The latter is clear, because (6.9) implies \([g, g] \cong g/\text{rad} \, g\).

To obtain the decomposition (6.9), consider the adjoint representation \( g_{ad} \). This representation is completely reducible by Weyl’s Theorem, because \( g \) acts through the semisimple quotient \( g/\mathcal{Z} \, g = g/\text{rad} \, g \). So \( g_{ad} = g_{ad}(1) \oplus g' \), where \( g' \) is the sum of the homogeneous components \( g_{ad}(S) \) with \( S \in \text{Irr} \, g \setminus \{1\} \). Clearly, \( g_{ad}(1) = \mathcal{Z} \, g \) and \( g \cdot S = S \) for all \( S \neq 1 \). Therefore, \( g' = [g, g'] = [g, g] \), proving (6.9).

The following proposition shows that reductive Lie algebras are a natural class of Lie algebras to consider. It shows, for example, that the Lie algebra \( gl_n(k) \) is reductive. Some further properties of reductive Lie algebras are explored in Exercises 6.2.2 and 6.2.3.

**Proposition 6.7** (char \( k = 0 \)). Let \( g \in \text{Lie}_k \) be arbitrary and let \( V \in \text{Rep}_{\text{fin}} \, g \) be completely reducible. Then the image of \( g \) in \( gl(V) \) is reductive.

**Proof.** Replacing \( g \) by its image in \( gl(V) \), we may assume that \( V \) is \( g \)-faithful and our goal is to show that \( \text{rad} \, g \subseteq \mathcal{Z} \, g \). Let \( \bar{k} \) be an algebraic closure of \( k \).

Then the representation \( \bar{k} \otimes V \) of the Lie algebra \( \bar{k} \otimes g \) is completely reducible (Exercise 1.4.7) and faithfulness is clearly preserved as well. It suffices to show that \( \text{rad}(\bar{k} \otimes g) \subseteq \mathcal{Z}(\bar{k} \otimes g) \), because this will imply \( \text{rad} \, g \subseteq \text{rad}(\bar{k} \otimes g) \cap g \subseteq \mathcal{Z}(\bar{k} \otimes g) \cap g \subseteq \mathcal{Z} \, g \) as desired. Thus, we may assume that \( \bar{k} \) is algebraically closed.

Write \( V = V_1 \oplus \cdots \oplus V_r \) with irreducible representations \( V_i \) and use Lie’s Theorem (§5.3.2) to select eigenvectors \( 0 \neq v_i \in V_i \) for \( \text{rad} \, g \), say \( v_i \) has weight \( \lambda_i \in (\text{rad} \, g)^* \).

Since each weight space \( (V_i)_{\lambda_i} \) is a subrepresentation of \( V_i \) (Lemma 5.17), we conclude that \( (V_i)_{\lambda_i} = V_i \) for all \( i \). Thus, \( \text{rad} \, g \) acts on all \( V_i \) by scalars, which yields the desired inclusion \( \text{rad} \, g \subseteq \mathcal{Z} \, g \) by faithfulness. \( \square \)

6.2.4. Abstract Jordan Decomposition

We finish this section with an application of Weyl’s Theorem that will be useful later on. An element \( x \in g \) will be called ad-semisimple if the operator \( \text{ad} \, x \in gl(\mathfrak{g}) \) is diagonalizable; similarly \( x \) is said to be ad-nilpotent if \( \text{ad} \, x \) is nilpotent.

**Proposition 6.8** (\( k \) algebraically closed, char \( k = 0 \)). Let \( g \in \text{Lie}_k \) be semisimple and let \( x \in g \). Then:

(a) **Abstract Jordan decomposition.** There are unique elements \( x_s, x_n \in g \) with \( x = x_s + x_n \), \( [x_s, x_n] = 0 \) and such that \( x_s \) is ad-semisimple while \( x_n \) is ad-nilpotent.
(b) **Preservation of Jordan decomposition.** For any $V \in \text{Rep}_{\text{fin}} \mathfrak{g}$, the ordinary Jordan decomposition of $x_V \in \text{End}_k(V)$ is given by $x_V = (x_s)_V + (x_n)_V$.

**Proof.** The crucial observation, for both (a) and (b), is contained in the following claim. We let $V \in \text{Rep}_{\text{fin}} \mathfrak{g}$ and $x \in \mathfrak{g}$ as in the statement of the proposition. Furthermore, we denote the image of the map $x \mapsto x_V$ by $\mathfrak{g}_V$.

**Claim.** Let $x_V = s + n$ is the ordinary Jordan decomposition of $x_V \in \text{End}_k(V)$, with $s$ diagonalizable, $n$ nilpotent, and $sn = ns$ (Proposition 5.19). Then $s, n \in \mathfrak{g}_V$.

In order to prove this, we may replace $\mathfrak{g}$ by $\mathfrak{g}_V$, thereby reducing to the case where $x \in \mathfrak{g} \subseteq \text{gl}(V)$ and $x = s + n$. We need to show that $s, n \in \mathfrak{g}$ and, of course, it suffices to show that $n \in \mathfrak{g}$. To this end, we will describe $\mathfrak{g}$ as an intersection of Lie subalgebras of $\text{gl}(V)$ that all contain $n$.

First, let $n = \{ y \in \text{gl}(V) \mid \text{ad} y.\mathfrak{g} \subseteq \mathfrak{g} \}$, the normalizer of $\mathfrak{g}$ in $\text{gl}(V)$. Invoking Proposition 5.19 and Lemma 5.20, we obtain $\text{ad} n = p(\text{ad} x)$ for some polynomial $p(t) \in k[t]$. Since $\text{ad} x$ stabilizes $\mathfrak{g}$, so does $\text{ad} n$. Therefore, $n \in n$.

Next, put $s_W = \{ z \in \text{gl}(V) \mid z.W \subseteq W \text{ and } \text{trace } z|_W = 0 \}$ for any subrepresentation $W \subseteq V$. Note that $\mathfrak{g} \subseteq s_W$, the vanishing of trace $\mathfrak{g}|_W$ being a consequence of the fact that $\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}]$ (Corollary 6.4). Since $x.W \subseteq W$ and $n$ is a polynomial in $x$ by Proposition 5.19, we also have $n.W \subseteq W$. Moreover, since $n$ is nilpotent, the trace of $n|_W$ vanishes. Thus, $n \in s_W$.

Finally, put $\mathfrak{s} := n \cap \bigcap_W s_W$, where $W$ runs over the irreducible subrepresentations of $V$. The claim will follow if we can show that $\mathfrak{g} = \mathfrak{s}$. But $\mathfrak{g}$ is a Lie subalgebra of the Lie algebra $\mathfrak{s}$, and so we consider the adjoint action of $\mathfrak{g}$ on $\mathfrak{s}$. Now Weyl's Theorem tells us that $\mathfrak{s} = \mathfrak{g} \oplus \mathfrak{c}$ for some $k$-subspace $\mathfrak{c}$ with $[\mathfrak{g}, \mathfrak{c}] \subseteq \mathfrak{c}$. Since $\mathfrak{c} \subseteq n$, we have $[\mathfrak{g}, \mathfrak{c}] \subseteq c \cap n = 0$. Schur's Lemma therefore implies that all elements of $\mathfrak{c}$ act by scalars on each irreducible $W$; these scalars must be 0 by the condition $\text{trace } z|_W = 0$ for $z \in \mathfrak{c}$. Inasmuch as $V$ is the sum of the various $W$ by Weyl's Theorem, it follows that $\mathfrak{c} = 0$, proving the claim.

It is now a simple matter to prove (a) and (b). For (a), apply the claim to the adjoint representation $\text{ad}: \mathfrak{g} \hookrightarrow \text{gl}(\mathfrak{g})$. It follows that if $\text{ad} x = s + n$ is the Jordan decomposition of $\text{ad} x \in \text{gl}(\mathfrak{g})$ for $x \in \mathfrak{g}$, then $s = \text{ad} x_s$ and $n = \text{ad} x_n$ for unique $x_s, x_n \in \mathfrak{g}$. Since $0 = [s, n] = \text{ad}[x_s, x_n]$, we also obtain $[x_s, x_n] = 0$. This proves (a). Note also that, for any epimorphism $f: \mathfrak{g} \twoheadrightarrow \mathfrak{b}$ in $\text{Lie}_k$, the abstract Jordan decomposition of $f(x) \in \mathfrak{b}$ is given by $f(x) = f(x_s) + f(x_n)$. Indeed, $[f(x_s), f(x_n)] = f([x_s, x_n]) = 0$ and the operators $\text{ad} f(x_s), \text{ad} f(x_n) \in \text{gl}(\mathfrak{b})$ are clearly diagonalizable and nilpotent, respectively.

Now for (b). By the foregoing, $x_V = (x_s)_V + (x_n)_V$ is the abstract Jordan decomposition of $x_V \in \mathfrak{g}_V$. On the other hand, if $x_V = s + n$ is the ordinary Jordan decomposition in $\text{gl}(V)$, then $[s, n] = 0$ and $s, n \in \mathfrak{g}_V$ by the claim. Moreover,
Lemma 5.20 tells us that the operators $\text{ad} s, \text{ad} n$ on $\mathfrak{gl}(V)$ are diagonalizable and nilpotent, respectively, and hence the same holds for their restrictions to $\mathfrak{g}_V \subseteq \mathfrak{gl}(\mathfrak{g})$. In other words, $x_V = s + n$ is the abstract Jordan decomposition of $x_V \in \mathfrak{g}_V$. By uniqueness, we must have $s = (x_s)_V$ and $n = (x_n)_V$, which completes the proof of the proposition. □

Exercises for Section 6.2

6.2.1 (Completely reducible representations). Let $\mathfrak{g} \in \text{Lie}_k$ be arbitrary and let $V \in \text{Rep} \mathfrak{g}$ be completely reducible. Show that $V = V^0 \oplus \mathfrak{g}.V$.

6.2.2 (Reductive Lie algebras). Assuming $\text{char} k = 0$ and $\mathfrak{g} \in \text{Lie}_k$ to be finite dimensional, show:
   
   (a) $\mathfrak{g}$ is reductive if and only if the adjoint representation $\mathfrak{g}_{\text{ad}}$ is completely reducible.
   
   (b) $\mathfrak{g}$ is reductive if and only if there exists $V \in \text{Rep}_{\text{fin}} \mathfrak{g}$ that is $\mathfrak{g}$-faithful and completely reducible
   
   (c) $\mathfrak{g}$ is reductive if and only if $\mathfrak{g}$ is isomorphic to a direct product of a finite-dimensional abelian and a semisimple Lie algebra.
   
   (d) Show that $\mathfrak{gl}_n$ is reductive and find the decomposition (6.9).

6.2.3 (Center-valued trace). Let $\mathfrak{g} \in \text{Lie}_k$ be reductive and let $U = \mathfrak{U}_\mathfrak{g}$. Show:
   
   (a) $U = \mathcal{Z} \oplus [U, U]$, where $\mathcal{Z} = \mathcal{Z}U$ and $[U, U] = \langle uv - vu \mid u, v \in U \rangle_k$.
   
   (Use Exercises 6.2.1 and 5.5.4.)
   
   (b) The projection $\tau: U \to \mathcal{Z}$ along $[U, U]$ is a $\mathcal{Z}$-linear map such that $\tau(uv) = \tau(vu)$ for all $u, v \in U$.

6.3. Cartan Subalgebras and the Root Space Decomposition

Our aim in this section is to generalize the familiar decomposition of $\mathfrak{sl}_2$ to arbitrary semisimple Lie algebras. Recall from Section 5.7 that

\[ \mathfrak{sl}_2 = k f \oplus k h \oplus k e = \mathfrak{g}_{-2} \oplus \mathfrak{g}_0 \oplus \mathfrak{g}_2 \]

with $\mathfrak{g}_i = \{ x \in \mathfrak{sl}_2 \mid [h, x] = \lambda x \}$; this is called the root space decomposition of $\mathfrak{sl}_2$ for the ad-semisimple element $h \in \mathfrak{sl}_2$. For a general semisimple Lie algebra $\mathfrak{g}$, the role of $k h \subseteq \mathfrak{sl}_2$, will be played by a so-called Cartan subalgebra of $\mathfrak{g}$.

Throughout this section, $\mathfrak{g}$ denotes a semisimple Lie $k$-algebra. The base field $k$ is understood to be algebraically closed and to have characteristic 0.
6.3.1. Cartan Subalgebras

A Lie subalgebra \( h \subseteq g \) is called a Cartan subalgebra if all elements of \( h \) are ad-semisimple, that is, \( \text{ad}_h x \) is diagonalizable for all \( x \in h \), and \( h \) is maximal with respect to this property.

**Existence.** If \( g \neq 0 \), then there are ad-semisimple elements \( 0 \neq x \in g \). Indeed, for any \( x \in g \), the abstract Jordan decomposition \( x = x_s + x_n \) gives an ad-semisimple \( x_s \in g \) (Proposition 6.8). Moreover, some \( x_s \) must be nonzero, because otherwise \( g \) would be nilpotent (Corollary 5.16). Thus, \( x_s \) is Lie subalgebra of \( g \) consisting of ad-semisimple elements, and it is certainly contained in a maximal one subject to this property. Therefore, a Cartan subalgebra \( h \subseteq g \) does exist and \( h \neq 0 \) if \( g \neq 0 \).

It is much less obvious, but true, that any two Cartan subalgebras of \( g \) are conjugate under the group of all Lie algebra automorphisms of \( g \) (even under the subgroup of all elementary automorphism of \( g \); see §8.6.2). For the proof of this fact, we refer to Bourbaki [28] or Humphreys [101]. We will prove later that the representation ring \( \mathcal{R}(g) \) is a polynomial ring in \( \dim \mathfrak{g} \) many variables over \( \mathbb{Z} \) (Theorem 8.15). Thus, \( \dim \mathfrak{g} \) is an invariant of \( g \); it is called the rank of \( g \).

**Weight Spaces.** Let us now fix a Cartan subalgebra \( h \subseteq g \). Our first goal will be to show that, in analogy with the situation for \( \mathfrak{s}\mathfrak{l}_2 \), the Lie algebra \( g \) is the direct sum of \( h \)-weight spaces

\[
\mathfrak{g}_\alpha = \{ x \in g \mid [z, x] = (\alpha, z)x \text{ for all } z \in h \} \quad (\alpha \in h^*)
\]

Note that \( \mathfrak{g}_0 = C_g(h) \), the centralizer of \( h \) in \( g \). In general, for any \( \alpha, \beta \in h^* \),

\[
[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subseteq \mathfrak{g}_{\alpha + \beta}.
\]

This follows from the following computation, for \( x \in \mathfrak{g}_\alpha \), \( y \in \mathfrak{g}_\beta \) and \( z \in h \):

\[
[z, [x, y]] = [[z, x], y] + [x, [z, y]] = \langle \alpha, z \rangle x y + x, \langle \beta, z \rangle y = \langle \alpha + \beta, z \rangle [x, y].
\]

Furthermore, if \( \alpha + \beta \neq 0 \), then \( \mathfrak{g}_\alpha \) and \( \mathfrak{g}_\beta \) are orthogonal to each other for the Killing form \( B = B_g \). To see this, fix \( z \in h \) such that \( \langle \alpha + \beta, z \rangle \neq 0 \) and let \( x \in \mathfrak{g}_\alpha \), \( y \in \mathfrak{g}_\beta \). Then

\[
\langle \alpha + \beta, z \rangle B(x, y) = B([z, x], y) + B(x, [z, y]) = B([z, x], y) + B([x, z], y) = 0,
\]

whence \( B(x, y) = 0 \). Thus, for the record,

\[
\alpha + \beta \neq 0 \quad \implies \quad B(\mathfrak{g}_\alpha, \mathfrak{g}_\beta) = 0.
\]

**Main Properties of Cartan Subalgebras.** Part (a) of the following theorem implies in particular that Cartan subalgebras can also be characterized as those Lie subalgebras \( h \subseteq g \) that consist of ad-semisimple elements on \( g \) and are maximal abelian Lie subalgebra of \( g \) or, equivalently, self-centralizing, that is, \( C_g(h) = h \).

**Theorem 6.9.** Let \( h \subseteq g \) be a Cartan subalgebra. Then:

(a) \( h = \mathfrak{g}_0 \); so \( h \) is a maximal abelian Lie subalgebra of \( g \).
We first show that \( \mathfrak{h} \) is abelian: \( \text{ad}_x x = 0 \) for all \( x \in \mathfrak{h} \). We certainly know that \( \text{ad}_x x \) is diagonalizable, because \( \text{ad}_x x \) is diagonalizable and stabilizes \( \mathfrak{h} \). Thus it suffices to show that every \( \text{ad}_x x \)-eigenvector \( y \in \mathfrak{h} \) satisfies \([x, y] = 0\). But \( \text{ad}_x y \) is also diagonalizable, say \( \mathfrak{h} = \bigoplus \mathbb{K} z_i \) with \([y, z_i] = \lambda_i z_i \) for \( \lambda_i \in \mathbb{K} \). Furthermore, \([x, y] \in \mathbb{K} y \subseteq C_0(y) = \bigoplus_{t: \lambda_i = 0} \mathbb{K} z_i \). On the other hand, writing \( x = \sum_i \xi_i z_i \) with \( \xi_i \in \mathbb{K} \), we have \([y, x] = \sum_i \xi_i \lambda_i z_i \). Therefore, we must have \( \xi_i = 0 \) if \( \lambda_i \neq 0 \). So \( x \in C_0(y) \) as needed. This proves that \( \mathfrak{h} \) is abelian or, equivalently, \( \mathfrak{h} \subseteq \mathfrak{g}_0 \).

Since commuting diagonalizable endomorphisms are simultaneously diagonalizable, we conclude that all \( \text{ad}_h z \) with \( z \in \mathfrak{h} \) are simultaneously diagonalizable, which is exactly what part (b) of the theorem states. Next, we make the following

**Claim.** \( a := \{ x \in \mathfrak{g}_0 | B(x, h) = 0 \} = 0 \).

This will prove (c) and it will also imply that \( \mathfrak{g}_0 \) embeds into \( \mathfrak{h}^* \). Since we already know that \( \mathfrak{h} \subseteq \mathfrak{g}_0 \), it will follow that \( \mathfrak{g}_0 = \mathfrak{h} \) for dimension reasons. Thus, the claim will also finish the proof of (a).

In order to prove the claim, consider an arbitrary element \( x \in \mathfrak{g}_0 \) and let \( x = x_s + x_n \) be the abstract Jordan decomposition of \( x \), viewed as an element of \( \mathfrak{g} \). We first show that \( x_s \in \mathfrak{h} \). Indeed, writing \( \text{ad} = \text{ad}_x \) for brevity, we know that \( \text{ad} x \in \mathfrak{g}(\mathfrak{g}) \) stabilizes all \( \mathfrak{h} \)-weight spaces \( \mathfrak{g}_\alpha \), because \([\mathfrak{g}_0, \mathfrak{g}_0] \subseteq \mathfrak{g}_\alpha \) by (6.10). Moreover, \( \text{ad} x_s \) is the semisimple part of the ordinary Jordan decomposition of \( \text{ad} x \), and hence \( \text{ad} x_s \) is a polynomial in \( \text{ad} x \) (Proposition 5.19). Consequently, \( \text{ad} x_s \) also stabilizes all \( \mathfrak{g}_\alpha \), acting as a diagonalizable operator on each of them. Since \( (\text{ad} x)(\mathfrak{h}) = 0 \), it also follows that \((\text{ad} x_s)(\mathfrak{h}) \subseteq \mathfrak{h} \). Thus, \( \mathfrak{h} + \mathbb{K} x_s \) is a Lie subalgebra of \( \mathfrak{g} \) whose elements are are ad-semisimple on \( \mathfrak{g} = \bigoplus \mathfrak{g}_\alpha \). By maximality of \( \mathfrak{h} \), we must have \( x_s \in \mathfrak{h} \) as claimed.

It now follows that \([x_s, \mathfrak{g}_0] \in [\mathfrak{h}, \mathfrak{g}_0] = 0 \) and so \( \text{ad}_x x = \text{ad}_{\mathfrak{g}_0} x_n \) is nilpotent. Since \( x \in \mathfrak{g}_0 \) was arbitrary, we conclude that the Lie algebra \( \mathfrak{g}_0 \) is nilpotent (Corollary 5.16). Note that \( a \) is an ideal of \( \mathfrak{g}_0 \). Suppose, for a contradiction, that \( a \neq 0 \).

Then there exists an element \( 0 \neq c \in a \cap \mathfrak{g}_0 \) (Proposition 5.13). We claim that \( B(c, x) = 0 \) for all \( x \in \mathfrak{g}_0 \). To see this, consider the abstract Jordan decomposition \( x = x_s + x_n \) as above. Then \( B(c, x_s) \in B(a, \mathfrak{h}) = 0 \) and so

\[
B(c, x) = B(c, x_s) + B(c, x_n) = B(c, x_n) = \text{trace}(\text{ad} c \text{ad} x_n) = 0.
\]

For the last equality, observe that \( \text{ad} c \text{ad} x_n \) is nilpotent, because \( c \in \mathfrak{g}_0 \) and \( x_n \in \mathfrak{g}_0 \); so \( \text{ad} c \) and \( \text{ad} x_n \) are commuting endomorphisms, with \( \text{ad} x_n \) being nilpotent. This proves that \( B(c, \mathfrak{g}_0) = 0 \). Since we already know by (6.11) that \( B(c, \mathfrak{g}_0) = 0 \) for all \( c \neq 0 \) it follows from (b) that \( B(c, \mathfrak{g}) = 0 \), contradicting non-degeneracy of the Killing form \( B \). Therefore, we must have \( a = 0 \) as claimed. \( \square \)
6.3.2. Root Space Decomposition

For a given Cartan subalgebra \( \mathfrak{h} \subseteq \mathfrak{g} \), we write the decomposition of \( \mathfrak{g} \) in Theorem 6.9 in the form

\[
\mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{\alpha \in \Phi} \mathfrak{g}_\alpha \quad \text{with} \quad \Phi \overset{\text{def}}{=} \{ \alpha \in \mathfrak{h}^* \setminus \{0\} \mid \mathfrak{g}_\alpha \neq 0 \}
\]

The elements \( \alpha \in \Phi \) are called the roots of \( \mathfrak{g} \) (for the given Cartan subalgebra) and the weight spaces \( \mathfrak{g}_\alpha \) are called the root spaces of \( \mathfrak{g} \).

The next theorem collects the main properties of \( \Phi \) and the decomposition (6.12). The proof will make significant use of the structure of finite-dimensional representations of \( \mathfrak{sl}_2 \) (Theorem 5.39). We will also use the fact that Theorem 6.9(c) allows us to identify \( \mathfrak{h} \) with its dual:

\[
\mathfrak{h} \overset{\sim}{\longrightarrow} \mathfrak{h}^*
\]

We let \( t_\alpha \in \mathfrak{h} \) denote the element corresponding to \( \alpha \in \mathfrak{h}^* \) under this isomorphism; so \( t_\alpha \) depends linearly on \( \alpha \) and is characterized by the condition

\[
\alpha = B(t_\alpha, \cdot) \quad \in \mathfrak{h}.
\]

**Theorem 6.10.** Let \( \mathfrak{h} \subseteq \mathfrak{g} \) be a Cartan subalgebra and let \( \Phi \subseteq \mathfrak{h}^* \setminus \{0\} \) be the corresponding set of roots. Then:

(a) **Finiteness, spanning property, and multiples.** \( \Phi \) is finite and spans \( \mathfrak{h}^* \). Furthermore, for each \( \alpha \in \Phi \), we have

\[
\mathbb{Z} \alpha \cap \Phi = \{ \pm \alpha \}.
\]

(b) **\( \mathfrak{sl}_2 \)-triples.** Let \( \alpha \in \Phi \). Then \( \dim \mathfrak{g}_\alpha = 1 \) and \( [\mathfrak{g}_\alpha, \mathfrak{g}_{-\alpha}] = \mathbb{Z} h_\alpha \), where

\[
h_\alpha := \frac{2t_\alpha}{B(t_\alpha, t_\alpha)} = \frac{2t_\alpha}{\langle \alpha, t_\alpha \rangle} \quad \in \mathfrak{h}.
\]

Moreover, \( s_\alpha := \mathfrak{g}_{-\alpha} \oplus \mathbb{Z} h_\alpha \oplus \mathfrak{g}_\alpha \) is a Lie subalgebra of \( \mathfrak{g} \) such that \( s_\alpha \cong \mathfrak{sl}_2 \), with \( h_\alpha \leftrightarrow h \in \mathfrak{sl}_2 \).

(c) **Cartan integers and root strings.** For all \( \alpha, \beta \in \Phi \),

\[
\langle \beta, h_\alpha \rangle \in \mathbb{Z} \quad \text{and} \quad \beta - \langle \beta, h_\alpha \rangle \alpha \in \Phi.
\]

If \( \beta \neq \pm \alpha \), then \( [\mathfrak{g}_\alpha, \mathfrak{g}_\beta] = \mathfrak{g}_{\alpha + \beta} \) and \( \langle \beta, h_\alpha \rangle = r - s \), where \( r, s \in \mathbb{Z}_+ \) are chosen maximal such that \( \beta - r \alpha \in \Phi \) and \( \beta + s \alpha \in \Phi \). Moreover, all \( \beta + i \alpha \) with \( -r \leq i \leq s \) belong to \( \Phi \) (the “\( \alpha \)-string through \( \beta \)”).
6.3. Cartan Subalgebras and the Root Space Decomposition

**Proof.** (a) Since \( g \) is finite dimensional and \( g_\alpha \neq 0 \) for \( \alpha \in \Phi \), it follows from (6.12) that \( \Phi \) must be finite. The fact that \( \Phi \) spans \( \mathfrak{h}^\perp \) is a consequence of the observation that if \( z \in \mathfrak{h} \) satisfies \( \langle \alpha, z \rangle = 0 \) for all \( \alpha \in \Phi \), then \( z \in Z(g) \) by (6.12), and so \( z = 0 \). Finally, if \( \alpha \in \Phi \) then \( B(g_\alpha, g_\beta) \neq 0 \) by non-degeneracy of the Killing form, while \( B(g_\alpha, g_\beta) = 0 \) for all \( \beta \neq -\alpha \) by (6.11). Therefore, (6.12) implies that

\[
B(g_\alpha, g_-\alpha) \neq 0.
\]

It follows that \( g_-\alpha \neq 0 \) and so \( -\alpha \in \Phi \). The fact that \( \pm \alpha \) are the only multiples of \( \alpha \) that belong to \( \Phi \) will be proved below.

(b) We first show that, for any \( \alpha \in \Phi \),

\[
x \in g_\alpha, y \in g_-\alpha \implies [x, y] = B(x, y)t_\alpha.
\]

Indeed, as functions on \( \mathfrak{h} \), we have \( B([x, y], \cdot) = B(x, [y, \cdot]) = \alpha B(x, y) \). Since \( [x, y] \in \mathfrak{g}_0 = \mathfrak{h} \) by (6.10) and Theorem 6.9(a), the latter equation can be stated as \( [x, y] = t_B(x, y)\alpha = B(x, y)t_\alpha \), proving (6.16). From (6.15) and (6.16) we deduce in particular that \( [g_-\alpha, g_\alpha] = \mathbb{K} t_\alpha \).

Next, we show that, for any \( \alpha \in \Phi \),

\[
B(t_\alpha, t_\alpha) \neq 0.
\]

Suppose (6.17) fails. Then \([t_\alpha, g_\alpha] = \langle \alpha, t_\alpha \rangle g_\alpha = B(t_\alpha, t_\alpha)g_\alpha = 0 \). Since \( g_-\alpha g_\alpha = \mathbb{K} t_\alpha \), we may choose elements \( x \in g_\alpha, y \in g_-\alpha \) with \([x, y] = t_\alpha \). Putting \( \tau = ad(g_\alpha t_\alpha), \xi = ad(x) \) and \( \eta = ad(y) \), we obtain \([\xi, \eta] = \tau \) and \([\tau, \xi] = 0 \), because \([t_\alpha, x] = 0 \). It follows that \( \tau^{m+1} = [\xi, \eta \tau^m] \) for all \( m \geq 0 \) and so trace \( \tau^{m+1} = 0 \). By the Newton formulas (3.56), this implies that the operator \( \tau \) is nilpotent. On the other hand, \( \tau \) is diagonalizable, because \( t_\alpha \in \mathfrak{h} \). Therefore, we must have \( \tau = 0 \), contradicting the fact that \( t_\alpha \neq 0 \) and \( ad(g_\alpha) \) is mono. This proves (6.17).

Now we may define \( h_\alpha := \frac{2t_\alpha}{B(t_\alpha, t_\alpha)} \in \mathfrak{h} \). Then \([g_-\alpha, g_\alpha] = \mathbb{K} h_\alpha \) and

\[
\langle \alpha, h_\alpha \rangle = B(t_\alpha, h_\alpha) = 2.
\]

Choose \( e_\alpha \in g_\alpha, f_\alpha \in g_-\alpha \) such that \([e_\alpha, f_\alpha] = h_\alpha \). Then \([h_\alpha, e_\alpha] = \langle \alpha, h_\alpha \rangle e_\alpha = 2e_\alpha \) and similarly \([h_\alpha, f_\alpha] = -2f_\alpha \). Thus,

\[
s_\alpha := \mathbb{K} f_\alpha \oplus \mathbb{K} h_\alpha \oplus \mathbb{K} e_\alpha
\]

is a Lie subalgebra of \( g \) that is isomorphic to \( sl_2 \) via \( f_\alpha \leftrightarrow f, e_\alpha \leftrightarrow e \) and \( h_\alpha \leftrightarrow h \).

To finish the proof of (b), we still need to show that \( \dim g_\alpha = 1 \) for all \( \alpha \in \Phi \). Fix \( \alpha \in \Phi \) and consider the following subspace of \( g \):

\[
V := \bigoplus_{\alpha' \in \mathbb{K} \Phi} g_{\alpha'} + \bigoplus_{\lambda \alpha \in \Phi} g_{\lambda \alpha}.
\]

By (6.10), \( V \) is an \( s_\alpha \)-subrepresentation of \( g_{ad + s_\alpha} \). The following claim will imply that \( \dim g_\alpha = 1 \) for all \( \alpha \in \Phi \) and also that \( \mathbb{K} \alpha \cap \Phi = \{ \pm \alpha \} \), thereby completing the proofs of (a) and (b).
Claim. \( V = \mathfrak{b} + \mathfrak{s}_\alpha \).

Clearly, \( U := \mathfrak{b} + \mathfrak{s}_\alpha = \mathfrak{b} \oplus \mathbb{K} f_\alpha \oplus \mathbb{K} e_\alpha \) is an \( \mathfrak{s}_\alpha \)-subrepresentation of \( V \). In order to prove the desired equality \( U = V \), recall that \( \mathfrak{s}_\alpha \cong \mathfrak{s} \mathfrak{l}_2 \) with \( h_\alpha \leftrightarrow h \). So the structure theory of finite-dimensional \( \mathfrak{s} \mathfrak{l}_2 \)-representations (Theorem 5.39 and Proposition 5.37) tells us that \( V \) is completely reducible and each irreducible constituent of \( V \) has the form \( V(m) \) for some \( m \in \mathbb{Z}_+ \), with \( h_\alpha \)-weights \( \{-m, -m+2, \ldots, m-2, m\} \). On the other hand, by definition of \( V \) and (6.18), the \( h_\alpha \)-weight spaces \( V_i = \{ v \in V \mid [h_\alpha, v] = iv \} \) are as follows: \( V_0 = \mathfrak{b} \) and \( V_{2i} = \mathfrak{g}_{i\alpha} \cdot \). Only the weights 0 and \( \pm 2 \) (from \( \lambda = \pm 1 \)) occur in \( U \). Since \( V_0 = \mathfrak{b} \subseteq U \), it follows that \( U \) has nonzero intersection with each irreducible constituent \( V(m) \) such that \( m \) is even, and hence \( U \) contains all these constituents. Therefore, the only even weights of \( V \) are 0 and \( \pm 2 \). In particular, \( 2\alpha \notin \Phi \), because otherwise \( V_4 \neq 0 \). Since \( \alpha \in \Phi \) was arbitrary, we have shown that twice a root is never a root. Consequently, \( \alpha \in \Phi \) forces \( \frac{1}{2} \alpha \notin \Phi \). Hence 1 is not a weight of \( V \), and so \( V \) has no irreducible constituents \( V(m) \) with \( m \) odd. Therefore, we must have \( V = U \) as claimed.

We already know that \( \langle \pm \alpha, h_\alpha \rangle = \pm 2 \) and \( \pm \alpha - \langle \pm \alpha, h_\alpha \rangle \alpha = \mp \alpha \in \Phi \). So let us assume that \( \beta \neq \pm \alpha \). Consider the following subspace of \( \mathfrak{g} \), with \( r, s \in \mathbb{Z}_+ \), as in the statement of (c):

\[
W := \bigoplus_{i \in \mathbb{Z}} \mathfrak{g}_{\beta + i\alpha} = \bigoplus_{-r \leq i \leq s, \beta + i\alpha \in \Phi} \mathfrak{g}_{\beta + i\alpha}.
\]

Again, it follows from (6.10) that \( W \) is a subrepresentation of \( \mathfrak{g} \rightarrow \mathfrak{s}_\alpha \cdot \). The \( h_\alpha \)-weight spaces of \( W \) are the various \( \mathfrak{g}_{\beta + i\alpha} \) for \( \beta + i\alpha \in \Phi \) \((-r \leq i \leq s\), with corresponding weights

\[
\langle \beta + i\alpha, h_\alpha \rangle = \langle \beta, h_\alpha \rangle + 2i.
\]

By Theorem 5.39, these weights must be integers; so we obtain \( \langle \beta, h_\alpha \rangle \in \mathbb{Z} \). We also see that all weights of \( W \) have the same parity, the smallest one being \( \langle \beta, h_\alpha \rangle - 2r \) and the largest \( \langle \beta, h_\alpha \rangle + 2s \), and we know that the weight spaces \( \mathfrak{g}_{\beta + i\alpha} \) of \( W \) are 1-dimensional by (b). Thus, Theorem 5.39 tells us that \( W \) is irreducible, say \( W \cong V(m) \) with \( m \in \mathbb{Z}_+ \); so the weights of \( W \) are \( \{-m, -m+2, \ldots, m-2, m\} \) (Proposition 5.37). Therefore, \( \langle \beta, h_\alpha \rangle - 2r = -m \) and \( \langle \beta, h_\alpha \rangle + 2s = m \), and so \( \langle \beta, h_\alpha \rangle = r - s \). Moreover, the entire string \( \{ \beta + i\alpha \mid -r \leq i \leq s \} \) must belong to \( \Phi \). For \( i = s - r \), we obtain in particular that \( \beta - \langle \beta, h_\alpha \rangle \alpha \in \Phi \). Finally, the equality \( \mathfrak{g}_\alpha = \mathfrak{s}_\alpha \alpha + \beta \) follows from the fact that the action of \( e_\alpha \) maps the weight space \( \mathfrak{g}_\beta = W(\beta, h_\alpha) \) onto \( \mathfrak{g}_{\beta + \alpha} = W(\beta, h_\alpha) + 2 \) (Proposition 5.37). This completes the proof of the theorem. \( \square \)
6.3.3. Simplicity

Let $\mathfrak{h}$ be a fixed Cartan subalgebra of $\mathfrak{g}$. We transport the Killing form $B|_{\mathfrak{h} \times \mathfrak{h}}$ to $\mathfrak{h}^* \times \mathfrak{h}^*$ using the identification $\mathfrak{h}^* \equiv \mathfrak{g}$, $\alpha \mapsto t_\alpha$, from (6.13) and (6.14):

(6.19) \[ (\mu, \nu) := B(t_\mu, t_\nu) \quad (\mu, \nu \in \mathfrak{h}^*). \]

So $(\mu, \nu) = \langle \mu, t_\nu \rangle = \langle \nu, t_\mu \rangle$. In view of Theorem 6.9, the space $\mathfrak{h}^*$ is thus equipped with a non-degenerate symmetric bilinear form.

Now let $\Phi \subseteq \mathfrak{h}^*$ be the set of roots as in (6.12). We will say that $\Phi$ is \textit{irreducible} if $\Phi \neq \emptyset$ (or, equivalently, $\mathfrak{g} \neq 0$) and it is not possible to write $\Phi$ as a disjoint union $\Phi = \Phi_1 \cup \Phi_2$ of nonempty subsets $\Phi_i$ that are orthogonal to each other in the sense that $(\Phi_1, \Phi_2) = \{0\}$.

**Proposition 6.11.** Let $\mathfrak{h} \subseteq \mathfrak{g}$ be a Cartan subalgebra and let $\Phi$ be the corresponding set of roots. Then $\mathfrak{g}$ is simple if and only if $\Phi$ is irreducible.

**Proof.** First, assume that $\mathfrak{g}$ is not simple. Then $\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{b}$ for nonzero ideals $\mathfrak{a}$ and $\mathfrak{b}$. Since $\mathfrak{a}$ is stable under the adjoint action of $\mathfrak{h}$, the root space decomposition (6.12) together with the fact that $\dim_\mathbb{C} \mathfrak{g}_\alpha = 1$ for $\alpha \in \Phi$ (Theorem 6.10) imply that

\[ a = (a \cap \mathfrak{h}) \oplus \bigoplus_{\alpha \in \Phi} (a \cap \mathfrak{g}_\alpha) = (a \cap \mathfrak{h}) \oplus \bigoplus_{\alpha \in \Phi} \mathfrak{g}_\alpha \]

with $\Phi_a = \{ \alpha \in \Phi \mid a \cap \mathfrak{g}_\alpha \neq 0 \} = \{ \alpha \in \Phi \mid a \supseteq \mathfrak{g}_\alpha \}$. Similarly, $b = (b \cap \mathfrak{h}) \oplus \bigoplus_{\alpha \in \Phi} \mathfrak{g}_\alpha$. Since $a \cap b = 0$ and $a + b = \mathfrak{g}$, we certainly have $\Phi_a \cap \Phi_b = \emptyset$ and $\Phi_a \cup \Phi_b = \Phi$. Furthermore, $\Phi_a \neq \emptyset$, because otherwise $a \subseteq \mathfrak{h}$ and so $a$ would be a nonzero abelian ideal of $\mathfrak{g}$, which does not exist. Likewise, $\Phi_b \neq \emptyset$. Finally, if $\alpha \in \Phi_a$ and $\beta \in \Phi_b$, then $\mathfrak{g}_\alpha \subseteq a$ and $t_\beta \in [\mathfrak{g}_\beta, \mathfrak{g}_\alpha] \subseteq b$ (Theorem 6.10). Thus, $\langle \alpha, t_\beta \rangle \mathfrak{g}_\alpha = [t_\beta, \mathfrak{g}_\alpha] \subseteq a \cap b = 0$, whence $0 = \langle \alpha, t_\beta \rangle = (\alpha, \beta)$. This shows that $\Phi$ is not irreducible.

Conversely, assume that $\Phi = \Phi_1 \cup \Phi_2$ for nonempty orthogonal subsets $\Phi_i$. Put $\mathfrak{t} := \{ z \in \mathfrak{h} \mid \langle z, \alpha \rangle = 0 \text{ for all } \alpha \in \Phi_2 \}$ and $\mathfrak{a} := \mathfrak{t} \oplus \bigoplus_{\alpha \in \Phi_1} \mathfrak{g}_\alpha$; this is a proper nonzero subspace of $\mathfrak{g}$. We will show that $\mathfrak{a}$ is in fact an ideal of $\mathfrak{g}$, whence $\mathfrak{a}$ is not simple. Evidently, $\mathfrak{a}$ is stable under $\mathfrak{ad} \mathfrak{h}$. Let $\alpha_i \in \Phi_i$ ($i = 1, 2$). Then $(\alpha_1 + \alpha_2, \alpha_i) = (\alpha_i, \alpha_i) \neq 0$ by (6.17). Therefore, $0 \neq \alpha_1 + \alpha_2 \notin \Phi$ and so $[\mathfrak{g}_{\alpha_1}, \mathfrak{g}_{\alpha_2}] \subseteq \mathfrak{g}_{\alpha_1 + \alpha_2} = 0$. Furthermore, $[\mathfrak{t}, \mathfrak{g}_{\alpha_1}] = \langle \alpha_2, \mathfrak{t} \rangle \mathfrak{g}_{\alpha_2} = 0$. This shows that $\mathfrak{ad} \mathfrak{g}_{\alpha_1}$ annihilates $\mathfrak{a}$. Finally,

\[ [\mathfrak{a}, \mathfrak{g}_{\alpha_1}] = [\mathfrak{t}, \mathfrak{g}_{\alpha_1}] + \sum_{\alpha \in \Phi_1} [\mathfrak{g}_\alpha, \mathfrak{g}_{\alpha_1}] \subseteq \mathfrak{g}_{\alpha_1} + \sum_{\alpha \in \Phi_1 \setminus \{-\alpha_1\}} \mathfrak{g}_{\alpha + \alpha_1} + \mathfrak{z} \tau_{\alpha_1}, \]

because $[\mathfrak{g}_{-\alpha_1}, \mathfrak{g}_{\alpha_1}] = \mathfrak{z} \tau_{\alpha_1}$ (Theorem 6.10). Here, $\mathfrak{g}_{\alpha_1} \subseteq a$ and $t_{\alpha_1} \in \mathfrak{t}$, because $\langle \alpha_2, t_{\alpha_1} \rangle = (\alpha_2, \alpha_1) = 0$ for all $\alpha_2 \in \Phi_2$. Moreover, if $\mathfrak{g}_{\alpha + \alpha_1} \neq 0$ for $\alpha \in \Phi_1 \setminus \{-\alpha_1\}$, then we must have $\alpha + \alpha_1 \in \Phi_1$, because $(\alpha + \alpha_1, \Phi_2) = \{0\}$. Therefore, $\mathfrak{a}$ is also stable under $\mathfrak{ad} \mathfrak{g}_{\alpha_1}$. Hence $\mathfrak{a}$ is an ideal of $\mathfrak{g}$ as was to be shown. \qed
6.3.4. Embedding into Euclidean Space

Let \( \mathfrak{h} \) be a fixed Cartan subalgebra of \( \mathfrak{g} \). In the following, we will write
\[
n = \dim_{\mathbb{R}} \mathfrak{h}.
\]
As we have remarked earlier, this is an invariant of \( \mathfrak{g} \), called the rank of \( \mathfrak{g} \). Our goal in this subsection is to replace \( \mathfrak{h}^* \cong \mathbb{R}^n \) by a real vector space \( \mathbb{E} \cong \mathbb{R}^n \) in such a way that the form \( (\ , \ ) \) in (6.19) becomes an inner product on \( \mathbb{E} \), that is, a positive definite symmetric bilinear form \( \mathbb{E} \times \mathbb{E} \to \mathbb{R} \). Specifically, we put \( \mathbb{E} := \mathbb{Q}(\Phi) \otimes_{\mathbb{Q}} \mathbb{R} \).

**Proposition 6.12.** The restriction of the form \( (\ , \ ) \) to \( \mathbb{Q}(\Phi) \subseteq \mathfrak{h}^* \) is \( \mathbb{Q} \)-valued. Extending \( (\ , \ ) \) from \( \mathbb{Q}(\Phi) \to \mathbb{E} \) by \( \mathbb{R} \)-bilinearity, we obtain an inner product on \( \mathbb{E} \). Moreover, \( \dim_{\mathbb{R}} \mathbb{E} = n \).

**Proof.** We first derive a useful expression for the value of \( (\mu, \nu) \) with \( \mu, \nu \in \mathfrak{h}^* \). For a basis of \( \mathfrak{g} \) chosen according to the root space decomposition (6.12), the matrix of \( \text{ad} z \) for \( z \in \mathfrak{h} \) is diagonal, with \( n \) entries equal to 0 and one entry \( \langle \alpha, z \rangle \) for each \( \alpha \in \Phi \). Therefore, \( (\mu, \nu) = \text{trace}(\text{ad} \mu, \text{ad} \nu) = \sum_{\alpha \in \Phi} \langle \alpha, t_{\mu} \rangle \langle \alpha, t_{\nu} \rangle \). Since \( \langle \alpha, t_{\mu} \rangle = (\alpha, \mu) \) by (6.14), we obtain
\[
(\mu, \nu) = \sum_{\alpha \in \Phi} (\alpha, \mu)(\alpha, \nu).
\]
The form \( (\ , \ ) \) is non-degenerate on \( \mathfrak{h}^* \). Since \( \mathbb{Q}(\Phi) \) contains a basis of \( \mathfrak{h}^* \) (Theorem 6.10), the restriction of \( (\ , \ ) \) to \( \mathbb{Q}(\Phi) \) is also non-degenerate. In order to prove that this restriction is \( \mathbb{Q} \)-valued, it suffices to show that \( (\alpha, \beta) \in \mathbb{Q} \) for \( \alpha, \beta \in \Phi \). In fact, we may take \( \alpha = \beta \), because \( \langle \alpha, \beta \rangle \in \mathbb{Z} \) by Theorem 6.10 and
\[
(\alpha, \beta) = \langle \alpha, t_{\beta} \rangle = \frac{1}{2} \langle \alpha, h_{\beta} \rangle (\beta, \beta).
\]
By (6.17) we know that \( (\beta, \beta) = B(t_{\beta}, t_{\beta}) \neq 0 \) and (6.20) gives \( (\beta, \beta) = \sum_{\alpha \in \Phi} (\alpha, \beta)^2 \). Therefore,
\[
\frac{1}{(\beta, \beta)} = \sum_{\alpha \in \Phi} (\alpha, \beta)^2 \quad (6.21) \quad \sum_{\alpha \in \Phi} \frac{1}{4} \langle \alpha, h_{\beta} \rangle^2  \in \mathbb{Q}.
\]
This proves that \( (\ , \ ) \) is \( \mathbb{Q} \)-valued on \( \mathbb{Q}(\Phi) \). Hence, we may consider the unique \( \mathbb{R} \)-bilinear extension of \( (\ , \ ) \) to \( \mathbb{E} \); this is a non-degenerate symmetric bilinear form on \( \mathbb{E} \). Denoting this extension by \( (\ , \ ) \) as well, formula (6.20) continues to hold for \( \mu, \nu \in \mathbb{E} \). In particular, \( (\mu, \mu) = \sum_{\alpha \in \Phi} (\alpha, \mu)^2 \geq 0 \) for all \( \mu \in \mathbb{E} \), and \( (\mu, \mu) = 0 \) forces \( (\mu, \nu) = 0 \) for all \( \nu \in \mathbb{E} \) and hence \( \mu = 0 \). This shows that \( (\ , \ ) \) is an inner product on \( \mathbb{E} \).

It remains to show that \( \dim_{\mathbb{R}} \mathbb{E} = n \) or, equivalently, \( \dim_{\mathbb{Q}} \mathbb{Q}(\Phi) = n \). By Theorem 6.10, there is a basis of \( \mathfrak{h}^* \) consisting of elements of \( \Phi \), say \( \alpha_1, \ldots, \alpha_n \).
Given any $\alpha \in \Phi$, we may write $\alpha = \sum_i k_i \alpha_i$ with $k_i \in \mathbb{K}$. The coefficients $k_i$ satisfy the following system of equations:

$$
\begin{pmatrix}
(\alpha_1, \alpha_1) & \cdots & (\alpha_1, \alpha_n) \\
\vdots & \ddots & \vdots \\
(\alpha_n, \alpha_1) & \cdots & (\alpha_n, \alpha_n)
\end{pmatrix}
= 
\begin{pmatrix}
k_1 \\
\vdots \\
k_n
\end{pmatrix}
$$

By the foregoing, the left hand side of the system and the coefficient matrix are both rational and the latter is non-singular by non-degeneracy of $(\cdot, \cdot)$. It follows that all $k_i \in \mathbb{Q}$. Therefore, $\mathbb{Q}\Phi = \bigoplus_{i=1}^n \mathbb{Q}\alpha_i$, which finishes the proof of the proposition. 

To summarize, we have the following diagram of inclusions:

$$
\begin{array}{ccc}
\mathfrak{h}^* & \cong & \mathbb{K}^n \\
\downarrow & & \downarrow \\
\mathbb{B} & \cong & \mathbb{R}^n \\
\Phi \cup & & \\
\mathbb{Q}\Phi & \cong & \mathbb{Q}^n
\end{array}
$$

Moreover, $\mathbb{B}$ is endowed with the inner product $(\cdot, \cdot)$, which in particular makes the notions of length and angle available: $\|\mu\| = \sqrt{(\mu, \mu)}$ and $(\mu, \nu) = \|\mu\| \|\nu\| \cos \theta$, where $\theta$ is the angle between $\mu, \nu \in \mathbb{B} \setminus \{0\}$. In order to study and visualize $\Phi$, we will generally view $\Phi \subseteq \mathbb{B}$. For convenience, we shall use the following shorthand notation:

$$
\langle \mu, \nu \rangle := 2 \frac{(\mu, \nu)}{\|\nu\|} = 2 \frac{\|\mu\|}{\|\nu\|} \cos \theta.
$$

Note that $\langle \mu, \nu \rangle$ is linear only in $\mu$. By (6.21), the connection with our earlier notation $\langle \cdot, \cdot \rangle : \mathfrak{h}^* \times \mathfrak{h} \to \mathbb{K}$ for the evaluation pairing is as follows:

$$
\langle \beta, \alpha \rangle = \langle \beta, h_\alpha \rangle = (\alpha, \beta) \quad (\alpha, \beta \in \Phi).
$$

These numbers are the Cartan integers of Theorem 6.10. With this, we have the following facts, which are all immediate from corresponding statements in Theorem 6.10:

- **R1:** $\Phi$ is a finite subset of $\mathbb{E} \setminus \{0\}$ that spans $\mathbb{E}$.
- **R2:** If $\alpha \in \Phi$, then $\mathbb{R}\alpha \cap \Phi = \{\pm \alpha\}$.
- **R3:** If $\alpha, \beta \in \Phi$, then $\beta - \langle \beta, \alpha \rangle \alpha \in \Phi$.
- **R4:** If $\alpha, \beta \in \Phi$, then $\langle \beta, \alpha \rangle \in \mathbb{Z}$.

Properties R1 – R4 will be taken as the defining axioms for root systems in Chapter 7.
Exercises for Section 6.3

In these exercises, $\mathfrak{h} \subseteq \mathfrak{g}$ is a Cartan subalgebra and $\Phi \subseteq \mathfrak{h}^*$ is the corresponding set of roots.

6.3.1 (Dimensions). Show that $\dim_k \mathfrak{g} = \dim_k \mathfrak{h} + |\Phi|$. Conclude that there are no semisimple Lie algebras of dimensions 4, 5 or 7.

6.3.2 (Cartan subalgebras are self-normalizing). Show that $\mathfrak{h}$ is self-normalizing, that is, $\mathfrak{h} = N_{\mathfrak{g}}(\mathfrak{h})$ (Example 5.4).

6.3.3 (Generators). Show that the root spaces $\mathfrak{g}_\alpha$ ($\alpha \in \Phi$) generate the Lie algebra $\mathfrak{g}$ and that each $\mathfrak{g}_\alpha$ consists of ad-nilpotent elements of $\mathfrak{g}$.

6.3.4 (Reductive centralizers). (a) For each $z \in \mathfrak{h}$, show that the centralizer $C_{\mathfrak{g}}(z)$ is reductive (§6.2.3).

(b) Show that there exists an element $z \in \mathfrak{h}$ such that $C_{\mathfrak{g}}(z) = \mathfrak{h}$.

6.4. The Classical Lie Algebras

So far, we have developed the theory of semisimple Lie algebras along the lines of one basic example, $\mathfrak{sl}_2$. It is time to discuss some further examples. In this section, we will describe four infinite series of semisimple (in fact, simple) Lie algebras that are generally called the classical Lie algebras of types $A_n$, $B_n$, $C_n$ and $D_n$; the subscript $n$ indicates the rank of the Lie algebra in question.

We continue to assume that $k$ is algebraically closed and $\text{char } k = 0$.

6.4.1. Checking for Semisimplicity

All examples described below will be constructed as Lie subalgebras of $\mathfrak{gl}(V)$ for some finite-dimensional $V \in \text{Vect}_k$. For such a Lie algebra $\mathfrak{g}$ to be semisimple, we must have $\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}]$ (Corollary 6.4) and so $\mathfrak{g} \subseteq \mathfrak{sl}(V)$. Semisimplicity will then be a consequence of the following proposition, which also tells us how to locate a Cartan subalgebra in each case.

Proposition 6.13. Let $V \in \text{Vect}_k$ be finite dimensional and let $\mathfrak{g} \subseteq \mathfrak{sl}(V)$ be a Lie subalgebra such that $V$ is irreducible in $\text{Rep} \mathfrak{g}$. Then $\mathfrak{g}$ is semisimple. Any self-centralizing Lie subalgebra of $\mathfrak{g}$ consisting of diagonalizable operators on $V$ is a Cartan subalgebra of $\mathfrak{g}$.

Proof. Consider an abelian ideal $\mathfrak{a}$ of $\mathfrak{g}$. There exists a common eigenvector in $V$ for all operators in $\mathfrak{a}$; in other words, there is a linear form $\lambda \in \mathfrak{a}^*$ such that the $\mathfrak{a}$-weight space $V_\lambda = \{v \in V \mid x.v = \langle \lambda, x \rangle v \text{ for all } x \in \mathfrak{a}\}$ is nonzero. Since $V_\lambda$ is a $\mathfrak{g}$-subrepresentation of $V$ (Lemma 5.17), we must have $V = V_\lambda$, because $V$ is assumed irreducible. Thus, trace $x = \langle \lambda, x \rangle \dim_k V$ for all $x \in \mathfrak{a}$. On the other
hand, all elements of \( g \) have trace zero; so \( \lambda = 0 \) and hence \( a = 0 \). We have thus shown that \( g \) is semisimple.

Now let \( h \) be a self-centralizing Lie subalgebra of \( g \) consisting of diagonalizable operators on \( V \). Then we know by Lemma 5.20 that all \( z \in h \) are ad-semisimple on \( gl(V) \) and hence on \( g \) as well. Thus, \( h \) is a self-centralizing Lie subalgebra of \( g \) consisting of ad-semisimple elements, and hence \( h \) is a Cartan subalgebra by Theorem 6.9. \( \square \)

### 6.4.2. The Special Linear Lie Algebra \( sl_{n+1} \) (Type \( A_n \))

By definition, the special linear Lie algebra \( sl_{n+1} = sl_{n+1} \left( \mathbb{K} \right) \) consists of all matrices in \( gl_{n+1} = gl_{n+1} \left( \mathbb{K} \right) \) whose trace equals 0. A \( \mathbb{K} \)-basis of \( sl_{n+1} \) is given by the matrices \( e_{i,j} \ (i \neq j) \), having 1 in position \((i, j)\) and 0 elsewhere, along with the diagonal matrices

\[
    h_i := e_{i,i} - e_{i+1,i+1} \quad (i = 1, \ldots, n).
\]

Since \( gl_{n+1} \) clearly acts irreducibly on \( V = \mathbb{K}^{n+1} \) and \( gl_{n+1} = \mathbb{K} \mathbf{1} \oplus sl_{n+1} \), it follows that \( sl_{n+1} \) also acts irreducibly. Proposition 6.13 now tells us that \( sl_{n+1} \) is semisimple.

**Cartan subalgebra.** Let \( b_{n+1} \subseteq gl_{n+1} \) be the Lie subalgebra consisting of all diagonal matrices and let \( e_{i} \in b_{n+1}^* \) denote the projection onto the \( i \)th diagonal entry. Then our earlier formula (5.12) takes the form

\[
    [d, e_{i,j}] = (e_i - e_j, d) e_{i,j} \quad (d \in b_{n+1}).
\]

It follows easily that \( b_{n+1} \) is self-centralizing in \( gl_{n+1} \). Putting

\[
    h = b_{n+1} \cap sl_{n+1} = \{ \text{diagonal matrices in } sl_{n+1} \} = \bigoplus_{i=1}^{n} \mathbb{K} h_i
\]

and noting that \( b_{n+1} = \mathbb{K} \mathbf{1} \oplus h \), we conclude that \( h \) is self-centralizing in \( sl_{n+1} \). So \( h \) is a Cartan subalgebra of \( sl_{n+1} \).

**Roots and root space decomposition.** Note that \( b_{n+1}^* = \mathbb{K}(e_1 + e_2 + \cdots + e_{n+1}) \oplus E \) with

\[
    E := \{ \sum_{i=1}^{n+1} x_i e_i \mid \sum_{i=1}^{n} x_i = 0 \} = \bigoplus_{i=1}^{n} \mathbb{K} \alpha_i \quad \text{and} \quad \alpha_i := e_i - e_{i+1}.
\]

Since \( e_1 + e_2 + \cdots + e_{n+1} \) vanishes on \( h \), the restriction map \( b_{n+1}^* \rightarrow h^* \) yields an isomorphism \( E \cong h^* \). We shall use this isomorphism to identify \( h^* \) with \( E \). In view of (6.25) the set of roots \( \Phi \) for the Cartan subalgebra \( h \) is given by

\[
    \Phi = \{ e_i - e_j \mid 1 \leq i \neq j \leq n + 1 \}
\]

\[
    = \{ \pm (e_i - e_j) \mid 1 \leq i < j \leq n + 1 \}
\]

\[
    = \{ \pm \sum_{i=1}^{j-1} \alpha_i \mid 1 \leq i < j \leq n + 1 \}.
\]
Thus, $|\Phi| = (n + 1)n$. The root space for the root $\varepsilon_i - \varepsilon_j$ is $\mathbb{k}e_{i,j}$ and the root space decomposition of $\mathfrak{sl}_{n+1}$ is

$$\mathfrak{sl}_{n+1} = \mathfrak{h} \oplus \bigoplus_{i \neq j} \mathbb{k}e_{i,j}.$$ 

$\mathfrak{sl}_2$-triples. For any semisimple Lie algebra $\mathfrak{g}$, the elements $h_\alpha \in \mathfrak{h} (\alpha \in \Phi)$ are characterized by the conditions $[g_\alpha, g_\gamma] = \mathbb{k}h_\alpha$ and $\langle \alpha, h_\gamma \rangle = 2$ (Theorem 6.10). With $\mathfrak{g} = \mathfrak{sl}_{n+1}$ and $\alpha = \varepsilon_i - \varepsilon_j \in \Phi$, the first condition becomes $\mathbb{k}h_{e_i - e_j} = [\mathbb{k}e_{i,j}, \mathbb{k}e_{j,i}] = \mathbb{k}(e_{i,i} - e_{j,j})$, and the second condition then gives

$$h_{e_i - e_j} = e_{i,i} - e_{j,j}.$$ 

Thus, the $\mathfrak{sl}_2$-triple $s_\alpha$ for the root $\alpha = \varepsilon_i - \varepsilon_j \in \Phi$ is given by

$$s_{e_i - e_j} = \mathbb{k}e_{j,i} \oplus \mathbb{k}(e_{i,i} - e_{j,j}) \oplus \mathbb{k}e_{i,j}.$$ 

Cartan integers. From (6.27) and (6.28) we obtain the Cartan integers $\langle \beta, \alpha \rangle = \langle \beta, h_\alpha \rangle$ for $\alpha, \beta \in \Phi$:

$$\langle e_i - e_j, e_r - e_s \rangle = \langle e_i - e_j, e_r, e_r - e_s \rangle.$$ 

The possible values are $\pm 2, \pm 1$ and 0. The $n \times n$ matrix $\langle \langle \alpha_i, \alpha_j \rangle \rangle_{i,j}$, with $\alpha_i$ as in (6.26), is called the Cartan matrix of $\mathfrak{sl}_{n+1}$; it will be discussed in a more general abstract setting in Section 7.2. Equation (6.29) implies in particular that

$$\langle \alpha_i, \alpha_j \rangle = \begin{cases} 2 & \text{for } i = j, \\ -1 & \text{for } |i - j| = 1, \\ 0 & \text{otherwise.} \end{cases}$$ 

Simplicity. We have already remarked earlier that $\mathfrak{sl}_{n+1}$ is simple (Exercise 5.2.6). This fact can now also be deduced rather easily from Proposition 6.11. We just need to show that the set of roots $\Phi$ in (6.27) cannot be partitioned as $\Phi = \Phi_1 \sqcup \Phi_2$ with nonempty subsets $\Phi_i$ satisfying $(\Phi_1, \Phi_2) = \{0\}$ or, equivalently, $(\Phi_1, \Phi_2) = \{0\}$. Indeed, if $\alpha_1 \in \Phi_1$, say, then (6.30) successively gives that $\alpha_2, \alpha_3, \ldots, \alpha_n$ all must belong to $\Phi_1$. Since an arbitrary root $\alpha \in \Phi$ can be written as a linear combination of $\alpha_1, \ldots, \alpha_n$ by (6.27), it follows that $\alpha$ is orthogonal to $\Phi_2$ and hence $\alpha \in \Phi_1$. Therefore, $\Phi = \Phi_1$.

Embedding into Euclidean space. In view of (6.23), the equalities (6.30) give in particular that, in the Euclidean space $\mathbb{E} \equiv \mathbb{R}^n$, we have $\|\alpha_i\| = \|\alpha_{i+1}\|$ for all $i$ and the angle $\theta$ between $\alpha_i$ and $\alpha_{i+1}$ satisfies $\cos \theta = -\frac{1}{2}$; so $\theta = \frac{2\pi}{3}$. The Cartan matrix (6.30) in conjunction with conditions $\mathbf{R1}$ and $\mathbf{R3}$ can be used to construct all roots in $\Phi$ from $\alpha_1, \ldots, \alpha_n$. First, $\mathbf{R1}$ gives all $-\alpha_i$. If we already have $\beta = \pm \sum_{i=1}^j \alpha_i \in \Phi$, then $\mathbf{R3}$ yields $\beta - \langle \beta, \alpha_{j+1} \rangle \alpha_{j+1} = \beta - \langle \pm \alpha_j, \alpha_{j+1} \rangle \alpha_{j+1} = \pm \sum_{i=1}^{j+1} \alpha_i \in \Phi$. This produces the entire set of roots $\Phi$ by (6.27). Figure 6.1 shows $\Phi$ for $\mathfrak{sl}_3$ in $\mathbb{E} \equiv \mathbb{R}^2$. 


6.4.3. The Classical Lie Algebras of Types $\mathbb{B}_n$, $\mathbb{C}_n$ and $\mathbb{D}_n$

The remaining classical Lie algebras are all constructed from a given bilinear form

$$b : V \times V \to \mathbb{k}$$

on a finite-dimensional $V \in \text{Vect}_\mathbb{k}$.

In detail, thinking of $b$ as an element of $(V \otimes V)^* \in \text{Rep} \mathfrak{gl}(V)$, we obtain a bilinear form $x.b \in (V \otimes V)^*$ for each $x \in \mathfrak{gl}(V)$. The following subspace is evidently a Lie subalgebra of $\mathfrak{gl}(V)$:

$$\mathfrak{o}(V, b) := \{ x \in \mathfrak{gl}(V) \mid x.b = 0 \}.$$

Explicitly, $(x.b)(v \otimes w) = -b(x.(v \otimes w)) = -b((x.v) \otimes w + v \otimes (x.w))$ for $v, w \in V$. So $x \in \mathfrak{o}(V, b)$ if and only if $b((x.v) \otimes w) = -b(v \otimes (x.w))$ for all $v$ and $w$. Let $B = \left( b(v_i, v_j) \right)_{i,j}$ be the matrix of the form $b$ for a given basis $(v_i)$ of $V$ and $X$ the matrix of $x \in \mathfrak{gl}(V)$. Then it is straightforward to check that

$$x \in \mathfrak{o}(V, b) \iff BX = -X^TB,$$

where $X^T$ denotes the transpose of $X$.

If the form $b$ is non-degenerate, the matrix condition (6.31) can be rewritten as $X = -B^{-1}X^TB$, which implies $\text{trace}(X) = -\text{trace}(X) = 0$. Therefore, $\mathfrak{o}(V, b) \subseteq \mathfrak{sl}(V)$ holds in this case. This will apply to all examples below. Furthermore, the following facts are not hard to check from the explicit descriptions of each of the Lie algebras $\mathfrak{g} = \mathfrak{o}(V, b)$ provided below:

- $V$ is irreducible in $\text{Rep} \mathfrak{g}$. By Proposition 6.13, this implies that $\mathfrak{g}$ is semisimple. In fact, all Lie algebras below are simple.
- The Lie subalgebra $\mathfrak{h}$ consisting of all diagonal matrices in $\mathfrak{g}$ is self-centralizing. By Proposition 6.13, it follows that $\mathfrak{h}$ is a Cartan subalgebra.

We leave the verifications of these facts to the reader (Exercise 6.4.1). Here are now the Lie algebras of types $\mathbb{B}_n$, $\mathbb{C}_n$ and $\mathbb{D}_n$. 

![Figure 6.1. Roots of $\mathfrak{sl}_3$](image-url)
The Orthogonal Lie Algebra $so_{2n+1}(k)$ (Type $B_n$). By definition, $so_{2n+1}(k)$ is the Lie algebra $o(V, b)$ for $V = k^{2n+1}$ and $b$ the bilinear form on $V$, whose matrix for the standard basis of $V$ is

$$B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1_{n \times n} \end{pmatrix}.$$  

Working out condition (6.31), one sees that $so_{2n+1} = so_{2n+1}(k)$ is the following Lie subalgebra of $gl_{2n+1}$, with blocks of the same sizes as for $B$:

$$so_{2n+1} = \begin{pmatrix} 0 & -e^T & -b^T \\ b & m & p \\ c & q & -m^T \end{pmatrix} \quad \text{with} \quad p = -p^T \quad \text{and} \quad q = -q^T.$$  

In order to exhibit a basis for $so_{2n+1}$, it will be convenient to label the rows and columns of $(2n + 1) \times (2n + 1)$-matrices by $0, 1, \ldots, 2n$. Denoting the subalgebra of diagonal matrices in $so_{2n+1}$ by $\mathfrak{h}$ as usual, we have

$$\mathfrak{h} = \bigoplus_{i=1}^{n} k h_i \quad \text{with} \quad h_i := e_{i,i} - e_{n+i,n+i}. $$  

To complement $\mathfrak{h}$ in the two $m$-blocks of (6.32), we pick the matrices

$$m_{i,j} := e_{i,j} - e_{n+j,n+i} \quad (1 \leq i \neq j \leq n).$$  

The $b$-blocks and $c$-blocks are covered by the matrices

$$b_i := e_{i,0} - e_{0,n+i} \quad \text{and} \quad c_i := e_{0,i} - e_{n+i,0} \quad (1 \leq i \leq n).$$  

Finally, for the blocks labeled $p$ and $q$, we choose

$$p_{i,j} := e_{i,n+j} - e_{j,n+i} \quad \text{and} \quad q_{j,i} := p_{i,j}^T \quad (1 \leq i < j \leq n).$$  

Altogether, these matrices form a basis of $so_{2n+1}$. In particular,  

$$\dim_k so_{2n+1} = 2n^2 + n.$$  

The choices made in the foregoing do in fact give weight vectors for the adjoint action of $\mathfrak{h}$. Specifically, letting $\{e_i\}$ denote the basis of $\mathfrak{h}^* \cong k^n$ that is dual to the basis $(h_i)$ of $\mathfrak{h}$ in (6.33), we have the following table:

<table>
<thead>
<tr>
<th>weight vector</th>
<th>$b_i$</th>
<th>$c_i$</th>
<th>$m_{i,j}$</th>
<th>$p_{i,j}$</th>
<th>$q_{j,i}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>weight</td>
<td>$e_i$</td>
<td>$-e_i$</td>
<td>$e_i - e_j$</td>
<td>$e_i + e_j$</td>
<td>$-(e_i + e_j)$</td>
</tr>
</tbody>
</table>

We leave it to the reader to check that $so_{2n+1}$ acts irreducibly on $k^{2n+1}$ and that $\mathfrak{h}$ is self-centralizing in $so_{2n+1}$ (Exercise 6.4.1). Thus, $so_{2n+1}$ is semisimple and $\mathfrak{h}$ is a Cartan subalgebra. By the above table, the corresponding set of roots is

$$\Phi = \{ \pm e_i \mid 1 \leq i \leq n \} \cup \{ \pm e_i \pm e_j \mid 1 \leq i < j \leq n \}.$$  

All roots can be expressed in terms of the following roots:

$$\alpha_i := e_i - e_{i+1} \quad (i = 1, \ldots, n-1) \quad \text{and} \quad \alpha_n := e_n.$$
Indeed,\

\[ \pm e_i = \pm \sum_{k=i}^{n} \alpha_k, \]

\[ (6.36) \]

\[ \pm (e_i - e_j) = \pm \sum_{k=i}^{j-1} \alpha_k, \]

\[ \pm (e_i + e_j) = \pm \left( \sum_{k=i}^{j-1} \alpha_k + 2 \sum_{k=j}^{n} \alpha_k \right). \]

Finally, we address the Cartan integers \( \langle \beta, \alpha \rangle = \langle \beta, h_\alpha \rangle \) for \( \alpha, \beta \in \Phi \). Recall that \( h_\alpha \) is determined by the conditions \([g_\alpha, g_{-\alpha}] = k h_\alpha \) and \( \langle \alpha, h_\alpha \rangle = 2 \) (Theorem 6.10). From the easily verified relations \([b_i, c_i] = h_i, [m_{i,j}, m_{j,i}] = h_i - h_j \) and \([p_{i,j}, q_{j,i}] = h_i + h_j \) in conjunction with \( \langle e_i, h_j \rangle = \delta_{i,j} \), one obtains

\[ (6.37) \]

\[ h_{\pm e_i} = \pm 2 h_i \quad \text{and} \quad h_{\pm e_i, \pm e_j} = \pm h_i \pm h_j. \]

Here is the Cartan matrix \( (\langle \alpha_i, \alpha_j \rangle)_{i,j} \) of \( so_{2n+1} \), with \( \alpha_i \) as in (6.35):

\[ (6.38) \]

\[ \langle \alpha_i, \alpha_j \rangle = \begin{cases} 
2 & \text{for } i = j, \\
-1 & \text{for } |i - j| = 1 \text{ with } i, j \neq n, \\
-1 & \text{for } (i, j) = (n, n - 1), \\
-2 & \text{for } (i, j) = (n - 1, n), \\
0 & \text{otherwise.} 
\end{cases} \]

This Cartan matrix is nearly identical to the one in (6.30) for \( sl_{n+1} \), the difference being that the entry in position \((n - 1, n)\) is now \(-2\) rather than \(-1\). In particular, we see exactly as for \( sl_{n+1} \) that the collection \( \{\alpha_1, \ldots, \alpha_n\} \) cannot be partitioned into nonempty orthogonal subsets. Since every root \( \alpha \in \Phi \) is a linear combination of \( \alpha_1, \ldots, \alpha_n \) by (6.36), it follows as in §6.4.2 that \( \Phi \) cannot be partitioned into nonempty orthogonal subsets either. Hence, the Lie algebra \( so_{2n+1} \) is simple (Proposition 6.11).

The Symplectic Lie Algebra \( sp_{2n}(\mathbb{k}) \) (Type \( C_n \)). Here we take \( V = \mathbb{k}^{2n} \) and \( B = \begin{pmatrix} 0 & 1_{n \times n} \\ -1_{n \times n} & 0 \end{pmatrix} \). The Lie algebra \( o(V, b) \) is denoted by \( sp_{2n}(\mathbb{k}) \) or just \( sp_{2n} \).

Condition (6.31) now gives

\[ (6.39) \]

\[ sp_{2n} = \left\{ \begin{pmatrix} m & p \\ q & -m^T \end{pmatrix} \right\} \quad \text{with} \quad p = p^T \quad \text{and} \quad q = q^T. \]

The dimension is the same as for \( so_{2n+1} \), namely

\[ \dim_\mathbb{k} sp_{2n} = 2n^2 + n. \]
We leave it to the reader to carry out a more detailed analysis of this case along the lines of \( \mathfrak{so}_{2n+1} \) above. Alternatively, he or she may wish to consult the literature, for example [66, Section 12.5], where the main features of \( \mathfrak{sp}_{2n} \) are well documented.

**The Orthogonal Lie Algebra** \( \mathfrak{so}_{2n}(\mathbb{k}) \) (Type \( D_n \)). With \( V = \mathbb{k}^{2n} \) and \( B = \begin{pmatrix} 0 & 1_{n \times n} \\ 1_{n \times n} & 0 \end{pmatrix} \), we obtain \( \mathfrak{so}_{2n}(\mathbb{k}) \). Explicitly,

\[
(6.40) \quad \mathfrak{so}_{2n} = \left\{ \begin{pmatrix} m & p \\ q & -m^T \end{pmatrix} \mid p = -p^T \text{ and } q = -q^T \right\}.
\]

Now, \( \dim \mathbb{k} \mathfrak{so}_{2n} = 2n^2 - n \). Again, we omit the details; see [66, Section 12.4], for example.

**Exercises for Section 6.4**

**6.4.1** (Details for the classical Lie algebras). For each of the Lie algebras \( \mathfrak{g} = \mathfrak{so}_{2n+1}(\mathbb{k}), \mathfrak{so}_{2n}(\mathbb{k}) \) and \( \mathfrak{sp}_{2n}(\mathbb{k}) \), let \( \mathfrak{h} \) denote the Lie subalgebra consisting of all diagonal matrices in \( \mathfrak{g} \). Prove:

(a) The \( \mathbb{k} \)-subalgebra of \( \text{End}_\mathbb{k}(V) \) that is generated by \( \mathfrak{h} \) is the algebra of all diagonal matrices in \( \text{End}_\mathbb{k}(V) \). Conclude that \( \mathfrak{h} \) is self-centralizing in \( \mathfrak{g} \).

(b) The \( \mathbb{k} \)-subalgebra of \( \text{End}_\mathbb{k}(V) \) that is generated by \( \mathfrak{g} \) is all of \( \text{End}_\mathbb{k}(V) \). Conclude that \( V \in \text{Rep} \mathfrak{g} \) is irreducible.

**6.4.2** (Killing forms of the classical Lie algebras). Let \( \mathfrak{g} \) be one of the Lie algebras of types \( A_n \) – \( C_n \). Using Lemma 6.6 show that the Killing form of \( \mathfrak{g} \) is given by \( B(x, y) = \lambda \text{trace}(xy) \) for \( x, y \in \mathfrak{g} \), with

\[
\lambda = \begin{cases} 
2(n + 1) & \text{for types } A_n \text{ and } C_n, \\
2n - 1 & \text{for type } B_n, \\
2(n - 1) & \text{for type } D_n.
\end{cases}
\]
Chapter 7

Root Systems

This chapter studies sets of roots arising from semisimple Lie algebras abstractly, taking properties R1 – R4 (§6.3.4) as the point of departure. Any collection of vectors in some Euclidean space satisfying these conditions is called a root system. It turns out that every abstract root system defined in this way is in fact the set of roots of a unique (up to isomorphism) semisimple Lie algebra. Lie algebras for the so-called classical root systems of types A – D have already been exhibited in Section 6.4. In general, following Serre [183], the Lie algebra can be constructed by generators and relations that depend only on the given root system. We will not discuss Serre’s Theorem in detail here and neither will we prove the celebrated Classification Theorem for irreducible root systems, although the latter will at least be stated in §7.3.3. Excellent sources for this material include the standard references Bourbaki [23], [28] and Humphreys [101] as well as the very accessible monograph by Erdmann and Wildon [66].

After presenting the basic definitions of root systems and their associated Weyl groups, we will study bases of roots systems; these will allow us to distinguish between “positive” and “negative” roots. Then we will discuss certain lattices that are associated to a given root system: the root lattice and the weight lattice. The chapter will close with an invariant theoretic result on the multiplicative invariants of weight lattices under the action of the Weyl group. This result will be important later in connection with representation rings of semisimple Lie algebras (Section 8.5).

Throughout this chapter, $\mathbb{E}$ denotes a Euclidean space, that is, a finite-dimensional $\mathbb{R}$-vector space that is equipped with an inner product $(\cdot, \cdot)$. We put $n = \dim \mathbb{E}$. 
7.1. Abstract Root Systems

For $\mu, \nu \in \mathbb{E} \setminus \{0\}$, define

$$\langle \mu, \nu \rangle := \frac{(\mu, \nu)}{(\nu, \nu)} = 2 \frac{\|\mu\|}{\|\nu\|} \cos \theta,$$

where $\theta$ is the angle between $\mu$ and $\nu$. The bracket $\langle \cdot, \cdot \rangle$ is linear only in the first variable, and it is insensitive to rescaling the given inner product $\langle \cdot, \cdot \rangle$.

A subset $\Phi \subseteq \mathbb{E}$ is called a root system of rank $n = \dim \mathbb{E}$ if the following conditions are satisfied:

- **R1:** $\Phi$ is a finite subset of $\mathbb{E} \setminus \{0\}$ that spans $\mathbb{E}$;
- **R2:** If $\alpha \in \Phi$, then $R_\alpha \cap \Phi = \{\pm \alpha\}$;
- **R3:** If $\alpha, \beta \in \Phi$, then $\beta - \langle \beta, \alpha \rangle \alpha \in \Phi$;
- **R4:** If $\alpha, \beta \in \Phi$, then $\langle \beta, \alpha \rangle \in \mathbb{Z}$.

Vectors $\alpha \in \Phi$ will be called roots and the integers $\langle \beta, \alpha \rangle$ are called the Cartan integers of the root system. In the literature, root systems as defined above are sometimes referred to as reduced ($R2$) crystallographic ($R4$) root systems, but we will not do so here.

An isomorphism of root systems $\Phi \subseteq \mathbb{E}$ and $\Phi' \subseteq \mathbb{E}'$ is an $\mathbb{R}$-linear isomorphism $f : \mathbb{E} \isom \mathbb{E}'$ satisfying $f(\Phi) = \Phi'$. We do not require $f$ to be an isometry. However, Lemma 7.2 below will show that $f$ automatically preserves the Cartan integers: $\langle f\alpha, f\beta \rangle = \langle \alpha, \beta \rangle$ for all $\alpha, \beta \in \Phi$. It follows that $f$ also preserves angles between pairs of roots and the ratios of lengths of non-orthogonal roots; see §7.1.1 below.

7.1.1. The Crystallographic Restriction

The “crystallographic” axiom $R4$ places severe restrictions on the possible angles between roots and on their relative lengths. Specifically, let $\alpha, \beta \in \Phi$ with $\beta \neq \pm \alpha$. Then (7.1) gives $\langle \alpha, \beta \rangle \langle \beta, \alpha \rangle = 4 \cos^2 \theta$, where $\theta$ is the angle between $\alpha$ and $\beta$.

Now, $0 \leq \cos^2 \theta \leq 1$ and, on the other hand, $\langle \alpha, \beta \rangle \langle \beta, \alpha \rangle$ must be an integer by $R4$. Since the value 4 is ruled out by our assumption that $\beta \neq \pm \alpha$, we are left with the following possibilities:

$$\langle \alpha, \beta \rangle \langle \beta, \alpha \rangle = 4 \cos^2 \theta \in \{0, 1, 2, 3\}$$

Note also that $\langle \alpha, \beta \rangle = 0$ if and only if $\langle \beta, \alpha \rangle = 0$, with both equations stating that $\alpha$ and $\beta$ are orthogonal. If $\langle \alpha, \beta \rangle$ and $\langle \beta, \alpha \rangle$ are nonzero, then (7.2) forces one of these values to be $\pm 1$, while the other is $\pm 2$ or $\pm 3$, and both values have the same sign. Thus, $\langle \alpha, \beta \rangle \langle \beta, \alpha \rangle = 1$ means that $\cos \theta = \pm \frac{1}{2}$ and $\|\alpha\| = \|\beta\|$. Finally, $\langle \alpha, \beta \rangle \langle \beta, \alpha \rangle \in \{2, 3\}$ if and only if $\cos \theta \in \{\pm \frac{1}{2} \sqrt{2}, \pm \frac{1}{2} \sqrt{3}\}$. Since $\langle \alpha, \beta \rangle = 2 \frac{\|\alpha\|}{\|\beta\|} \cos \theta \in \mathbb{Z}$, we must have $\|\alpha\| \neq \|\beta\|$ in this case. In fact, if $\|\beta\| > \|\alpha\|$, say,
then $\langle \beta, \alpha \rangle = 2 \frac{\|\beta\|}{\|\alpha\|} \cos \theta \in \{ \pm 2, \pm 3 \}$ and so $\frac{\|\beta\|}{\|\alpha\|} \in \{ \sqrt{2}, \sqrt{3} \}$. The results of this discussion are summarized in Table 7.1.

<table>
<thead>
<tr>
<th>$\langle \alpha, \beta \rangle$</th>
<th>$\langle \beta, \alpha \rangle$</th>
<th>$\theta$</th>
<th>$|\beta|/|\alpha|$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>$\pi/2$</td>
<td>undetermined</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>$\pi/3$</td>
<td>1</td>
</tr>
<tr>
<td>$-1$</td>
<td>$-1$</td>
<td>$2\pi/3$</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>$\pi/4$</td>
<td>$\sqrt{2}$</td>
</tr>
<tr>
<td>$-1$</td>
<td>$-2$</td>
<td>$3\pi/4$</td>
<td>$\sqrt{2}$</td>
</tr>
<tr>
<td>1</td>
<td>3</td>
<td>$\pi/6$</td>
<td>$\sqrt{3}$</td>
</tr>
<tr>
<td>$-1$</td>
<td>$-3$</td>
<td>$5\pi/6$</td>
<td>$\sqrt{3}$</td>
</tr>
</tbody>
</table>

Table 7.1. Pairs of roots $\alpha, \beta \in \Phi$ with $\beta \neq \pm \alpha$, $\|\beta\| \geq \|\alpha\|$.

Lemma 7.1. Let $\alpha, \beta \in \Phi$ with $\beta \neq \pm \alpha$. If $(\alpha, \beta) > 0$, then $\pm (\alpha - \beta) \in \Phi$; and if $(\alpha, \beta) < 0$, then $\pm (\alpha + \beta) \in \Phi$.

Proof. First assume that $(\alpha, \beta) > 0$. Then at least one of $\langle \alpha, \beta \rangle$ or $\langle \beta, \alpha \rangle$ must be 1, say $\langle \alpha, \beta \rangle = 1$. Then $\alpha - \beta \in \Phi$ by R3 and so $\pm (\alpha - \beta) \in \Phi$ by R2. The case $(\alpha, \beta) < 0$ follows by applying the foregoing to the root $-\beta$ in place of $\beta$. □

7.1.2. Root Systems of Ranks $\leq 2$

Evidently, $\Phi = \emptyset$ is the only root system of rank 0. In view of axiom R2, there is also a unique root system of rank 1, up to isomorphism; it is called the root system of type $A_1$:

\[ -\alpha \leftarrow \cdots \rightarrow \alpha \]

In Section 5.7, we have encountered this root system as the set of roots of the Lie algebra $\mathfrak{sl}_2 = \mathbb{K}e \oplus \mathbb{K}h \oplus \mathbb{K}e$, with $\langle \alpha, h \rangle = 2$.

Figure 7.1 displays all possible root systems of rank 2, up to isomorphism, with their standard labels. In each case, $\Phi$ is depicted in the ordinary Euclidean plane $\mathbb{E} = \mathbb{R}^2$. Axioms R1 - R4 are straightforward to check for all four root systems. Also, these root systems are plainly non-isomorphic, having different numbers of roots. In order to verify that they are the only possibilities in rank 2, start with a root $\alpha \in \Phi$ of minimal length. By R1, there must be a root $\beta \neq \pm \alpha$ in $\Phi$. By considering $-\beta$ if necessary, we may assume that $(\alpha, \beta) \leq 0$; so the angle $\theta$ in Table 7.1 is one
of $\pi/2$, $2\pi/3$, $3\pi/4$ or $5\pi/6$. Using Table 7.1 and \textbf{R2}, \textbf{R3}, one easily checks that these four cases result in the four displayed root systems.

![Image of root systems](image)

**Figure 7.1.** Rank 2 root systems ($\mathbb{E} = \mathbb{R}^2$)

### 7.1.3. Automorphism Group and Weyl Group

For a given root system $\Phi \subseteq \mathbb{E}$, we define

$$\text{Aut } \Phi \overset{\text{def}}{=} \{ f \in \text{GL}(\mathbb{E}) \mid f\Phi = \Phi \}.$$  

It is clear that $\text{Aut } \Phi$ is a subgroup of $\text{GL}(\mathbb{E})$. Moreover, it follows from \textbf{R1} that the restriction map $\text{Aut } \Phi \rightarrow S_{\Phi}$, $f \mapsto f|_{\Phi}$ is a monomorphism of $\text{Aut } \Phi$ into the group of permutations of $\Phi$. Therefore, $\text{Aut } \Phi$ is a finite group.

For each $0 \neq \nu \in \mathbb{E}$, consider the map $s_{\nu} : \mathbb{E} \rightarrow \mathbb{E}$ that is defined by

$$s_{\nu}\mu = \mu - \langle \mu, \nu \rangle \nu \quad (\mu \in \mathbb{E}).$$

This map is obviously linear; it sends $\nu$ to $-\nu$; and it is the identity on the hyperplane $\nu^\perp = \{ \mu \in \mathbb{E} \mid \langle \mu, \nu \rangle = 0 \}$. Moreover, it is easy to see that $s_{\nu}$ is an orthogonal transformation of $\mathbb{E}$ (Exercise 7.1.1). The map $s_{\nu}$ is called the \textit{reflection} of $\mathbb{E}$ that is associated with $\nu$. Axiom \textbf{R3} can now be restated as follows:
In other words, \( s_\alpha \in \text{Aut } \Phi \) for all \( \alpha \in \Phi \). The \textbf{Weyl group} of \( \Phi \) is the subgroup of \( \text{Aut } \Phi \) that is generated by these reflections:

\[ \mathcal{W} = \langle s_\alpha \mid \alpha \in \Phi \rangle \subseteq \text{Aut } \Phi. \]

Since \( \text{Aut } \Phi \) is a finite group, \( \mathcal{W} \) is likewise. The following lemma shows that \( \mathcal{W} \) is in fact a normal subgroup of \( \text{Aut } \Phi \); part (b) also justifies the earlier claim that isomorphism of root systems preserve \( \langle \cdot, \cdot \rangle \) on roots. Below, we shall generally dispense with the composition symbol \( \circ \) if this leads to no misunderstandings.

\begin{lemma}
Let \( \Phi \subseteq \mathbb{E} \) and \( \Phi' \subseteq \mathbb{E}' \) be isomorphic root systems via an \( \mathbb{R} \)-linear isomorphism \( f: \mathbb{E} \xrightarrow{\sim} \mathbb{E}' \) satisfying \( f\Phi = \Phi' \). Then:

(a) For all \( \alpha \in \Phi \), the equality \( fs_\alpha f^{-1} = s_{f\alpha} \) holds in \( \text{GL}(\mathbb{E}') \).

(b) If \( \alpha, \beta \in \Phi \), then \( \langle f\alpha, f\beta \rangle = \langle \alpha, \beta \rangle \).
\end{lemma}

\begin{proof}
(a) Put \( \alpha' := f\alpha \in \Phi' \) and write \( s := s_{\alpha'} \) and \( r := fs_\alpha f^{-1} \) for brevity. Observe that \( s \) and \( r \) share the following properties:

(i) both \( s \) and \( r \) belong to the subgroup \( \text{Aut } \Phi' \) of \( \text{GL}(\mathbb{E}') \);

(ii) \( sa' = -a' = r\alpha' \); and

(iii) both induce the identity transformation on \( \mathbb{E}'/\mathbb{R}a' \).

Consider the element \( t := sr \in \text{Aut } \Phi' \). We wish to show that \( t = \text{Id} \). Since \( \text{Aut } \Phi' \) is finite, we certainly know that \( t \) has finite order. Furthermore, \( ta' = a' \) by (ii) and \( t \) also has property (iii). Thus, for any \( \mu \in \mathbb{E}' \), we have \( t\mu = \mu + c\alpha' \) for some \( c \in \mathbb{R} \). It follows that \( t^n\mu = \mu + nca' \). Since \( t^n = \text{Id} \) for some \( n > 0 \), we must have \( c = 0 \), and so \( t = \text{Id} \) as desired.

(b) Since \( f\beta \neq 0 \), the following computation proves (b):

\[ f\alpha - \langle \alpha, \beta \rangle f\beta = f(s_{f\beta}\alpha) = s_{f\beta}(f\alpha) = f\alpha - \langle f\alpha, f\beta \rangle f\beta. \quad \square \]

If \( f \) is an isomorphism of root systems as in the lemma, then the group isomorphism \( f \cdot f^{-1}: \text{GL}(\mathbb{E}) \xrightarrow{\sim} \text{GL}(\mathbb{E}') \) restricts to isomorphisms

\[ \text{Aut } \Phi \xrightarrow{\sim} \text{Aut } \Phi' \quad \text{and} \quad \mathcal{W}_\Phi \xrightarrow{\sim} \mathcal{W}_{\Phi'}. \]

In particular, \( \mathcal{W}_\Phi \) is a normal subgroup of \( \text{Aut } \Phi \); see Exercise 7.2.4 for a more precise description of the relationship between these two groups.
7.1.4. The Classical Root Systems

Type $A_n$

Let $\mathbb{B} = \{(x_1, x_2, \ldots, x_{n+1}) \in \mathbb{R}^{n+1} \mid \sum x_i = 0\} \cong \mathbb{R}^n$ and equip $\mathbb{B}$ with the inner product $\langle \cdot, \cdot \rangle$ that is obtained by restricting the standard dot product of $\mathbb{R}^{n+1}$ to $\mathbb{B}$. Thus, $\mathbb{B} = L^\perp$ with $L := \mathbb{R}(e_1 + \cdots + e_{n+1})$, where $e_i \in \mathbb{R}^{n+1}$ denotes the standard basis vector, with $1$ in position $i$ and $0$ elsewhere. Then

$$\Phi = \{e_i - e_j \mid 1 \leq i \neq j \leq n + 1\} = \{\pm (e_i - e_j) \mid 1 \leq i < j \leq n + 1\}$$

is a root system in $\mathbb{B}$. Indeed, $\mathbf{R1}$ and $\mathbf{R2}$ are both obvious. The “crystallographic” axiom $\mathbf{R4}$ follows from the computation

$$\langle e_i - e_j, e_r - e_s \rangle = \frac{2(e_i - e_j, e_r - e_s)}{(e_r - e_s, e_r - e_s)} = (e_i - e_j, e_r - e_s) \in \{0, \pm 1, \pm 2\}.$$

To describe the reflections $s_{\alpha}$ ($\alpha \in \Phi$), write $\mathbb{R}^{n+1} = L \oplus \mathbb{B}$ and consider the embedding $\text{GL}(\mathbb{B}) \hookrightarrow \text{GL}(\mathbb{R}^{n+1})$, $f \mapsto \text{Id}_L \oplus f$. Then $s_{\alpha} \mu = \mu - (\mu, \alpha)\alpha$ for $\mu \in \mathbb{R}^{n+1}$, since $(\alpha, \alpha) = 2$ and so $\langle \mu, \alpha \rangle = (\mu, \alpha)$. In particular,

$$s_{e_i - e_j}(e_l) = e_l - (e_i, e_l - e_j)(e_i - e_j) = \begin{cases} e_l & \text{for } l \notin \{i, j\}, \\ e_j & \text{for } l = i, \\ e_i & \text{for } l = j. \end{cases}$$

Thus, $s_{e_i - e_j}$ permutes the basis $(e_l)_{l=1}^{n+1}$ of $\mathbb{R}^{n+1}$, acting as the transposition $(i, j) \in S_{n+1}$. Since these transpositions generate the symmetric group $S_{n+1}$, the Weyl group $\mathcal{W} = \mathcal{W}_\Phi$ is isomorphic to $S_{n+1}$ via

$$\mathcal{W} \xrightarrow{\sim} S_{n+1}$$

(7.4)

$$w \xmapsto{\sim} w \mid_{\{e_1, \ldots, e_{n+1}\}}$$

Viewing this isomorphism as an identification, we have $w(e_i - e_j) = e_{wi} - e_{wj}$. Thus, $\mathcal{W}$ maps $\Phi$ to itself, and hence $\mathbf{R3}$ is also satisfied.

We have seen the root system $\Phi$ before: the same collection of vectors was realized in (6.27) as the set of roots of the Lie algebra $\mathfrak{sl}_{n+1}$ inside $E = \{\sum_{i=1}^{n+1} x_i e_i \mid \sum_{i=1}^n x_i = 0\} \subseteq \bigoplus_{i=1}^{n+1} \mathbb{R} e_i$ and then viewed inside $\mathbb{B} \cong \mathbb{R}^n$, equipped with the inner product that comes from the Killing form of $\mathfrak{sl}_{n+1}$. We did not discuss the specifics of the latter inner product earlier—see Exercise 6.4.2 for this—but the Cartan integers do of course agree with those determined earlier in (6.29), as must be the case by Lemma 7.2.
7.1. Abstract Root Systems

Here is the familiar picture of $A_2$ again, now as part of the (gray) hyperplane $E = (\varepsilon_1 + \varepsilon_2 + \varepsilon_3)^\perp \subseteq \mathbb{R}^3$:

$$A_2$$

To picture $A_3$ in $\mathbb{R}^3$, note that the roots in $A_n$ can also be written as $\mu_i - \mu_j$ for $1 \leq i, j \leq n + 1$, where we have put

$$\mu_k := \varepsilon_k - \frac{1}{n + 1} \sum_{j=1}^{n+1} \varepsilon_j.$$

All $\mu_k$ belong to $E$; they all have the same length; and the angles between any two of them are the same. Furthermore, $\sum_{k=1}^{n+1} \mu_k = 0$. For $n = 3$, we visualize $\mu_1, \dots, \mu_4$ as the vectors to every other vertex of a unit cube centered at the origin. Here is the resulting picture of $\tilde{A}_3$ with the $\mu_k$ represented by black dots:

$$\mu_1 \mu_2 \mu_3 \mu_4$$

We have also highlighted the roots $\alpha_i := \varepsilon_i - \varepsilon_{i+1} = \mu_i - \mu_{i+1}$ in bright red; they are the same as in (6.26). The significance of these particular roots and the differences in the coloring of the other roots will be explained in Section 7.2; see Example 7.3.

**Type $B_n$**

Now we let $E = \mathbb{R}^n$, with the ordinary dot product for $(\cdot, \cdot)$ and standard basis $\varepsilon_1, \dots, \varepsilon_n$ as before. Define

$$\Phi = \{ \pm \varepsilon_i \mid 1 \leq i \leq n \} \cup \{ \pm \varepsilon_i \pm \varepsilon_j \mid 1 \leq i < j \leq n \}.$$
This is the set of roots of the Lie algebra $\mathfrak{so}_{2n+1}$; see (6.34). In contrast with the root system of type $A_n$, not all $\alpha \in \Phi$ have the same length if $n \geq 2$: there are $2n$ roots with $||\alpha|| = 1$ and $2n(n-1)$ roots with $||\alpha|| = \sqrt{2}$. Once again, $R_1$ and $R_2$ are evident and $R_4$ also holds: $(\alpha, \beta) \in \{0, \pm 1, \pm 2\}$ for all $\alpha, \beta \in \Phi$ and $(\alpha, \beta) = 2(\alpha, \beta)$ if $\beta$ is short, while $(\alpha, \beta) = (\alpha, \beta)$ if $\beta$ is long. Now for the reflections $s_\alpha$ with $\alpha \in \Phi$. Since $s_\alpha = s_{-\alpha}$, the first $\pm$ in the description of $\alpha \in \Phi$ can always be taken to be $\pm$. If $\alpha = e_j$ is short, then $s_\alpha$ sends $e_\ell \mapsto -e_\ell$ and fixes all other $e_i$. For long $\alpha = e_i \pm e_j$, we have

$$s_{e_i \pm e_j}(e_\ell) = e_\ell - (e_i, e_\ell) (e_i \pm e_j) = \begin{cases} e_\ell & \text{for } \ell \notin \{i, j\}, \\ \mp e_j & \text{for } \ell = i, \\ \mp e_i & \text{for } \ell = j. \end{cases}$$

Thus the matrices of the reflections $s_\alpha$ for the basis $(e_\ell)_n$ of $\mathbb{R}^n$ are quasi-permutation integer matrices: each row and each column contains exactly one entry $\pm 1$ while all other entries are 0. Since $\Phi$ is evidently stable under all quasi-permutation matrices, we have verified $R_3$. To describe the Weyl group $W = \mathcal{W}_\Phi$, observe that the matrices of $s_\alpha$, for short $\alpha$ generate the subgroup of diagonal matrices in $GL_n(\mathbb{Z})$, which is isomorphic to $\{\pm 1\}^n$, and the matrices of $s_\alpha$ for long $\alpha = e_i - e_j$ generate the subgroup of all permutation matrices in $GL_n(\mathbb{Z})$, which is isomorphic to $S_n$. Together, these matrices generate the subgroup of all quasi-permutation integer matrices, which is isomorphic to the direct product $\{\pm 1\}^n \rtimes S_n$ with $S_n$ permuting the factors of $\{\pm 1\}^n$ according to the standard permutation action on $[n]$. In summary,

$$(7.7) \quad \mathcal{W} = \{\pm 1\}^n \rtimes S_n.$$

The root system of type $B_1$ is of course isomorphic to $A_1$, the only root system of rank 1 up to isomorphism, but this is no longer true for $n \geq 2$, because $B_n$ has $2n^2$ roots whereas $A_n$ has $n^2 + n$. The picture of $B_2$ is part of Figure 7.1.

**Type $C_n$**

This root system may be defined as the dual root system of the root system $B_n$ in $\mathbb{R} = \mathbb{R}^n$, again with the ordinary dot product; see Exercise 7.1.3. Explicitly, if $(e_\ell)_n$ is the standard basis of $\mathbb{R}$ as usual, then $C_n$ has the following $2n^2$ roots:

$$\Phi = \{ \pm 2e_i \mid 1 \leq i \leq n \} \cup \{ \pm e_i \pm e_j \mid 1 \leq i < j \leq n \}.$$ 

Since the reflections $s_\alpha$ are unaltered by a rescaling of $\alpha$, the Weyl group is the same as for $B_n$; see (7.7):

$$(7.8) \quad \mathcal{W} = \{\pm 1\}^n \rtimes S_n.$$

However, while $2n$ of the roots of $B_n$ are short and $2n(n-1)$ long, these proportions are reversed for the root system $C_n$. Rescaling $\mathbb{R}^2$ by $\sqrt{2}$ and rotating by $\pi/4$, one obtains $C_2$ from $B_2$; so $C_2$ is isomorphic to $B_2$. However, for $n \geq 3$, the root system
C_n is not isomorphic to A_n, since the root numbers are different, nor to B_n, because any isomorphism of root systems preserves ratios of lengths of non-orthogonal roots.

**Type D_n**

Keeping the notation from the previous two cases, B_n and C_n, the last of the classical root systems consists of all vectors \( \alpha \) in the lattice \( \mathbb{Z}^n = \bigoplus_i \mathbb{Z} e_i \subseteq \mathbb{E} = \mathbb{R}^n \) satisfying \( (\alpha, \alpha) = 2 \):

\[ \Phi = \{ \pm e_i \pm e_j \mid 1 \leq i < j \leq n \} . \]

Thus, there are \( 2n(n-1) \) root vectors, all having the same length, and all root vectors are also part of the root systems B_n and C_n. Therefore, the Weyl group of D_n is a subgroup of (7.7), (7.8). Explicitly, letting \( D \leq \{ \pm 1 \}^{\times n} \) denote the subgroup consisting of all \( (d_j) \) with \( \prod d_j = 1 \), we have (Exercise 7.1.4)

\[ W = D \rtimes S_n \cong \{ \pm 1 \}^{\times (n-1)} \rtimes S_n . \]

It is easy to see that \( D_2 \cong A_1 \times A_1 \) and \( D_3 \cong A_3 \), but there are no further isomorphisms to any of the other classical root systems A_n, B_n or C_n.

**Exercises for Section 7.1**

7.1.1 (Orthogonality of reflections). Show that all reflections \( s = s_\nu \) (0 \( \neq \) \( \nu \) \subseteq \( \mathbb{E} \)) are orthogonal transformations of \( \mathbb{E} \) : \( (s\lambda, s\mu) = (\lambda, \mu) \) for all \( \lambda, \mu \subseteq \mathbb{E} \).

7.1.2 (Rank 2 root systems). Using Figure 7.1, show that the root system \( A_1 \times A_1 \) has Weyl group \( S_2 \times S_2 \) and that the Weyl group of \( G_2 \) is \( D_6 \), the dihedral group of order 12.

7.1.3 (Dual root systems). Let \( \Phi \) be a root system in \( \mathbb{E} \). For 0 \( \neq \) \( \mu \subseteq \mathbb{E} \), put \( \mu^\vee = \frac{2\mu}{(\mu, \mu)} \subseteq \mathbb{E} \). Prove:

(a) \( \mu^{\vee \vee} = \mu \), \( \langle \lambda, \mu \rangle = \langle \mu^\vee, \lambda^\vee \rangle \) and \( s_\mu^\vee(\lambda^\vee) = s_\mu(\lambda)^\vee \) for all 0 \( \neq \) \( \lambda, \mu \subseteq \mathbb{E} \).

(b) \( \Phi^\vee := \{ \alpha^\vee \mid \alpha \subseteq \Phi \} \) is a root system in \( \mathbb{E} \), with \( (\Phi^\vee)^\vee = \Phi \). The root system \( \Phi^\vee \) is called the **dual root system** of \( \Phi \) and \( \alpha^\vee \) is called the **coroot** of \( \alpha \).

(c) \( W_\Phi = W_{\Phi^\vee} \).

7.1.4 (Weyl group of D_n). Check the description of the Weyl group given in (7.9).

**7.2. Bases of a Root System**

We keep the notations from Section 7.1. In particular, \( \Phi \subseteq \mathbb{E} \cong \mathbb{R}^n \) will always denote a root system in this section.

A subset \( \Delta \subseteq \Phi \) is called a **base** of \( \Phi \) if the following conditions are satisfied:
**B1** \( \Delta \) is an \( \mathbb{R} \)-basis of \( \mathbb{E} \), and

**B2** Each \( \beta \in \Phi \) has the form \( \beta = \sum_{\alpha \in \Delta} z_\alpha \alpha \) with all \( z_\alpha \in \mathbb{Z}_+ \) or all \( z_\alpha \in -\mathbb{Z}_+ \).

The roots in \( \Delta \) are called **simple roots**; those with all \( z_\alpha \in \mathbb{Z}_+ \) are called **positive**; and those with all \( z_\alpha \in -\mathbb{Z}_+ \) are the **negative** roots (for the given base \( \Delta \)). The partition of \( \Phi \) into positive and negative roots is written as follows:

\[
\Phi = \Phi_+ \sqcup \Phi_-
\]

Thus, \( \Phi_- = -\Phi_+ \). Note also that **B1** amounts to requiring \( \Delta \) to be linearly independent; the fact that \( \Delta \) is a basis of \( \mathbb{E} \) then follows from **B2** and **R1**.

### 7.2.1. Examples: The Classical Root Systems

It is not immediately obvious that every root system \( \Phi \) does in fact have a base. In §7.2.2 we will show that this is indeed the case. Later, we will also see that bases are essentially unique (Theorem 7.9). For now, we exhibit bases for each of the classical root systems \( A_n - D_n \) (§7.1.4).

**Example 7.3** (Base for \( A_n \)). The root system of type \( A_n \) has roots \( \Phi = \{ \pm (\varepsilon_i - \varepsilon_j) \mid 1 \leq i < j \leq n+1 \} \). The subset \( \Delta = \{ \alpha_i := \varepsilon_i - \varepsilon_{i+1} \mid i = 1, 2, \ldots, n \} \) is clearly linearly independent and **B2** has already been verified in (6.27). Thus, \( \Delta \) a base of \( \Phi \) and \( \Phi_+ = \{ \varepsilon_i - \varepsilon_j \mid 1 \leq i < j \leq n+1 \} \) are the positive roots for this base. The picture above, a rotated version of (7.5) with \( \alpha_1 = (1, -1, 0) \) and \( \alpha_2 = (0, 1, -1) \), depicts the root system \( A_2 \), with \( \Phi_+ \) in red. For \( A_3 \), (7.6) shows the base \( \Delta \) in bright red; the non-simple positive roots \( \Phi_+ \) are a duller red while the negative roots are green.

**Example 7.4** (Base for \( B_n \)). Here, the roots are given by \( \pm \varepsilon_i \) \((1 \leq i \leq n)\) and \( \pm \varepsilon_i \pm \varepsilon_j \) \((1 \leq i < j \leq n)\). Putting

\[
\alpha_i := \varepsilon_i - \varepsilon_{i+1} \quad (i = 1, \ldots, n-1) \quad \text{and} \quad \alpha_n := 2\varepsilon_n
\]

we obtain a base \( \Delta = \{ \alpha_1, \ldots, \alpha_n \} \) for \( \Phi \); see (6.36) for the verification of property **B2**. The picture shows the root system \( B_2 \), again with \( \Phi_+ \) in red.

**Example 7.5** (Base for \( C_n \)). The roots in this case are \( \pm 2\varepsilon_i \) \((1 \leq i \leq n)\) and \( \pm \varepsilon_i \pm \varepsilon_j \) \((1 \leq i < j \leq n)\). The roots

\[
\alpha_i := \varepsilon_i - \varepsilon_{i+1} \quad (i = 1, \ldots, n-1) \quad \text{and} \quad \alpha_n := 2\varepsilon_n
\]
are linearly independent and satisfy

\[ 2\varepsilon_i = 2 \sum_{k=1}^{n-1} \alpha_k + \alpha_n, \]

(7.10)

\[ \varepsilon_i - \varepsilon_j = \sum_{k=i}^{j-1} \alpha_k, \]

\[ \varepsilon_i + \varepsilon_j = \sum_{k=i}^{j-1} \alpha_k + 2 \sum_{k=j}^{n-1} \alpha_k + \alpha_n. \]

Thus, \( B2 \) holds and hence \( \Delta = \{ \alpha_1, \ldots, \alpha_n \} \) is a base of \( \Phi \). The roots in (7.10) are the positive roots for this base.

**Example 7.6** (Base for \( D_n \)). Now the roots are \( \pm \varepsilon_i \pm \varepsilon_j \) (\( 1 \leq i < j \leq n \)). A linearly independent subset is given by

\[ \alpha_i := \varepsilon_i - \varepsilon_{i+1} \quad (i = 1, \ldots, n-1) \quad \text{and} \quad \alpha_n := \varepsilon_{n-1} + \varepsilon_n. \]

\( B2 \) also holds:

\[ \varepsilon_i - \varepsilon_j = \sum_{k=i}^{j-1} \alpha_k \quad (1 \leq i < j \leq n), \]

(7.11)

\[ \varepsilon_i + \varepsilon_n = \sum_{k=i}^{n-2} \alpha_k + \alpha_n \quad (1 \leq i < n), \]

\[ \varepsilon_i + \varepsilon_j = \sum_{k=i}^{j-1} \alpha_k + 2 \sum_{k=j}^{n-2} \alpha_k + \alpha_{n-1} + \alpha_n \quad (1 \leq i < j < n). \]

Thus, \( \Delta = \{ \alpha_1, \ldots, \alpha_n \} \) is a base of \( \Phi \) and (7.11) gives the positive roots.

### 7.2.2. Existence

In preparation for the existence proof for bases in general, note that \( \bigcup_{\alpha \in \Phi} \alpha^\perp \) is a proper subset of \( \mathbb{E} \), because \( \mathbb{E} \) is not the union of a finite number of proper subspaces (Exercise 7.2.1). Any element \( \gamma \in \mathbb{E} \setminus \bigcup_{\alpha \in \Phi} \alpha^\perp \) will be called **regular**. Thus, \( \gamma \in \mathbb{E} \) is regular if and only if \( \Phi \cap \gamma^+ = \emptyset \). The hyperplane \( \gamma^\perp \) perpendicular to \( \gamma \) divides \( \mathbb{E} \) into two open half-spaces:

\[ \mathbb{E} = \gamma^\perp \cup \mathbb{E}_+(\gamma) \cup \mathbb{E}_-(\gamma), \]

with \( \mathbb{E}_+(\gamma) = \{ \mu \in \mathbb{E} \mid (\mu, \gamma) > 0 \} \) and \( \mathbb{E}_-(\gamma) = \{ \mu \in \mathbb{E} \mid (\mu, \gamma) < 0 \} = -\mathbb{E}_+(\gamma) \). Letting \( \Phi_+(\gamma) = \Phi \cap \mathbb{E}_+(\gamma) \) denote the set of all roots belonging to the same open half-space as \( \gamma \), we have

\[ \Phi = \Phi_+(\gamma) \cup -\Phi_+(\gamma). \]
An element $\alpha \in \Phi_+(\gamma)$ will be called **indecomposable** if $\alpha$ cannot be written as $\alpha = \beta + \beta'$ with $\beta, \beta' \in \Phi_+(\gamma)$. The picture above shows the root system $B_2$ and a particular choice of regular $\gamma$, with the indecomposable elements of $\Phi_+(\gamma)$ drawn in red.

**Proposition 7.7.** Let $\gamma \in \mathbb{B}$ be regular. Then the set $\Delta = \Delta(\gamma)$ consisting of all indecomposable elements of $\Phi_+(\gamma)$ is a base of $\Phi$ and all bases of $\Phi$ arise in this manner.

**Proof.** First, observe that if $\alpha \in \Phi_+(\gamma)$ can be written as $\alpha = \beta + \beta'$ with $\beta, \beta' \in \Phi_+(\gamma)$ then $(\alpha, \gamma)$ is bigger than each of $(\beta, \gamma)$ and $(\beta', \gamma)$. Since $\Phi_+(\gamma)$ is finite, it follows that each $\alpha \in \Phi_+(\gamma)$ is a finite sum of indecomposables. In view of the decomposition $\Phi = \Phi_+(\gamma) \sqcup -\Phi_+(\gamma)$, this proves $B_2$ for $\Delta$.

Before addressing the linear independence requirement $B_1$, we make the following

**Claim.** If $\alpha, \beta$ are distinct elements of $\Delta$, then $(\alpha, \beta) \leq 0$.

Indeed, if $(\alpha, \beta) > 0$ then $\alpha - \beta$ and $\beta - \alpha$ belong to $\Phi$ (Lemma 7.1). Hence, $\alpha - \beta \in \Phi_+(\gamma)$ or $\beta - \alpha \in \Phi_+(\gamma)$. In the former case, $\alpha = (\alpha - \beta) + \beta$ contradicts indecomposability of $\alpha$, and in the later case, $\beta = (\beta - \alpha) + \alpha$ gives a contradiction. This proves the claim.

Now suppose that there is a non-trivial $\mathbb{R}$-linear dependence among the elements of $\Delta$. Then we obtain an equation $\sum_{\alpha \in \Delta'} r_{\alpha} \alpha = \sum_{\beta \in \Delta''} r_{\beta} \beta$ with positive real coefficients $r_{\alpha}, r_{\beta}$ and disjoint subsets $\Delta', \Delta'' \subseteq \Delta$ with $\Delta' \neq \emptyset$, say. Putting $\mu = \sum_{\alpha \in \Delta'} r_{\alpha} \alpha$, the claim gives $(\mu, \mu) = \sum_{\alpha, \beta} r_{\alpha} r_{\beta} (\alpha, \beta) \leq 0$, which forces $\mu = 0$. On the other hand, $(\mu, \gamma) = \sum_{\alpha} r_{\alpha} (\alpha, \gamma)$ is a positive real number. This contradiction completes the proof that $\Delta$ is a base of $\Phi$.

It remains to show that every base $\Delta$ of $\Phi$ has the form $\Delta = \Delta(\gamma)$ for a suitable $\gamma$. To this end, pick any vector $\gamma \in \mathbb{B}$ such that $(\alpha, \gamma) > 0$ holds for all $\alpha \in \Delta$; this can be always be done (Exercise 7.2.1). Then $\gamma$ is regular by $B_2$, and $\Phi_+ \subseteq \Phi_+(\gamma)$. Since $\Phi_+ \sqcup -\Phi_+ = \Phi = \Phi_+(\gamma) \sqcup -\Phi_+(\gamma)$, we must have $\Phi_+ = \Phi_+(\gamma)$. Clearly, $\Delta$ is exactly the set of indecomposable elements of $\Phi_+ = \Phi_+(\gamma)$; so $\Delta = \Delta(\gamma)$ as desired. $\square$

### 7.2.3. Uniqueness

If $\Delta$ is a base of the root system $\Phi$ and $f \in \text{Aut} \Phi$, then it is trivial to see that $f(\Delta)$ is also a base of $\Phi$. Our next goal is to show that all bases of $\Phi$ arise in this manner from a fixed base $\Delta$ and we may even take $f \in W = W_{\Phi}$.

**A Technical Lemma.** The following lemma collects a number of technicalities concerning the Weyl group that will be needed in the proof of the main result.
Lemma 7.8. Let $\Delta$ be a base of $\Phi$.

(a) For each $\alpha \in \Delta$, the reflection $s_{\alpha}$ permutes the set $\Phi_+ \setminus \{\alpha\}$.

(b) $s_{\alpha}\rho = \rho - \alpha$ for all $\alpha \in \Delta$.

(c) Assume that $1 \neq w \in W$ has the form $w = s_{\alpha_1} s_{\alpha_2} \ldots s_{\alpha_r}$, with $\alpha_i \in \Delta$ (not necessarily distinct) and that $t$ is chosen minimal. Then $w\alpha_t \in \Phi_-$.

Proof. (a) Each $\beta \in \Phi_+ \setminus \{\alpha\}$ can be uniquely written as $\beta = \sum_{\delta \in \Delta} z_\delta \delta$ with $z_\delta \in \mathbb{Z}_+$ for all $\delta$ and we must have $z_\delta > 0$ for some $\delta \neq \alpha$ by R2. The root

$$s_{\alpha} \beta = \beta - \langle \beta, \alpha \rangle \alpha = \sum_{\delta \in \Delta \setminus \{\alpha\}} z_\delta \delta + (z_{\alpha} - \langle \beta, \alpha \rangle) \alpha$$

has the same positive coefficient $z_\delta$, whence $s_{\alpha} \beta \in \Phi_+ \setminus \{\alpha\}$. This proves (a).

(b) follows immediately from (a) and the equation $s_{\alpha} \alpha = -\alpha$.

(c) Suppose, to the contrary, that $w\alpha_t \in \Phi_+$. Then, writing $s_t = s_{\alpha_t}$ and observing that $w\alpha_t = s_is_{i+1} \ldots s_{t-1}(\alpha_t)$, we obtain $s_is_{i+1} \ldots s_{t-1}(\alpha_t) \in \Phi_-$ for $0 \leq i \leq t - 1$, we have $\beta_0 \in \Phi_-$ and $\beta_{t-1} = \alpha_t \in \Phi_+$. Choose $s$ so that $\beta_s = \beta_s$, but $\beta_{s-1} = s_{\alpha_s} \in \Phi_+$. Then we know from (a) that $\beta_s = \alpha_s$. Now put $u = s_{s+1} \ldots s_{t-1}$; so $u\alpha_t = \beta_s = \alpha_s$. Then $u_{\alpha_t}^{-1} = s_{\alpha_t}^{-1} s_t^{-1} \ldots s_{s+1}^{-1}$ by Lemma 7.2(a), and this further implies

$$w = s_is_{i+1} \ldots s_{t-1}(\alpha_t) = s_{s+1} \ldots s_{t-1} s_t u_{\alpha_t}^{-1} = s_{s+1} \ldots s_{t-1} s_t u,$$

contradicting minimality of $t$. This completes the proof of the lemma. \hfill \Box

**Weyl Chambers.** It will be useful to reprise some ideas from §7.2.2. Recall that a vector $\gamma \in \mathbb{B}$ is said to be regular if $\gamma$ belongs to $\mathbb{B}^\circ := \mathbb{B} \setminus \bigcup_{\alpha \in \Phi} \alpha^\perp$. In this case, we have put $\Phi_+(\gamma) = \{\alpha \in \Phi \mid \langle \alpha, \gamma \rangle > 0\}$, the set of all roots in the same half-space as $\gamma$, and $\Delta(\gamma)$ was defined as the set of indecomposable elements of $\Phi_+(\gamma)$. The connected components of $\mathbb{B}^\circ$ are called (open) **Weyl chambers** and the hyperplanes $\alpha^\perp$ for $\alpha \in \Phi$ are often referred to as the **walls**. The Weyl chambers are the fibers of the continuous map $\mathbb{B}^\circ \to \{\pm\}^\Phi$, $\gamma \mapsto (\text{sgn}(\alpha, \gamma))_{\alpha \in \Phi}$; they partition $\mathbb{B}^\circ$ into convex open subsets. Thus, each regular $\gamma$ belongs to exactly one Weyl chamber, which will be denoted by $\mathcal{C}(\gamma)$, and $\mathcal{C}(\gamma) = \mathcal{C}(\gamma')$ if and only if $\langle \alpha, \gamma \rangle$ and $\langle \alpha, \gamma' \rangle$
have the same sign for all \( \alpha \in \Phi \). Therefore,

\[
\mathcal{C}(\gamma) = \mathcal{C}(\gamma') \iff \Phi_+(\gamma) = \Phi_+(\gamma') \iff \Delta(\gamma) = \Delta(\gamma').
\]

In view of Proposition 7.7, this says that the Weyl chambers are in bijection with the set of bases of \( \Phi \):

\[
\begin{array}{c}
\{ \text{bases of } \Phi \} \\
\sim \\
\Delta = \Delta(\gamma) \\
\mathcal{C}(\Delta) = \mathcal{C}(\gamma)
\end{array}
\]

The Weyl chamber \( \mathcal{C}(\Delta) \) is called the fundamental Weyl chamber for the base \( \Delta \); it is explicitly given by

\[
\begin{align*}
\mathcal{C}(\Delta) &= \{ \beta \in \mathbb{E} \mid (\beta, \alpha) > 0 \text{ for all } \alpha \in \Phi_+ \} \\
&= \{ \beta \in \mathbb{E} \mid (\beta, \alpha) > 0 \text{ for all } \alpha \in \Delta \}.
\end{align*}
\]

The picture on the right shows \( \mathcal{C} = \mathcal{C}(\Delta) \) for \( A_2 \) with base \( \Delta = \{ \alpha_1, \alpha_2 \} \) as in Example 7.3.

**Bases and the Weyl group.** Since \( W \) stabilizes \( \Phi \) and consists of orthogonal transformations of \( \mathbb{E} \), the hyperplanes \( \alpha \perp (\alpha \in \Phi) \) are permuted by \( W \), and hence so are the Weyl chambers. Explicitly, if \( \gamma \) is regular and \( w \in W \), then \( w\gamma \) is regular as well and \( \mathcal{C}(w\gamma) = w\mathcal{C}(\gamma) \). Similarly, \( w\Phi_+(\gamma) = \Phi_+(w\gamma) \) and so \( w\Delta(\gamma) = \Delta(w\gamma) \). Thus the bijection (7.13) is equivariant for the action of \( W \) on both sets.

**Theorem 7.9.** Let \( \Phi \subseteq \mathbb{E} \) be a root system, \( \Delta \) a base of \( \Phi \), and \( W' = W_\Phi \) the Weyl group. Then:

(a) \( W' \) acts simply transitively on the set of all bases of \( \Phi \): they are exactly the subsets of the form \( w\Delta \) with \( w \in W' \) and \( w\Delta = \Delta \) forces \( w = 1 \).

(b) Each \( W' \)-orbit in \( \Phi \) meets \( \Delta \).

(c) \( W' \) is generated by the reflections \( s_\alpha \) with \( \alpha \in \Delta \). Moreover, for each \( w \in W' \), the minimal \( \ell \) such that \( w = s_{\alpha_1}s_{\alpha_2} \ldots s_{\alpha_\ell} \) with \( \alpha_i \in \Delta \) (not necessarily distinct) is equal to \( \# \{ \alpha \in \Phi_+ \mid w\alpha \in \Phi_- \} \).

Just as the roots \( \alpha \in \Delta \) are called simple, the corresponding reflections \( s_\alpha \) are frequently referred to as simple reflections. Thus, (c) states that \( W' \) is generated by the simple reflections. The number \( \ell = \ell(w) = \# \{ \alpha \in \Phi_+ \mid w\alpha \in \Phi_- \} \) in (c) is called the length of the element \( w \in W' \).

**Proof of Theorem 7.9.** Let \( W \) denote the subgroup of \( W' \) that is generated by the simple reflections \( s_\alpha \) (\( \alpha \in \Delta \)). We will show that the key assertions in the theorem all hold for \( W \) in place of \( W' \) and then use this fact to conclude that \( W = W' \).
Step 1: \( W \) acts simply transitively on the set of all bases of \( \Phi \). Given some base \( \Delta(\gamma) \) of \( \Phi \), with \( \gamma \in \mathbb{B} \) regular, choose \( w \in W \) such that \((w\gamma, \rho)\) is maximal. Then, using orthogonality of reflections (Exercise 7.1.1) and Lemma 7.8(b), we obtain \((w\gamma, \rho) \geq (s_{\rho}w\gamma, \rho) = (w\gamma, s_{\rho}\rho) = (w\gamma, \rho) - (w\gamma, \rho)\) for all \( \alpha \in \Delta \). Hence, \((w\gamma, \alpha) \geq 0\) and this inequality is in fact strict, since \( w\gamma \) is regular. Therefore, \( w\gamma \in \mathcal{C}(\Delta) \) and so \( w^{-1}\mathcal{C}(\gamma) = \mathcal{C}(\Delta) \) and \( w\Delta(\gamma) = \Delta \). This proves that \( W \) acts transitively on the set of all bases of \( \Phi \). Now suppose that \( 1 \neq w \in W \) satisfies \( w\Delta = \Delta \) and write \( w = s_{\alpha_1}s_{\alpha_2}\ldots s_{\alpha_t} \) with \( \alpha_i \in \Delta \) and \( t \) minimal. We know from Lemma 7.8 that \( w\alpha_i \in \Phi_+ \). But this contradicts the fact that \( w\alpha_i \notin \Delta \), thereby finishing Step 1.

Step 2: Each \( W \)-orbit in \( \Phi \) meets \( \Delta \). In view of Step 1, it suffices to show that each \( \alpha \in \Phi \) belongs to some \( \Delta(\gamma) \) with \( \gamma \in \mathbb{B} \) regular. To prove this, note that if \( \beta \in \Phi \) satisfies \( \beta^\perp = \alpha^\perp \), then \( \mathbb{R}\beta = \beta^\perp = \alpha^\perp = \mathbb{R}\alpha \) and so \( \beta = \pm \alpha \) by R2. Therefore, \( \alpha^\perp \notin \bigcup_{t \neq \pm \alpha} \beta^\perp \) (Exercise 7.2.1); so there exists \( \lambda \in \mathbb{B} \) with \( (\lambda, \alpha) = 0 \) but \( (\lambda, \beta) \neq 0 \) for all \( \beta \in \Phi \setminus \{\pm \alpha\} \). Choosing \( \gamma \) close enough to \( \lambda \), we can arrange that \( 0 < (\gamma, \alpha) < \min\{|(\gamma, \beta)| \mid \beta \in \Phi \setminus \{\pm \alpha\}\} \). Then \( \gamma \) is regular and \( \alpha \) is clearly an indecomposable element of \( \Phi_+(\gamma) \). Therefore, \( \alpha \in \Delta(\gamma) \).

Step 3: \( W = \mathcal{W} \). Given \( \alpha \in \Phi \), we know by Step 2 that \( w\alpha \in \Delta \) for some \( w \in W \). Hence, \( ws_{\alpha}w^{-1} = s_{w\alpha} \in W \) and so \( s_{\alpha} \in W \). Since the reflections \( s_{\alpha} \) generate \( \mathcal{W} \), it follows that \( \mathcal{W} = W \).

Step 4: Length. For each \( w \in \mathcal{W} \), put \( N(w) = \{\alpha \in \Phi_+ \mid w\alpha \in \Phi_-\} \). We will show by induction on \( \ell(w) \) that \( \ell(w) = \#N(w) \). This is clear for \( \ell(w) = 0 \); in this case, \( w = 1 \). Now assume that \( \ell = \ell(w) > 0 \) and that the equality holds for all shorter elements of \( \mathcal{W} \). Writing \( w = s_{\alpha_1}s_{\alpha_2}\ldots s_{\alpha_t} \) as in (c), we have \( \ell(ws_{\alpha_t}) = \ell - 1 \) and so \( \#N(ws_{\alpha_t}) = \ell - 1 \) as well. From Lemma 7.8 we further know that \( s_{\alpha_t} \) permutes \( \Phi_+ \setminus \{\alpha_t\} \) and that \( w\alpha_t \notin \Phi_+ \). It follows that \( N(ws_{\alpha_t}) = N(w) \setminus \{\alpha_t\} \). Therefore, \( \#N(w) = \ell \), completing the proof of the theorem.  \( \square \)

**Example 7.10** (Length of elements in \( S_{n+1} \)). Consider the root system \( \Phi \) of type \( A_n \) and use the base \( \Delta = \{\varepsilon_i - \varepsilon_{i+1} \mid i = 1, \ldots, n\} \) as in Example 7.3; so \( \Phi_+ = \{\varepsilon_i - \varepsilon_j \mid 1 \leq i < j \leq n + 1\} \). Under the isomorphism \( \mathcal{W} \cong S_{n+1} \) in (7.4), the reflection \( s_{\varepsilon_i} \in \mathcal{W} \) corresponds to the transposition \( s_i = (i, i+1) \in S_{n+1} \). Thus, Theorem 7.9 yields the standard fact that \( s_1, \ldots, s_n \) generate \( S_{n+1} \). Moreover, viewing the isomorphism \( \mathcal{W} \cong S_{n+1} \) as an identification, we have the following description of the set \( N(w) \) in Step 4 above:

\[
\{\alpha \in \Phi_+ \mid w\alpha \in \Phi_-\} \; \xrightarrow{\sim} \; \{(i, j) \in [n + 1]^2 \mid i < j \text{ but } wi > wj\}
\]

\[
\varepsilon_i - \varepsilon_j \; \xrightarrow{\sim} \; (i, j)
\]
The pairs \((i, j) \in [n + 1]^2\) on the right are called the inversions of the permutation \(w \in S_{n+1}\). Thus, the length \(\ell(w)\) of any \(w \in S_{n+1}\) in terms of the generators \(s_1, \ldots, s_n\) is equal to the number of inversions of \(w\).

### Exercises for Section 7.2

#### 7.2.1 (Regular vectors). Prove:

(a) \(E\) is not a finite union of proper subspaces.

(b) If \(B\) is an \(\mathbb{R}\)-basis of \(E\), then there exists \(\gamma \in E\) with \((\beta, \gamma) > 0\) for all \(\beta \in B\).

#### 7.2.2 (Height reduction). Let \(\Delta\) be a base of \(\Phi\). The height of a root \(\beta = \sum\alpha \in \Delta \alpha \in \Phi\) (relative to \(\Delta\)) is the integer \(\sum_{\alpha \in \Delta} z_\alpha\); so simple roots are the same as roots of height 1. Now let \(\beta \in \Phi_+ \setminus \Delta\). Show that \(\beta - \alpha \in \Phi_+ \) for some \(\alpha \in \Delta\).

#### 7.2.3 (Bases of dual root systems). Let \(\Phi^\vee = \{\alpha^\vee = \frac{2\alpha}{\langle \alpha, \alpha \rangle} \mid \alpha \in \Phi\}\) be the dual root system (Exercise 7.1.3) and let \(\Delta\) be a base of \(\Phi\). Show:

(a) The Weyl chambers of \(\Phi^\vee\) are the same as those of \(\Phi\).

(b) \(\Delta^\vee := \{\alpha^\vee \mid \alpha \in \Delta\}\) is a base of \(\Phi^\vee\).

(c) If \(\mu \in E\) satisfies \(\langle \mu, \alpha \rangle \in \mathbb{Z}\) for all \(\alpha \in \Delta\), then \(\langle \mu, \beta \rangle \in \mathbb{Z}\) for all \(\beta \in \Phi\).

#### 7.2.4 (Decomposition of \(\text{Aut} \Phi\)). (a) Let \(\Phi' \subseteq \Phi\) be a root system that is isomorphic to \(\Phi\) via \(f : E \rightarrow \mathbb{E}\)', \(f(\Phi) = \Phi\)' and let \(\Delta\) be a base of \(\Phi\). Show that \(f\Delta\) is a base of \(\Phi\)'.

(b) For a given base \(\Delta \subseteq \Phi\), put \(I_\Delta = \{f \in \text{Aut} \Phi \mid f\Delta = \Delta\}\). Show that \(\text{Aut} \Phi = \mathcal{W} \rtimes I_\Delta\).

#### 7.2.5 (Rank 2 root systems). Use Exercise 7.2.4 to determine \(\text{Aut} \Phi\) for \(\Phi = A_1 \times A_1, A_2, B_2, G_2\).

#### 7.2.6 (The longest element of \(\mathcal{W}\)). Let \(\Delta\) be a base of \(\Phi\). Since \(-\Delta\) is clearly a base of \(\Phi\) as well, there is a unique \(w_0 \in \mathcal{W}\) such that \(w_0\Delta = -\Delta\). Prove:

(a) \(\ell(w_0) = \#\Phi_+ > \ell(w)\) for all \(w \in \mathcal{W} \setminus \{w_0\}\). In particular, \(w_0^{-1} = w_0\).

(b) \(\ell(w_0w) = \ell(w_0) - \ell(w)\) for all \(w \in \mathcal{W}\).

(c) Determine \(w_0 \in S_{n+1}\) for \(\Phi\) of type \(A_n\) and \(\Delta\) as in Example 7.3.

### 7.3. Classification

We keep the notations from the previous sections. In particular, \(\Phi \subseteq \mathbb{E} \cong \mathbb{R}^n\) will continue to denote a root system throughout this section.

#### 7.3.1. Irreducibility

A root system \(\Phi \neq \emptyset\) is said to be irreducible (as in §6.3.3) if \(\Phi\) cannot be partitioned into two nonempty subsets that are orthogonal to each other: it is not possible to write \(\Phi = \Phi_1 \sqcup \Phi_2\) with \(\Phi_1 \neq \emptyset\) and \((\Phi_1, \Phi_2) = \{0\}\) or, equivalently, \((\Phi_1, \Phi_2) = \{0\}\).
Proposition 7.11. Let \( \Phi \neq \emptyset \) and let \( \Delta \) be a base of \( \Phi \). Then the following are equivalent:

(i) \( \Phi \) is irreducible;

(ii) \( \Delta \) cannot be partitioned as a disjoint union of nonempty subsets that are orthogonal to each other;

(iii) \( \mathbb{E} \) is an irreducible representation of \( \mathcal{W}_q \) over \( \mathbb{R} \).

Proof. First observe that if \( 0 \neq \alpha, \beta \in \mathbb{E} \) are orthogonal to each other, then the reflection \( s_\alpha \) fixes \( \beta \) and \( s_\beta \) fixes \( \alpha \). Furthermore, \( s_\alpha s_\beta \mu = \mu - \langle \mu, \alpha \rangle \alpha - \langle \mu, \beta \rangle \beta \) for any \( \mu \in \mathbb{E} \); so \( s_\alpha s_\beta = s_\beta s_\alpha \).

Now assume that \( \mathbb{E} \) is an irreducible representation of \( \mathcal{W} = \mathcal{W}_q \) but \( \Delta = \Delta_1 \sqcup \Delta_2 \) with \( \Delta_i \neq \emptyset \) and \( \Delta_1 \perp \Delta_2 \), that is, \( (\Delta_1, \Delta_2) = \{0\} \). Letting \( \mathcal{E}_i \) denote the \( \mathbb{R} \)-span of \( \Delta_i \) in \( \mathbb{E} \), we have

\[
\mathbb{E} = \mathcal{E}_1 \oplus \mathcal{E}_2 \quad \text{and} \quad \mathcal{E}_1 \perp \mathcal{E}_2 .
\]

Define \( \mathcal{W}_i \) to be the subgroup of \( \mathcal{W} \) that is generated by the reflections \( s_\alpha \) with \( \alpha \in \Delta_i \). Then each \( \mathcal{W}_i \) stabilizes \( \mathcal{E}_i \). Furthermore, since \( \mathcal{W}_1 \) and \( \mathcal{W}_2 \) generate \( \mathcal{W} \) by Theorem 7.9, our observation above gives

\[
\mathcal{W} = \mathcal{W}_1 \times \mathcal{W}_2 \quad \text{with} \quad \mathcal{W}_1|_{\mathcal{E}_2} = \text{Id}, \quad \mathcal{W}_2|_{\mathcal{E}_1} = \text{Id} .
\]

Therefore, both summands \( \mathcal{E}_i \) are \( \mathcal{W} \)-subrepresentations of \( \mathbb{E} \), contradicting irreducibility. This proves the implication (iii) \( \Rightarrow \) (ii).

For (ii) \( \Rightarrow \) (i), just note that any partition \( \Phi = \Phi_1 \sqcup \Phi_2 \) with \( \Phi_1 \neq \emptyset \) and \( \Phi_1 \perp \Phi_2 \) gives rise to the partition \( \Delta = \Delta_1 \sqcup \Delta_2 \) with \( \Delta_i = \Delta \cap \Phi_i \). Now (ii) forces \( \Delta_2 = \emptyset \), say, and hence \( \Delta \perp \Phi_2 \). Since \( \Delta \) spans \( \mathbb{E} \), it follows that \( \mathbb{E} \perp \Phi_2 \), which is absurd.

Finally, assume that there is a \( \mathcal{W} \)-stable subspace \( 0 \neq V \subsetneq \mathbb{E} \). Then \( V^\perp \) is also \( \mathcal{W} \)-stable, because \( \mathcal{W} \) consists of orthogonal transformations of \( \mathbb{E} \), and \( \mathbb{E} = V \oplus V^\perp \). Let \( \alpha \in \Phi \) be given and assume that \( \alpha \notin V \). Since \( s_\alpha \mu = \mu - \langle \mu, \alpha \rangle \alpha \in V \) for \( \mu \in V \), we must have \( \langle \mu, \alpha \rangle = 0 \) and so \( \alpha \in V^\perp \). Thus, we obtain the orthogonal partition \( \Phi = (\Phi \cap V) \sqcup (\Phi \cap V^\perp) \) and neither part is empty, because \( \Phi \) spans \( \mathbb{E} \). This proves the implication (i) \( \Rightarrow \) (iii).

For an arbitrary root system \( \Phi \subseteq \mathbb{E} \), define an equivalence relation \( \sim \) by declaring \( \alpha \sim \beta \) if there is a chain \( \alpha = \alpha_1, \alpha_2, \ldots, \alpha_r = \beta \) with \( \alpha_i \in \Phi \) and \( (\alpha_i, \alpha_{i+1}) \neq 0 \) for all \( i \). The collection \( \{\Phi_i\} \) of \( \sim \)-equivalence classes of \( \Phi \) gives a partition of \( \Phi \) into nonempty subsets that are pairwise orthogonal to each other.

Let \( \mathcal{E}_i \) denote the \( \mathbb{R} \)-subspace of \( \mathbb{E} \) that is spanned by \( \Phi_i \) and let \( \mathcal{W}_i \) denote the subgroup of \( \mathcal{W}_q \) that is generated by the reflections \( s_\alpha \) with \( \alpha \in \Phi_i \). As in the proof of Proposition 7.11, one sees that \( \mathbb{E} = \bigoplus_i \mathcal{E}_i \) with \( \mathcal{E}_i \perp \mathcal{E}_j \) for \( i \neq j \) and

\[
(7.15) \quad \mathcal{W}_q = \prod_i \mathcal{W}_i \quad \text{with} \quad \mathcal{W}_i|_{\mathcal{E}_j} = \text{Id} \quad (i \neq j) .
\]
It follows that \( \Phi_i \) is an irreducible root system in \( \mathbb{E}_i \) with Weyl group \( W_{\Phi_i} \cong W_i \). The \( \Phi_i \) are called the **irreducible components** of the root system \( \Phi \).

### 7.3.2. Cartan Matrix and Dynkin Diagram

Fix a base \( \Delta \) of \( \Phi \) and choose an ordering of the simple roots, say they are \( \alpha_1, \ldots, \alpha_n \). The **Cartan matrix** of \( \Phi \) is the matrix

\[
C = \left( \langle \alpha_i, \alpha_j \rangle \right)_{i,j} \in \text{Mat}_n(\mathbb{Z}).
\]

Of course, \( C \) depends on the chosen ordering of \( \Delta \), but the choice of the base \( \Delta \) is insubstantial thanks to Theorem 7.9(a). We make the following remarks:

- **The diagonal entries of \( C \) are all 2; the other entries are 0, –1, –2 or –3:** The assertion about the diagonal entries is clear. From the claim in the proof of Proposition 7.7, we also know that if \( \alpha, \beta \) are distinct elements of \( \Delta \) then \( \langle \alpha, \beta \rangle \leq 0 \). The values of the off-diagonal entries therefore follow from Table 7.1.

- **The determinant of \( C \) is a strictly positive integer:** Up to a rescaling of the columns by the positive factors \( \frac{2}{\langle \alpha_j, \alpha_j \rangle} \), the matrix \( C \) is the matrix of the inner product \( (\cdot, \cdot) \) for the \( \mathbb{R} \)-basis \( \Delta \) of \( \mathbb{E} \). Since \( (\cdot, \cdot) \) is positive definite symmetric, the determinant of this matrix is strictly positive.

The information contained in the Cartan matrix can also be graphically recorded in the so-called **Dynkin diagram** of \( \Phi \). This is a graph with \( n \) vertices, one for each root vector \( \alpha_i \), and the \( i^{th} \) and \( j^{th} \) vertex are connected by \( \langle \alpha_i, \alpha_j \rangle \langle \alpha_j, \alpha_i \rangle \) edges. Thus, from Table 7.1 we know that there are 0, 1, 2 or 3 edges between any two vertices. Specifically, if there are no edges between the \( i^{th} \) and \( j^{th} \) vertex, then this means that \( \alpha_i \) and \( \alpha_j \) are orthogonal; there is one edge iff \( \alpha_i \) and \( \alpha_j \) are non-orthogonal with \( ||\alpha_i|| = ||\alpha_j|| \); and multiple edges occur iff \( \alpha_i \) and \( \alpha_j \) are non-orthogonal of different lengths. To specify relative lengths, an arrow pointing to the vertex for the shorter root is added to any \( m \)-fold edge with \( m > 1 \). If the arrow points to the \( i^{th} \) vertex, say, then \( \langle \alpha_i, \alpha_j \rangle = -1 \) and \( \langle \alpha_j, \alpha_i \rangle = -m \) (Table 7.1). In this way, one can recover the Cartan matrix from the Dynkin diagram.

Note that condition (ii) in Proposition 7.11 is equivalent to connectedness of the Dynkin diagram of \( \Phi \) in the usual graph theoretical sense. Thus:

\[
(7.16) \quad \text{\( \Phi \) is irreducible if and only if the Dynkin diagram of \( \Phi \) is connected}
\]

In general, the connected components of the Dynkin diagram of \( \Phi \) correspond to the irreducible components of \( \Phi \). From the bases exhibited in Examples 7.3 – 7.6, it is straightforward to obtain the Cartan matrices and Dynkin diagrams of the classical (irreducible) root systems \( \mathbf{A}_n \) – \( \mathbf{D}_n \). They are displayed in Table 7.2; see also (6.30) for the Cartan matrix of \( \mathbf{A}_n \).
7.3. Classification

We have seen in §7.3.2, the Cartan matrix and the Dynkin diagram of a root system determine each other. Moreover, either one determines the entire root system up to isomorphism:

Proposition 7.12. Let $\Phi \subseteq \mathbb{E}$ and $\Phi' \subseteq \mathbb{E}'$ be root systems with respective bases $\Delta = \{\alpha_1, \ldots, \alpha_n\}$ and $\Delta' = \{\alpha'_1, \ldots, \alpha'_n\}$. If $\langle \alpha_i, \alpha_j \rangle = \langle \alpha'_i, \alpha'_j \rangle$ for all $i$ and $j$, then the $\mathbb{R}$-linear isomorphism $f : \mathbb{E} = \bigoplus_i \mathbb{R} \alpha_i \to \mathbb{E}' = \bigoplus_i \mathbb{R} \alpha'_i$ that is given by $f \alpha_i = \alpha'_i$ satisfies $f \Phi = \Phi'$. Thus, root systems with identical Cartan matrices are isomorphic.

Proof. The map $f$ is indeed an $\mathbb{R}$-linear isomorphism, since bases of root systems are $\mathbb{R}$-bases of the ambient Euclidean space by B1. Furthermore, by (7.3) our hypothesis $\langle \alpha_i, \alpha_j \rangle = \langle \alpha'_i, \alpha'_j \rangle$ is equivalent to $f(s_{\alpha_j} \alpha_i) = s_{\alpha'_j} \alpha'_i$ for all $i$ and $j$, which in turn states that $s_{\alpha'_j} = f s_{\alpha_j} f^{-1}$ for all $j$. Since the reflections $s_{\alpha_j}$ and $s_{\alpha'_j}$ generate the respective Weyl groups $W = \mathcal{W}_\Phi$ and $W' = \mathcal{W}_{\Phi'}$ (Theorem 7.9), we obtain a group isomorphism $W \simeq W'$, $w \mapsto f w f^{-1}$. Finally, $\Phi = W \Delta$ and likewise for $\Phi'$ (Theorem 7.9 again). Therefore, $f \Phi = f' W \Delta = f' W f^{-1} \Delta' = W' \Delta' = \Phi'$ as desired.

The converse of Proposition 7.12 is clear: isomorphic roots systems $\Phi$ and $\Phi'$ have identical Cartan matrices, because any isomorphism $\Phi \simeq \Phi'$ preserves $\langle \cdot, \cdot \rangle$ (Lemma 7.2) and sends a base of $\Phi$ to a base of $\Phi'$. Thus, in order to

<table>
<thead>
<tr>
<th>$A_n$</th>
<th>$B_n$</th>
</tr>
</thead>
</table>
| \[
\begin{pmatrix}
2 & -1 & -1 & 0 \\
-1 & 2 & -1 & -2 \\
-1 & -1 & 2 & -1 \\
\vdots & \vdots & \vdots & \vdots \\
-1 & -1 & -1 & 2 \\
\end{pmatrix}
\] | \[
\begin{pmatrix}
2 & -1 & -1 & 0 \\
-1 & 2 & -1 & -2 \\
-1 & -1 & 2 & -1 \\
\vdots & \vdots & \vdots & \vdots \\
-1 & -1 & -1 & 2 \\
\end{pmatrix}
\] |

<table>
<thead>
<tr>
<th>$C_n$</th>
<th>$D_n$</th>
</tr>
</thead>
</table>
| \[
\begin{pmatrix}
2 & -1 & -1 & 0 \\
-1 & 2 & -1 & -2 \\
-1 & -1 & 2 & -1 \\
\vdots & \vdots & \vdots & \vdots \\
-1 & -1 & -1 & 2 \\
\end{pmatrix}
\] | \[
\begin{pmatrix}
2 & -1 & -1 & 0 \\
-1 & 2 & -1 & -2 \\
-1 & -1 & 2 & -1 \\
\vdots & \vdots & \vdots & \vdots \\
-1 & -1 & -1 & 2 \\
\end{pmatrix}
\] |

Table 7.2. Cartan matrices and Dynkin diagrams of the classical root systems

7.3.3. Classification Theorem

As we have seen in §7.3.2, the Cartan matrix and the Dynkin diagram of a root system determine each other. Moreover, either one determines the entire root system up to isomorphism:

Proposition 7.12. Let $\Phi \subseteq \mathbb{E}$ and $\Phi' \subseteq \mathbb{E}'$ be root systems with respective bases $\Delta = \{\alpha_1, \ldots, \alpha_n\}$ and $\Delta' = \{\alpha'_1, \ldots, \alpha'_n\}$. If $\langle \alpha_i, \alpha_j \rangle = \langle \alpha'_i, \alpha'_j \rangle$ for all $i$ and $j$, then the $\mathbb{R}$-linear isomorphism $f : \mathbb{E} = \bigoplus_i \mathbb{R} \alpha_i \to \mathbb{E}' = \bigoplus_i \mathbb{R} \alpha'_i$ that is given by $f \alpha_i = \alpha'_i$ satisfies $f \Phi = \Phi'$. Thus, root systems with identical Cartan matrices are isomorphic.

Proof. The map $f$ is indeed an $\mathbb{R}$-linear isomorphism, since bases of root systems are $\mathbb{R}$-bases of the ambient Euclidean space by B1. Furthermore, by (7.3) our hypothesis $\langle \alpha_i, \alpha_j \rangle = \langle \alpha'_i, \alpha'_j \rangle$ is equivalent to $f(s_{\alpha_j} \alpha_i) = s_{\alpha'_j} \alpha'_i$ for all $i$ and $j$, which in turn states that $s_{\alpha'_j} = f s_{\alpha_j} f^{-1}$ for all $j$. Since the reflections $s_{\alpha_j}$ and $s_{\alpha'_j}$ generate the respective Weyl groups $W = \mathcal{W}_\Phi$ and $W' = \mathcal{W}_{\Phi'}$ (Theorem 7.9), we obtain a group isomorphism $W \simeq W'$, $w \mapsto f w f^{-1}$. Finally, $\Phi = W \Delta$ and likewise for $\Phi'$ (Theorem 7.9 again). Therefore, $f \Phi = f' W \Delta = f' W f^{-1} \Delta' = W' \Delta' = \Phi'$ as desired.

The converse of Proposition 7.12 is clear: isomorphic roots systems $\Phi$ and $\Phi'$ have identical Cartan matrices, because any isomorphism $\Phi \simeq \Phi'$ preserves $\langle \cdot, \cdot \rangle$ (Lemma 7.2) and sends a base of $\Phi$ to a base of $\Phi'$. Thus, in order to
classify root systems up to isomorphism, it suffices to classify their Cartan matrices (up to a labeling of the bases) or, equivalently, their Dynkin diagrams (up to an isomorphism of diagrams). Furthermore, by the discussion in §7.3.1 and §7.3.2, it is enough to classify the connected Dynkin diagrams. This task can be achieved by elementary methods from Euclidean geometry, but the proof is rather lengthy and not particularly germane to the main theme of representation theory. Therefore, we omit it here and only state the result for the sake of completeness. In turns out that, in addition to the four “classical” series displayed in Table 7.2, there are five “exceptional” Dynkin diagrams; they are shown in Table 7.3 along with the corresponding Cartan matrices. The classical Dynkin diagrams have already been realized as coming from a root system with the same label (§7.1.4). A root system for the exceptional diagram $G_2$ was displayed in Figure 7.1. For the construction of an actual root system for each of the remaining exceptional Dynkin diagrams as well as for the proof of the classification theorem, the reader is referred to the literature on Lie algebras (e.g., Humphreys [101, 11.4 and 12.1]).

Table 7.3. Cartan matrices and Dynkin diagrams of the exceptional root systems

<table>
<thead>
<tr>
<th>Root System</th>
<th>Cartan Matrix</th>
</tr>
</thead>
<tbody>
<tr>
<td>$E_6$</td>
<td>$\begin{pmatrix} 2 &amp; 0 &amp; -1 &amp; 0 &amp; 0 &amp; 0 \ 0 &amp; 2 &amp; 0 &amp; -1 &amp; 0 &amp; 0 \ -1 &amp; 0 &amp; 2 &amp; -1 &amp; 0 &amp; 0 \ 0 &amp; 0 &amp; 0 &amp; -1 &amp; 2 &amp; -1 \ 0 &amp; 0 &amp; 0 &amp; 0 &amp; -1 &amp; 2 \end{pmatrix}$</td>
</tr>
<tr>
<td>$E_7$</td>
<td>$\begin{pmatrix} 2 &amp; 0 &amp; -1 &amp; 0 &amp; 0 &amp; 0 &amp; 0 \ 0 &amp; 2 &amp; 0 &amp; -1 &amp; 0 &amp; 0 &amp; 0 \ -1 &amp; 0 &amp; 2 &amp; -1 &amp; 0 &amp; 0 &amp; 0 \ 0 &amp; -1 &amp; -1 &amp; 2 &amp; -1 &amp; 0 &amp; 0 \ 0 &amp; 0 &amp; 0 &amp; -1 &amp; 2 &amp; -1 &amp; 0 \ 0 &amp; 0 &amp; 0 &amp; 0 &amp; -1 &amp; 2 &amp; -1 \end{pmatrix}$</td>
</tr>
<tr>
<td>$E_8$</td>
<td>$\begin{pmatrix} 2 &amp; 0 &amp; -1 &amp; 0 &amp; 0 &amp; 0 &amp; 0 &amp; 0 \ 0 &amp; 2 &amp; 0 &amp; -1 &amp; 0 &amp; 0 &amp; 0 &amp; 0 \ -1 &amp; 0 &amp; 2 &amp; -1 &amp; 0 &amp; 0 &amp; 0 &amp; 0 \ 0 &amp; -1 &amp; -1 &amp; 2 &amp; -1 &amp; 0 &amp; 0 &amp; 0 \ 0 &amp; 0 &amp; 0 &amp; -1 &amp; 2 &amp; -1 &amp; 0 &amp; 0 \ 0 &amp; 0 &amp; 0 &amp; 0 &amp; -1 &amp; 2 &amp; -1 &amp; 0 \ 0 &amp; 0 &amp; 0 &amp; 0 &amp; 0 &amp; -1 &amp; 2 &amp; -1 \end{pmatrix}$</td>
</tr>
<tr>
<td>$F_4$</td>
<td>$\begin{pmatrix} 2 &amp; -1 &amp; 0 &amp; 0 \ -1 &amp; 2 &amp; -2 &amp; 0 \ 0 &amp; -1 &amp; 2 &amp; -2 \ 0 &amp; 0 &amp; 0 &amp; -1 \end{pmatrix}$</td>
</tr>
<tr>
<td>$G_2$</td>
<td>$\begin{pmatrix} 2 &amp; -1 \ -3 &amp; 2 \end{pmatrix}$</td>
</tr>
</tbody>
</table>

Here now, for the record, is the statement of the classification theorem; the restrictions on the rank for the classical types are imposed so as to avoid duplication.

**Theorem 7.13.** Let $\Phi$ be an irreducible root system of rank $n$. Then $\Phi$ is either isomorphic to one of the classical root systems $A_n$ ($n \geq 1$), $B_n$ ($n \geq 2$), $C_n$ ($n \geq 3$) or $D_n$ ($n \geq 4$) or else to one of the exceptional root systems $E_6$, $E_7$, $E_8$, $F_4$ or $G_2$ (value of $n$ indicated by the subscript).
Exercises for Section 7.3

7.3.1 (Some steps towards the Classification Theorem). Let \( \Gamma \) denote the Dynkin diagram of the root system \( \Phi \) with base \( \Delta \). Without referring to the Classification Theorem (Theorem 7.13), show:

(a) \( \Gamma \) contains no cycles: if \( \alpha_1, \ldots, \alpha_k \in \Delta \) are distinct with \( (\alpha_i, \alpha_{i+1}) \neq 0 \) for \( i = 1, \ldots, k - 1 \) and \( (\alpha_k, \alpha_1) \neq 0 \), then \( k \leq 2 \). (Estimate \( (\eta, \eta) \) for \( \eta = \sum_{i=1}^{k} \frac{\alpha_i}{\|\alpha_i\|} \).)

(b) At most three edges can originate from any vertex of \( \Gamma \): for any \( \alpha \in \Delta \), we have \( d(\alpha) := \sum_{\beta \in \Delta \setminus \{\alpha\}} \langle \alpha, \beta \rangle \langle \beta, \alpha \rangle \leq 3 \). (Using (a) show that \( 4\|\alpha\|^2 > \|\alpha\|^2 d(\alpha) \).)

7.3.2 (Determinants of Cartan matrices). Use the Cartan matrices \( C \) in Tables 7.2 and 7.3 to verify the following determinants:

<table>
<thead>
<tr>
<th>root system</th>
<th>( A_n )</th>
<th>( B_n )</th>
<th>( C_n )</th>
<th>( D_n )</th>
<th>( E_6 )</th>
<th>( E_7 )</th>
<th>( E_8 )</th>
<th>( F_4 )</th>
<th>( G_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \det C )</td>
<td>( n + 1 )</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

7.3.3 (Cartan matrices of dual root systems). (a) Let \( \Phi^\vee \) be the dual root system of \( \Phi \) (Exercises 7.1.3 and 7.2.3). Show that, with a suitable labeling of bases, the Cartan matrices \( C = C_\Phi \) and \( C^\vee = C_{\Phi^\vee} \) are related by \( C^\vee = C^T \), the transpose of \( C \).

(b) Use the Cartan matrices \( C \) in Tables 7.2 and 7.3 to show that all irreducible root systems are isomorphic to their duals, except for \( B_n \) and \( C_n \), which are dual to each other.

7.3.4 (Absolute irreducibility). Assume that \( \Phi \) is irreducible and let \( V = \mathbb{Q}\Phi \subseteq \mathbb{B} \) denote the \( \mathbb{Q} \)-subspace generated by \( \Phi \). Show that \( V \) is an absolutely irreducible representation of \( \mathcal{W}_\Phi \) over \( \mathbb{Q} \).

7.4. Lattices Associated to a Root System

Throughout this section, \( \Phi \) denotes a root system in Euclidean space \( \mathbb{B} \cong \mathbb{R}^n \) and \( \Delta = \{\alpha_1, \ldots, \alpha_n\} \) is a fixed base of \( \Phi \).

7.4.1. Root and Weight Lattice

The root lattice of \( \Phi \), by definition, is the sublattice \( \mathbb{Z}\Phi = \sum_{\alpha \in \Phi} \mathbb{Z}\alpha \subseteq \mathbb{B} \); it will be denoted by \( L \) or \( L_\Phi \). By \( B_1 \) and \( B_2 \), we can write \( L \) as follows:

\[
L = \bigoplus_{i=1}^{n} \mathbb{Z}\alpha_i \cong \mathbb{Z}^n
\]
Axiom R4 implies that the root lattice \( L \) is contained in the so-called weight lattice\(^1\) of \( \Phi \), which is defined as follows; see Exercise 7.2.3 for the last equality:

\[
\Lambda = \Lambda_\Phi \overset{\text{def}}{=} \{ \lambda \in \mathbb{E} \mid \langle \lambda, \alpha \rangle \in \mathbb{Z} \text{ for all } \alpha \in \Phi \} = \{ \lambda \in \mathbb{E} \mid \langle \lambda, \alpha_i \rangle \in \mathbb{Z} \text{ for } i = 1, \ldots, n \}.
\]

Note that \( \Lambda \) corresponds to the standard integer lattice \( \mathbb{Z}^n \subseteq \mathbb{R}^n \) under the following \( \mathbb{R} \)-isomorphism, which is a consequence of B1:

\[
\begin{array}{ccc}
\mathbb{B} & \overset{\sim}{\longrightarrow} & \mathbb{R}^n \\
\psi & \mapsto & (\langle \mu, \alpha_i \rangle)_{i=1}^n \\
\mu & & \mu
\end{array}
\]

The preimages \( \lambda_i \in \Lambda \) of the standard basis vectors \( \varepsilon_i = (\delta_{i,j})_{j=1}^n \in \mathbb{Z}^n \) under this isomorphism are called the fundamental weights with respect to the base \( \Delta = \{ \alpha_1, \ldots, \alpha_n \} \); they form an \( \mathbb{R} \)-basis of \( \mathbb{E} \) such that \( \langle \lambda_i, \alpha_j \rangle = \delta_{i,j} \). Thus,

\[
(7.17) \quad \mu = \sum_{i=1}^n \langle \mu, \alpha_i \rangle \lambda_i \quad (\mu \in \mathbb{E})
\]

and

\[
(7.18) \quad \Lambda = \bigoplus_{i=1}^n \mathbb{Z} \lambda_i \cong \mathbb{Z}^n
\]

Using the Cartan matrix \( C = (\langle \alpha_i, \alpha_j \rangle)_{i,j} \) of \( \Phi \), the equations \( \alpha_i = \sum_j \langle \alpha_i, \alpha_j \rangle \lambda_j \) \( (i = 1, \ldots, n) \) can be written as

\[
(7.19) \quad \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = C \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}.
\]

**Lemma 7.14.** \( |\Lambda/L| = \det C \).

**Proof.** This is a consequence of the following standard result (e.g., Cor. 3 in [29, Chap. VII §4 n° 7]): if \( A \) is a lattice and \( f \in \text{End}(A) \) satisfies \( \det f \neq 0 \), then \( |A/fA| = |\det f| \). Now let \( A = \Lambda \) and let \( f \in \text{End}(\Lambda) \) be given by \( f \lambda_i = \alpha_i \) for all \( i \). By (7.19) the matrix of \( f \) for the \( \mathbb{Z} \)-basis \( (\lambda_i)_{i=1}^n \) of \( \Lambda \) is the transpose of the Cartan matrix \( C \). Thus, \( \Lambda/L = \Lambda/f\Lambda \) has order \( \det C \).

\(^1\)The name “weight lattice” comes from the fact that \( \Lambda \) consists precisely of the weights of finite-dimensional representations of the semisimple Lie algebra \( g \) that has \( \Phi \) as its set of roots; see (8.19).
7.4. Lattices Associated to a Root System

Example 7.15 (Root and weight lattice for $A_n$). The Cartan matrix $C$ of the root system of type $A_n$ is given in Table 7.2. One can check that $\det C = n + 1$ (Exercise 7.3.2) and $C^{-1} = \frac{1}{n+1} (d_{i,j})$ with $d_{i,j} = (n + 1) \min \{i, j\} - ij$. The fundamental weights for the base $\alpha_i = e_i - e_{i+1}$ ($i = 1, 2, \ldots, n$) as in Example 7.3 now can be obtained from (7.19):

$$\lambda_i = \mu_1 + \cdots + \mu_i \quad \text{with} \quad \mu_k = e_k - \frac{1}{n+1} \sum_{j=1}^{n+1} e_j.$$

It follows that

$$\Lambda = \bigoplus_{i=1}^{n} \mathbb{Z} \lambda_i = \bigoplus_{k=1}^{n} \mathbb{Z} \mu_k = L + \Lambda_1 \mathbb{Z}.$$

For the last equality, observe that $L + \lambda_1 \mathbb{Z}$ contains the elements $\mu_k = \lambda_1 - \sum_{i<k} \alpha_i$. Since the order of $\lambda_1$ modulo $L$ is $n + 1$, we obtain

$$\Lambda / L \cong \mathbb{Z} / (n + 1) \mathbb{Z}.$$

For $n = 2$, (7.20) gives $\lambda_1 = \frac{1}{3} (2\alpha_1 + \alpha_2)$ and $\lambda_2 = \frac{1}{3} (\alpha_1 + 2\alpha_2)$. Figure 7.2 shows the root lattice $L$ and the weight lattice $\Lambda$ of the root system of type $A_2$, the former being indicated by black dots and the latter as the intersections of the red lines. The gray region and the role of $\rho$ will be discussed in §7.4.2.

![Figure 7.2. Root and weight lattice for $A_2$](image)

7.4.2. Dominant Weights

We will consider the partial order $\preceq$ on $\mathbb{E}$ that is defined by

$$\mu \preceq \nu \iff \nu - \mu \in L_+,$$

where we have put

$$L_+ \overset{\text{def}}{=} \mathbb{Z}_+ \Phi_+ = \mathbb{Z}_+ \Delta = \bigoplus_{i=1}^{n} \mathbb{Z}_+ \alpha_i.$$
Since \( L_+ \) is a submonoid of \( \mathbb{B} \) and \( L_+ \cap -L_+ = \{0\} \), this does indeed define a partial order \( \preceq \). Moreover, \( \preceq \) is compatible with addition: if \( \mu \preceq \mu' \) and \( \nu \preceq \nu' \), then \( \mu + \nu \preceq \mu' + \nu' \).

We will also consider the following submonoid of the weight lattice \( \Lambda \):

\[
\Lambda_+ \overset{\text{def}}{=} \left\{ \lambda \in \mathbb{B} \mid \langle \lambda, \alpha_i \rangle \in \mathbb{Z}_+ \text{ for } i = 1, \ldots, n \right\}
\]

\[
= \left\{ \lambda \in \Lambda \mid \langle \lambda, \alpha_i \rangle \geq 0 \text{ for } i = 1, \ldots, n \right\}
\]

\[
= \left\{ \lambda \in \Lambda \mid \langle \lambda, \alpha \rangle \geq 0 \text{ for all } \alpha \in \Phi_+ \right\}
\]

\[
= \bigoplus_{i=1}^{n} \mathbb{Z}_+ \lambda_i.
\]

The weights \( \lambda \in \Lambda \) belonging to \( \Lambda_+ \) are called **dominant** with respect to \( \Delta \). The second description of \( \Lambda_+ \) above shows that \( \Lambda_+ = \Lambda \cap \mathcal{C}(\Delta) \), where \( \mathcal{C}(\Delta) \) is the closure of the fundamental Weyl chamber for the base \( \Delta \); see (7.14). Observe that \( \Lambda \cap \mathcal{C}(\Delta) = \Lambda_+ \setminus \bigcup_{\alpha \in \Phi} \mathcal{A}_+ = \bigoplus_{i=1}^n \mathbb{N}_+ \lambda_i \). The elements of this subset of \( \Lambda_+ \) are called **strongly dominant**. By Lemma 7.8, we have \( s_{\alpha_i}(\rho) = \rho - \alpha_i \) or, equivalently, \( \langle \rho, \alpha_i \rangle = 1 \) for \( i = 1, \ldots, n \). Thus (7.17) gives the expression

\[
(7.21) \quad \rho = \lambda_1 + \lambda_2 + \cdots + \lambda_n.
\]

Therefore,

\[
(7.22) \quad \Lambda \cap \mathcal{C}(\Delta) = \bigoplus_{i=1}^n \mathbb{N}_+ \lambda_i = \rho + \Lambda_+.
\]

A word of caution about the notation is in order. Namely, despite the fact that \( L \subseteq \Lambda \), it is usually not true that \( L_+ \subseteq \Lambda_+ \). Also, while \( L \cap \Lambda_+ = L \cap \mathcal{C}(\Delta) \) is a subset of \( \Lambda_+ \), it is generally a proper subset. Figure 7.2 depicts the situation for \( \mathfrak{A}_2 \); the gray region, without boundary, is \( \mathcal{C}(\Delta) \) for \( \Delta = \{\alpha_1, \alpha_2\} \).

### 7.4.3. The Action of the Weyl Group

By their very definition as subgroups of \( \text{GL}(\mathbb{B}) \), the groups \( \text{Aut} \Phi \) and \( \mathcal{W} = \mathcal{W}_\Phi \) act naturally on \( \mathbb{B} \). The root lattice \( L = \Lambda_\Phi \) and the weight lattice \( \Lambda = \Lambda_\Phi \) are both stable under these actions. In fact, much more can be said:

**Proposition 7.16.**

(a) \( \Lambda \) and \( L \) are stable under \( \text{Aut} \Phi \) and hence under \( \mathcal{W} \). Moreover, \( \Lambda/L = (\mathbb{E}/L)^\mathcal{W} \), the invariants of the action \( \mathcal{W} \subseteq \mathbb{E}/L \).

(b) If \( \lambda \in \Lambda_+ \), then \( w\lambda \preceq \lambda \) for all \( w \in \mathcal{W} \) and \( \#\{\mu \in \Lambda_+ \mid \mu \preceq \lambda\} < \infty \).

(c) \( \Lambda_+ \) is a fundamental domain for the action of \( \mathcal{W} \) on \( \Lambda \): each orbit \( \mathcal{W}\lambda \) (\( \lambda \in \Lambda \)) intersects \( \Lambda_+ \) in exactly one point. If \( \lambda \) is strongly dominant, then the isotropy group \( \mathcal{W}_\lambda \) is trivial.

---

2Bourbaki, besides denoting \( \Lambda \) by \( \mathbb{P} \) (for *poids*), uses the subscript ++ for the monoid of dominant weights, reserving the subscript + for the monoid \( \Lambda \cap \mathbb{E}_+, L_+ \), which contains the dominant weights (generally properly).
7.4. Lattices Associated to a Root System

Proof. (a) Stability of $L = \mathbb{Z}\Phi$ under $\text{Aut}\Phi$ is clear by definition of $\text{Aut}\Phi$, and stability of $\Lambda = \{\lambda \in \mathbb{E} | \langle \lambda, \alpha \rangle \in \mathbb{Z}\}$ for all $\alpha \in \Phi$ follows from the fact that $\text{Aut}\Phi$ preserves $\Phi$ and the bracket $\langle \ldots \rangle$ (Lemma 7.2). It remains to check that $\Lambda/L = (\mathbb{E}/L)^W$ or, equivalently,

$$\Lambda = \{\lambda \in \mathbb{E} | \lambda - w\lambda \in L \text{ for all } w \in W\}.$$ 

For $\subseteq$, it suffices to show that $\lambda - s_\alpha \lambda \in L$ holds for all $\alpha \in \Phi$ and $\lambda \in \Lambda$, because the reflections $s_\alpha$ generate $W$. But $\lambda - s_\alpha \lambda = \langle \lambda, \alpha \rangle \alpha \in \mathbb{Z}\alpha \subseteq L$ as desired. For the reverse inclusion, let $\lambda \in \mathbb{E}$ be such that $\lambda - w\lambda \in L$ for all $w \in W$. Specializing to $w = s_\alpha$ with $\alpha \in \Delta$, this gives $\langle \lambda, \alpha \rangle \alpha \in L$, and since $L = \bigoplus_{\alpha \in \Delta} \mathbb{Z}\alpha$, we must have $\langle \lambda, \alpha \rangle \in \mathbb{Z}$. Therefore, $\lambda \in \{\mu \in \mathbb{E} | \langle \mu, \alpha \rangle \in \mathbb{Z} \text{ for all } \alpha \in \Delta\} = \Lambda$. This proves $\supseteq$, and hence (a) is proved.

(b) Let $\lambda \in \Lambda_+$ and $w \in W$. We need to show that $w\lambda \leq \lambda$, that is, $\lambda - w\lambda \in L_+$. For this, we argue by induction on the length $\ell = \ell(w)$ (Theorem 7.9). The case $\ell = 0$ being trivial, assume $\ell \geq 1$ and write $w = w's_{\alpha}$ with $w' \in W$ having length $\ell(w') = \ell - 1$ and $\alpha \in \Delta$. Then $ws_{\alpha} = w'$ and

$$\lambda - w\lambda = \lambda - w'\lambda + w'(s_\alpha \lambda - \lambda) = \lambda - w'\lambda - \langle \lambda, \alpha \rangle w\alpha.$$ 

Here, $\lambda - w'\lambda \in L_+$ by induction. Furthermore, $-\langle \lambda, \alpha \rangle w(\alpha) \in -\mathbb{Z}_+\Phi_+ = L_+$, because $\lambda \in \Lambda_+$ gives $\langle \lambda, \alpha \rangle \in \mathbb{Z}_+$ and Lemma 7.8 gives $w\alpha \in \Phi_-$. Therefore, $\lambda - w\lambda \in L_+$ as claimed.

As for the finiteness assertion, let $\mu \in \Lambda_+$ with $\mu \leq \lambda$. Thus, $\lambda - \mu \in L_+$ and $\lambda + \mu \in \Lambda_+$, because $\Lambda_+$ is a monoid. By definition of $L_+$ and $\Lambda_+$, it follows that $\langle \lambda + \mu, \lambda - \mu \rangle \geq 0$. Since $\langle \lambda + \mu, \lambda - \mu \rangle = ||\lambda||^2 - ||\mu||^2$, this means $||\mu|| \leq ||\lambda||$. Consequently, $\{\mu \in \Lambda^+ | \mu \leq \lambda\}$ is contained in ball of radius $||\lambda||$ about the origin of $\mathbb{E}$, which is compact, as well as in the discrete set $\Lambda$, and hence it is a finite set.

(c) Since $\mathbb{E}$ is the union of the closures $\overset{\circ}{\mathcal{C}}_\Delta$ of the various Weyl chambers $\mathcal{C}_\Delta$, any given $\lambda \in \Lambda$ belongs to some $\overset{\circ}{\mathcal{C}}_\Delta$. By Theorem 7.9 and the correspondence (7.13), there exists a unique $w \in W$ such that $w^\vee \overset{\circ}{\mathcal{C}}_\Delta = \mathcal{C}_\Delta$, the fundamental Weyl chamber for the base $\Delta$. Therefore, $w\lambda \in \Lambda \cap \overline{\mathcal{C}}(\Delta) = \Lambda_+$; so each $W$-orbit in $\Lambda$ intersects $\Lambda_+$. Finally assume that $\lambda \in \Lambda_+$ and $w\lambda \in \Lambda_+$ for some $w \in W$. Then, applying (b) to $\lambda$ and to $w\lambda$, we obtain $w\lambda \leq \lambda$ and $\lambda = w^{-1}w\lambda \leq w\lambda$. Therefore, $w\lambda = \lambda$. Finally, if $\lambda \in \Lambda \cap \overline{\mathcal{C}}(\Delta)$ and $w\lambda = \lambda$, then $\mathcal{C}(\Delta) \cap w^\vee \mathcal{C}(\Delta) \neq \emptyset$, which forces $w = 1$ by (7.13) and Theorem 7.9. This finishes the proof of the proposition. $\square$

7.4.4. Multiplicative Invariants of Weight Lattices

Continuing with the notation of the previous section, we now focus on the weight lattice $\Lambda = \Lambda_\Phi$ and on the action of the Weyl group $W = W_\Phi$ on the group algebra $\mathbb{k}\Lambda$ of $\Lambda$ over an arbitrary commutative base ring $\mathbb{k}$.

Recall that $\mathbb{k}\Lambda$ contains a copy of $\Lambda$ as a subgroup of the group of units $(\mathbb{k}\Lambda)^\times$ and this copy of $\Lambda$ forms a $\mathbb{k}$-basis of $\mathbb{k}\Lambda$. As in Example 3.1, we will pass from the
customary additive notation of $\Lambda$ to a multiplicative notation when working inside $k\Lambda$, writing the basis element of $k\Lambda$ corresponding to $\lambda \in \Lambda$ as $x^\lambda$. Any choice of a $\mathbb{Z}$-basis $(\mu_i)_{i=1}^n$ of $\Lambda$—the most common choice are the fundamental weights (7.17)—gives rise to an isomorphism of $k\Lambda$ with the Laurent polynomial algebra over $k$:

$$k\Lambda \xrightarrow{\sim} k[x_1^{\pm 1}, \ldots, x_n^{\pm 1}], \quad x^{\mu_i} \mapsto x_i.$$ 

The action $\mathcal{W} \subset \Lambda$ extends uniquely to an action by $k$-algebra automorphisms on $k\Lambda$:

$$(7.23) \quad w \cdot \sum_{\lambda \in \Lambda} k_{\lambda} x^\lambda = \sum_{\lambda \in \Lambda} k_{\lambda} x^{w.\lambda} \quad (w \in \mathcal{W}).$$

This action is called a multiplicative action and the subalgebra of $k\Lambda$ consisting of all elements that are fixed under this action is called the multiplicative invariant algebra of the weight lattice $\Lambda$ over $k$:

$$(k\Lambda)^{\mathcal{W}} \overset{\text{def}}{=} \{ f \in k\Lambda \mid w.f = f \text{ for all } w \in \mathcal{W} \}.$$ 

The structure of $(k\Lambda)^{\mathcal{W}}$ as a $k$-module is easy to pin down, because $k\Lambda$ is a permutation representation of $\mathcal{W}$ over $k$: the action of $\mathcal{W}$ stabilizes the $k$-basis $(x^\lambda)_{\lambda \in \Lambda}$ of $k\Lambda$. Therefore, by (3.24), a $k$-basis of the $\mathcal{W}$-invariants $(k\Lambda)^{\mathcal{W}}$ is provided by the distinct $\mathcal{W}$-orbit sums in $k\Lambda$, that is, elements of the form

$$\sigma_\lambda := \sum_{\mu \in \mathcal{W}(\lambda)} x^\mu \quad (\lambda \in \Lambda).$$

Since each $\mathcal{W}$-orbit in $\Lambda$ meets the monoid $\Lambda_+$ of dominant weights in exactly one point (Proposition 7.16), the distinct orbit sums $\sigma_\lambda$ are obtained by letting $\lambda$ range over $\Lambda_+$. Thus,

$$(7.24) \quad (k\Lambda)^{\mathcal{W}} = \bigoplus_{\lambda \in \Lambda_+} k \sigma_\lambda.$$ 

It remains to sort out the ring theoretic structure of $(k\Lambda)^{\mathcal{W}}$. If $\Phi$ is of type $A_1$, then $\Lambda \cong \mathbb{Z}$ and $\mathcal{W} \cong S_2$ acts by $\lambda \mapsto -\lambda$. The invariant algebra was dealt with in Proposition 5.46 (with $k = \mathbb{Z}$): $(k\Lambda)^{\mathcal{W}} \cong k[t^{\pm 1}]^{S_2} = k[\sigma_t]$ with $\sigma_t = t + t^{-1}$. The following theorem of Bourbaki [23, chap. VI §3] covers all root systems $\Phi$.

**Theorem 7.17** (notation as above). The invariant algebra $(k\Lambda)^{\mathcal{W}}$ over any commutative base ring $k$ is a polynomial algebra in $n = \text{rank } \Phi$ variables: $(k\Lambda)^{\mathcal{W}} \cong k\Lambda_+$. More precisely, given any collection of invariants of the form

$$f_i = x^{\lambda_i} + \sum_{\mu \in \mathcal{W}(\lambda_i)} k_{i,\mu} x^\mu \in (k\Lambda)^{\mathcal{W}} \quad (i = 1, \ldots, n),$$

**Footnote**: Multiplicative invariant algebras arising from group actions on other lattices generally do not have such a simple description; see [137].
where the $\lambda_i$ are the fundamental weights with respect to a base $\Delta$ of $\Phi$, the map $t_i \mapsto f_i$ yields a $k$-algebra isomorphism $k[t_1, \ldots, t_n] \rightarrow (k\Lambda)^W$. In particular, we may take $f_i = \sigma_{\lambda_i}$.

**Proof.** The last statement of the theorem, that the $W$-orbit sums $\sigma_{\lambda_i}$ do indeed have the required form for the elements $f_i$, follows from the fact that the $W$-orbit of each $\lambda_i$ consists of $\lambda_i$ together with elements of $\Lambda$ that are strictly smaller than $\lambda_i$ for the partial order $\preceq$ (Proposition 7.16).

Now let $f_1, \ldots, f_n \in (k\Lambda)^W$ be as stipulated in the theorem and let $\lambda \in \Lambda_+$ be given. Write $\lambda = \sum_i m_i \lambda_i$ with unique $m_i \in \mathbb{Z}_+$ and put

$$f_\lambda := \prod_i t_i^{m_i} \in (k\Lambda)^W.$$ 

The theorem asserts that $F := (f_\lambda)_{\lambda \in \Lambda_+}$ is a $k$-basis of $(k\Lambda)^W$. To prove this, note that each $f_\lambda$ has the form

$$f_\lambda = x^\lambda + \sum_{\mu \preceq \lambda} k_{\lambda, \mu} x^\mu$$

with $k_{\lambda, \mu} \in k$; this is a consequence of the fact that the partial order $\preceq$ on $\Lambda$ is compatible with addition.

Assume that $k_1 f_{\lambda_1} + k_2 f_{\lambda_2} + \cdots + k_r f_{\lambda_r} = 0$ with $k_j \in k$ and distinct $\lambda_i \in \Lambda_+$; say $\lambda_1$ is maximal among the $\lambda_i$ for $\preceq$. Expressing $k_1 f_{\lambda_1} + k_2 f_{\lambda_2} + \cdots + k_r f_{\lambda_r}$ in terms of the standard basis $(x^\nu)_{\nu \in \Lambda}$ of $k\Lambda$, it follows from (7.25) that the coefficient of $x^{\lambda_1}$ is $k_1$. So we must have $k_1 = 0$, which shows that $F$ is $k$-independent.

Since $f_\lambda$ is $W$-invariant, (7.25) also implies that

$$f_\lambda = \sigma_\lambda + \sum_{\nu \preceq \lambda} k_{\lambda, \nu} \sigma_\nu.$$ 

It follows that each orbit sum $\sigma_\lambda$ ($\lambda \in \Lambda_+$) is a $k$-linear combination of elements in $F$. For, otherwise the finiteness statement in Proposition 7.16(c) would allow us to pick a counterexample $\lambda \in \Lambda_+$ that is minimal with respect to $\preceq$. Thus, all $\sigma_\nu$ in (7.26) can be expressed in terms of $F$, and hence $\sigma_\lambda$ as well, which is a contradiction. Thus, the $k$-linear span of $F$ is all of $(k\Lambda)^W$ by (7.24). This completes the proof of the theorem.

**7.4.5. Anti-invariants**

The material in this subsection will play a role in connection with Weyl’s Character Formula (Section 8.7), with $k = \mathbb{Z}$, but nowhere else in this book. Here, we continue to work over a commutative ring $k$; we only need to assume now that $k$ has no 2-torsion: $k \neq -k$ for $0 \neq k \in k$. 

As in the familiar special case of the symmetric groups (for type $\mathfrak{A}$), the Weyl group $\mathcal{W} = \mathcal{W}_\Phi$ of any root system $\Phi$ has a “sign” representation:

$$\text{sgn}: \mathcal{W} \longrightarrow \{\pm 1\}$$

(7.27)

Indeed, $\mathcal{W}$ is generated by reflections $s_\alpha \in \text{GL}(E)$, which all have determinant $-1$. This representation allows us to define anti-invariants for $\mathcal{W}$ in the group ring $\mathbb{k}\Lambda$, for the multiplicative action $\mathcal{W} \subset \mathbb{k}\Lambda$ considered in §7.4.4:

$$\begin{align*}
(\mathbb{k}\Lambda)^\pm & \overset{\text{def}}{=} \{ f \in \mathbb{k}\Lambda \mid w.f = \text{sgn}(w)f \text{ for all } w \in \mathcal{W} \} \\
\end{align*}$$

Since the $\mathcal{W}$-action respects the multiplication of $\mathbb{k}\Lambda$, the inclusions $(\mathbb{k}\Lambda)^W(\mathbb{k}\Lambda)^\pm \subseteq (\mathbb{k}\Lambda)^\pm$ and $(\mathbb{k}\Lambda)^W(\mathbb{k}\Lambda)^\pm \subseteq (\mathbb{k}\Lambda)^\pm$ hold in $\mathbb{k}\Lambda$.

Anti-invariants can easily be produced by means of the “antisymmetrizer,”

$$a := \sum_{w \in \mathcal{W}} \text{sgn}(w)w \in \mathbb{k}\mathcal{W}.$$ 

For the symmetric groups, the antisymmetrizer was introduced in §3.8.1, normalized so as to be an idempotent. In general, for any $w \in \mathcal{W}$, it is straightforward to check the equalities $wa = aw = \text{sgn}(w)a$ in the group algebra $\mathbb{k}\mathcal{W}$. Therefore, considering the algebra map $\mathbb{k}\mathcal{W} \rightarrow \text{End}_\mathbb{k}(\mathbb{k}\Lambda)$ that comes from the multiplicative action $\mathcal{W} \subset \mathbb{k}\Lambda$, we have $a.\mathbb{k}\Lambda \subseteq (\mathbb{k}\Lambda)^\pm$. In particular, we obtain the anti-invariants

$$a_{\lambda} := a.x^\lambda = \sum_{w \in \mathcal{W}} \text{sgn}(w)x^{w.\lambda} \in (\mathbb{k}\Lambda)^\pm \quad (\lambda \in \Lambda).$$

**Proposition 7.18.** A $\mathbb{k}$-basis of $(\mathbb{k}\Lambda)^\pm$ is given by the elements $a_{\lambda}$, where $\lambda$ runs over the set $\Lambda \cap \bar{\mathcal{C}}(\Delta) = \rho + \Lambda_+$ of strongly dominant weights (7.22).

**Proof.** Each $a_{\lambda}$ involves only terms $x^\mu$ with $\mu$ belonging to the orbit $\mathcal{W}(\lambda)$. Since the orbits $\mathcal{W}(\lambda)$ with $\lambda$ strongly dominant are all distinct and $|\mathcal{W}(\lambda)| = |\mathcal{W}|$ (Proposition 7.16), the elements $a_{\lambda}$ are $\mathbb{k}$-independent.

Now let $f = \sum_{\mu \in \Lambda} k_{\mu} x^\mu \in (\mathbb{k}\Lambda)^\pm$ be arbitrary. If $(\mu, \alpha) = 0$ for some $\alpha \in \Phi$, then $s_\alpha \mu = \mu$. Therefore, the coefficient of $x^\mu$ in $s_\alpha.f = \sum_{\mu \in \Lambda} k_{\mu} x^{s_\alpha \mu}$ is also $k_{\mu}$ while, on the other hand, $s_\alpha.f = -f$. It follows that $k_{\mu} = -k_{\mu}$ and so $k_{\mu} = 0$. Thus, all $\mu$ occurring in $f$ belong to $\Lambda \cap \bar{\mathcal{B}}$, where $\bar{\mathcal{B}} = \mathbb{B} \setminus \bigcup_{\alpha \in \Phi} \alpha^\perp$. By Proposition 7.16, $\mu = w.\lambda$ for unique $w \in \mathcal{W}$ and $\lambda \in \Lambda \cap \bar{\mathcal{C}}(\Delta) = \rho + \Lambda_+$, the collection of strongly dominant weights. Consequently,

$$f = \sum_{\lambda \in \rho + \Lambda_+} \sum_{w \in \mathcal{W}} k_{w.\lambda} x^{w.\lambda}.$$
Finally, since $w.f = \sum_{\mu} k_{w\mu} x^{w\mu} = \text{sgn}(w) \sum_{\mu} k_{\mu} x^{\mu}$, it follows that $k_{w\mu} = \text{sgn}(w)k_{\mu}$ for all $\mu$. The above expression for $f$ therefore becomes $f = \sum_{\lambda \in \rho + \Lambda_+} k_{\lambda}a_{\lambda}$, which wraps up the proof.

The anti-invariant $a_{\lambda}$ for the strongly dominant weight $\lambda = \rho$ will play a special role later on. In order to evaluate $a_{\rho}$, consider the element $x^\rho \in \mathbb{k}\Lambda$ and observe that, in the group algebra $\mathbb{k}[\mathcal{B}] \supset \mathbb{k}\Lambda$, we may write $x^\rho = \prod_{\alpha \in \Phi_+} x^{\frac{\alpha}{2}}$ by (7.12). Now put

$$d := \prod_{\alpha \in \Phi_+} (x^{\frac{\alpha}{2}} - x^{\frac{\alpha}{2}}) = x^\rho \prod_{\alpha \in \Phi_+} (1 - x^{-\alpha}) = x^{-\rho} \prod_{\alpha \in \Phi_+} (x^{\alpha} - 1) \in \mathbb{k}\Lambda.$$

**Proposition 7.19.**

(a) $d = a_{\rho}$.

(b) $(\mathbb{k}\Lambda)^{\pm} = d(\mathbb{k}\Lambda)^{W}$.

**Proof.** (a) Since the reflection $s_{\alpha}$ permutes the set $\Phi_+ \setminus \{\alpha\}$ for $\alpha \in \Delta$ (Lemma 7.8) and $s_{\alpha}\alpha = -\alpha$, we obtain

$$s_{\alpha}(d) = (x^{-\frac{\alpha}{2}} - x^{\frac{\alpha}{2}}) \prod_{\beta \in \Phi_+ \setminus \{\alpha\}} (x^\beta - x^{-\beta}) = \text{sgn}(s_{\alpha})d.$$

It follows that $d \in (\mathbb{k}\Lambda)^{\pm}$, because the reflections $s_{\alpha}$ generate $W$ (Theorem 7.9).

Expanding the expression $d = x^\rho \prod_{\alpha \in \Phi_+} (1 - x^{-\alpha})$ in the group algebra $\mathbb{k}\Lambda$ results in a finite sum of the following form:

$$d = x^\rho + \sum_{\mu \neq \rho} k_{\mu} x^{\mu}.$$

Indeed, $x^\rho$ has this form (with all $k_{\mu} = 0$) and multiplication of any sum of the above form by $1 - x^{-\alpha}$ with $\alpha \in \Delta$ yields such a sum. On the other hand, $d = \sum_{\lambda \in \rho + \Lambda_+} c_{\lambda}a_{\lambda}$ by Proposition 7.18, with each $a_{\lambda}$ having the form $a_{\lambda} = x^{\lambda} + \sum_{\mu \neq \lambda} d_{\lambda\mu}x^{\mu}$ (Proposition 7.16). Comparison of the two expressions for $d$ shows that we must have $c_{\lambda} = 0$ if $\lambda \neq \rho$ and $c_{\rho} = 1$, proving (a).

(b) The inclusion $(\mathbb{k}\Lambda)^{\pm} \supset d(\mathbb{k}\Lambda)^{W}$ is clear, because $d \in (\mathbb{k}\Lambda)^{\pm}$. For the other inclusion, we may assume that $\mathbb{k} = \mathbb{Z}$; the general case then follows by extension of scalars, since $(\mathbb{k}\Lambda)^{\pm} = \mathbb{Z} \otimes_{\mathbb{Z}} (\mathbb{Z}\Lambda)^{\pm}$ by Proposition 7.18. Since $\mathbb{Z}\Lambda$ is a unique factorization domain (a Laurent polynomial algebra over $\mathbb{Z}$), it suffices to show:

(i) Each $a_{\lambda}$ ($\lambda \in \Lambda$) is divisible in $\mathbb{Z}\Lambda$ by all $x^{\alpha} - 1$ with $\alpha \in \Phi$, and

(ii) if $\alpha, \beta \in \Phi$ and $\beta \neq \pm\alpha$, then $x^{\alpha} - 1, x^{\beta} - 1 \in \mathbb{Z}\Lambda$ are relatively prime.

It will follow from (i) and (ii) that each $a_{\lambda}$ is divisible by $d = x^{-\rho} \prod_{\alpha \in \Phi_+} (x^{\alpha} - 1)$ in $\mathbb{Z}\Lambda$, because $x^{-\rho}$ is a unit in $\mathbb{Z}\Lambda$. The quotient $a_{\lambda}/d$ is evidently $W$-invariant; so $a_{\lambda} \in d(\mathbb{Z}\Lambda)^{W}$. The desired inclusion $(\mathbb{Z}\Lambda)^{\pm} \subseteq d(\mathbb{Z}\Lambda)^{W}$ then follows from Proposition 7.18.
The proof of (ii), in essence a purely ring-theoretic fact, is left as an exercise to the reader (Exercise 7.4.6). For (i), consider the reflection \( s_\alpha \in W \) and let \( T \subseteq W \) be a transversal for the cosets \( \langle s_\alpha \rangle w \). Then

\[
a_\lambda = \sum_{w \in T} \text{sgn}(w)x^{w(\lambda)} + \sum_{w \in T} \text{sgn}(s_\alpha w)x^{s_\alpha w(\lambda)}
= \sum_{w \in T} \text{sgn}(w)x^{w(\lambda)} - \sum_{w \in T} \text{sgn}(w)x^{w - (w, \lambda)\alpha}
= \sum_{w \in T} \text{sgn}(w)x^{w(\lambda)}(1 - x^{-(w, \lambda)\alpha}).
\]

But \( n_w := -(w, \alpha) \in \mathbb{Z} \), because \( w(\lambda) \in \Lambda \). If \( n_w \geq 0 \), then the computation

\[
1 - x^{-(w, \lambda)\alpha} = 1 - x^{n_w\alpha} = (1 - x^\alpha)(1 + x^\alpha + \cdots + x^{(n_w-1)\alpha})
\]

shows that \( 1 - x^{-(w, \lambda)\alpha} \) is divisible in \( \mathbb{Z}\Lambda \) by \( x^\alpha - 1 \). If \( n_w < 0 \), then the same conclusion holds, because \( 1 - x^{n_w\alpha} = -x^{n_w\alpha}(1 - x^{-n_w\alpha}) \). Thus, all terms in the last sum for \( a_\lambda \) are divisible by \( x^\alpha - 1 \) and hence so is \( a_\lambda \). This proves (ii) and finishes the proof of the proposition.

\[\square\]

**Exercises for Section 7.4**

7.4.1 (Root and weight lattice for \( B_n \)). Using the base \( \Delta \) in Example 7.4 and the Cartan matrix of \( B_n \) in Table 7.2, show:

(a) The fundamental weights for the base \( \Delta \) are given by \( \lambda_i = e_1 + \cdots + e_i (i = 1, \ldots, n - 1) \) and \( \lambda_n = 1/2(e_1 + \cdots + e_n) \).

(b) The root lattice is \( L = \bigoplus_{i=1}^n \mathbb{Z}e_i \) and the weight lattice is \( \Lambda = L + \mathbb{Z}\lambda_n \). Thus, \( \Lambda/L \cong \mathbb{Z}/2\mathbb{Z} \).

7.4.2 (Root and weight lattice for \( C_n \)). Using the base \( \Delta \) in Example 7.5 and the Cartan matrix of \( C_n \) in Table 7.2, show:

(a) The fundamental weights for the base \( \Delta \) are given by \( \lambda_i = e_1 + \cdots + e_i (i = 1, \ldots, n) \).

(b) The weight lattice is \( \Lambda = \bigoplus_{i=1}^n \mathbb{Z}e_i \) and the root lattice is the sublattice \( L = \{ \sum_i x_i e_i \in \Lambda \mid \sum_i x_i \equiv 0 \text{ mod } 2 \} \). Thus, \( \Lambda/L \cong \mathbb{Z}/2\mathbb{Z} \).

7.4.3 (Root and weight lattice for \( D_n \)). Using the base \( \Delta \) in Example 7.6 and the Cartan matrix of \( D_n \) in Table 7.2, show:

(a) The fundamental weights for the base \( \Delta \) are given by \( \lambda_i = e_1 + \cdots + e_i (i = 1, \ldots, n - 2), \lambda_{n-1} = 1/2(e_1 + \cdots + e_{n-1} - e_n), \) and \( \lambda_n = 1/2(e_1 + \cdots + e_{n-1} + e_n) \).

(b) The root lattice is \( L = \{ \sum_i x_i e_i \in \bigoplus_{i=1}^n \mathbb{Z}e_i \mid \sum_i x_i \equiv 0 \text{ mod } 2 \} \) and the weight lattice is \( \Lambda = \bigoplus_{i=1}^n \mathbb{Z}e_i + \mathbb{Z}\lambda_n \).
7.4. Lattices Associated to a Root System

(c) \( \Lambda/L \equiv \begin{cases} \mathbb{Z}/4\mathbb{Z} & \text{for } n \text{ odd}, \\ \mathbb{Z}/2\mathbb{Z} \oplus \mathbb{Z}/2\mathbb{Z} & \text{for } n \text{ even}. \end{cases} \)

7.4.4 (Vanishing anti-invariants). Show that the following are equivalent for \( \mu \in \Lambda \):

(i) \( \mu \) lies on a wall: \( \mu \in \alpha^\perp \) for some \( \alpha \in \Phi \); (ii) the isotropy group \( \mathcal{W}_\mu \) is nontrivial; and (iii) \( a_\mu = 0 \).

7.4.5. Show that \( \prod_{\alpha \in \Phi_+} (w.\alpha) = \text{sgn}(w) \prod_{\alpha \in \Phi_+} (\alpha,\alpha) \) for \( \lambda \in \mathbb{E} \) and \( w \in \mathcal{W} \).

7.4.6 (A detail for Proposition 7.19). Let \( \Gamma \equiv \mathbb{Z}^r \) be a lattice and let \( \mu, \nu \in \Gamma \) be such that the sublattice \( \mathbb{Z}\mu + \mathbb{Z}\nu \subseteq \Gamma \) has rank 2. Show that \( \mathfrak{x}^\mu - 1, \mathfrak{x}^\nu - 1 \in \mathbb{Z}\Gamma \) are relatively prime. (It suffices to show that these two elements are relatively prime in \( \mathbb{C}\Gamma \). Factor \( \mathfrak{x}^\mu - 1 \) into irreducibles in \( \mathbb{C}\Gamma \) and show that none of them divides \( \mathfrak{x}^\nu - 1 \).)

The next two problems calculate some multiplicative invariant algebras of weight lattices using, instead of Theorem 7.17, the following more basic Fundamental Theorem of \( S_n \)-invariants, which determines the algebra of all symmetric polynomials in the variables \( x_1, \ldots, x_n \) (e.g., [29, chap. IV §6 Thm. 1]): Let \( S_n \) act on the polynomial algebra \( \mathbb{k}[x_1, \ldots, x_n] \) over a commutative ring \( \mathbb{k} \) by permuting the variables: \( s_i x_i = x_{s(i)} \). Then the invariant algebra \( \mathbb{k}[x_1, \ldots, x_n]^{S_n} \) is generated by the elementary symmetric polynomials \( e_i = e_i(x_1, \ldots, x_n) \) with \( 1 \leq i \leq n \).

7.4.7 (Multiplicative invariants of the weight lattice for \( A_n \)). The weight lattice for the root system of type \( A_n \) is given by \( \Lambda = \bigoplus_{i=1}^n \mathbb{Z}\mu_i \subseteq \mathbb{R}^{n+1} \), where \( \mu_i = e_i - \frac{1}{n+1} \sum_{j=1}^{n+1} e_j \) and \( (e_i)_{i=1}^{n+1} \) is the standard basis of \( \mathbb{R}^{n+1} \) (Example 7.15). The Weyl group \( \mathcal{W} = S_{n+1} \) permutes the \( e_i \) and hence also the \( \mu_i \); see (7.4). Use the Fundamental Theorem to show that \( (\mathbb{k}\Lambda)^{\mathcal{W}} = \mathbb{k}[s_1, \ldots, s_n] \) with \( s_i = e_i(x^{e_1}, \ldots, x^{e_n}) \).

7.4.8 (Multiplicative invariants of the weight lattice for \( C_n \)). Let \( \Lambda = \bigoplus_{i=1}^n \mathbb{Z}e_i \) be the weight lattice for type \( C_n \) (Exercise 7.4.2) and let \( \mathcal{W} = \{ \pm 1 \} \times \mathbb{S}_n \) be the Weyl group; see (7.8). Using the fact that \( \mathbb{k}[i^{\pm 1}]^{S_2} = \mathbb{k}[i + i^{-1}] \) (Proposition 5.46) and the Fundamental Theorem of \( S_n \)-invariants, show that \( (\mathbb{k}\Lambda)^{\mathcal{W}} = \mathbb{k}[s_1, \ldots, s_n] \) with \( s_i = e_i(x^{e_1} + x^{-e_1}, \ldots, x^{e_n} + x^{-e_n}) \).
Chapter 8

Representations of Semisimple Lie Algebras

After our excursion into Euclidean space in Chapter 7, we now resume the main thread of the representation theory of Lie algebras. Focusing on the case of a semisimple Lie algebra $\mathfrak{g}$, our main goal in this chapter is to describe the finite-dimensional representations of $\mathfrak{g}$ along the lines of the earlier description for $\mathfrak{sl}_2$ (Section 5.7). Our attention will be primarily on the set $\text{Irr}^\text{fin} \mathfrak{g}$ of all finite-dimensional irreducible representations of $\mathfrak{g}$. Indeed, by Weyl's Theorem (Section 6.2), an arbitrary $V \in \text{Rep}^\text{fin} \mathfrak{g}$ decomposes into finitely many irreducible constituents and the isomorphism type of $V$ is determined by the multiplicities of these constituents.

Certain infinite-dimensional representations will also play an important role: the so-called highest weight representations. It turns out that both finite-dimensional representations and highest weight representations are completely reducible for the Cartan subalgebra $\mathfrak{h}$. Moreover, the intersection of the two classes of representations is exactly $\text{Irr}^\text{fin} \mathfrak{g}$ (Corollary 8.4).

This chapter ends our coverage of Lie algebras—some will say, rightfully, just where things start to get interesting. To the reader wishing to see more, we recommend the monographs by Dixmier [60], Humphreys [104] and Jantzen [112].

Throughout this chapter, $\mathfrak{g}$ will denote a semisimple Lie algebra and the base field $\mathbb{K}$ is assumed to be algebraically closed and to have characteristic 0. Furthermore, $\mathfrak{h} \subseteq \mathfrak{g}$ will be fixed Cartan subalgebra, $\Phi \subseteq \mathfrak{h}^*$ the corresponding set of roots, $\Delta = \{\alpha_1, \ldots, \alpha_n\}$ a fixed base of $\Phi$, and $W = W_\Phi$ the Weyl group.
8. Representations of Semisimple Lie Algebras

8.1. Reminders

8.1.1. The Setting

Recall that the $Q$-space $Q\Phi \cong Q^n$ is contained in $h^* = k \otimes_Q Q\Phi \cong k^n$ as well as in $E = R \otimes_Q Q\Phi \cong R^n$; see (6.22). In this chapter, we shall view $Q\Phi$ in its native environment, namely $h^*$, rather than in $E$. This also applies to the various structures associated to $\Phi$ that were introduced in Chapter 7, such as the base $\Delta = \{\alpha_1, \ldots, \alpha_n\}$ of $\Phi$ and the root lattice,

$$L = \sum_{\beta \in \Phi} \mathbb{Z}\beta = \bigoplus_{i=1}^n \mathbb{Z}\alpha_i \subseteq Q\Phi.$$ 

By (6.24), the Cartan integers $\langle \beta, \alpha \rangle \in \mathbb{Z}$ for $\alpha, \beta \in \Phi$ have the form $\langle \beta, \alpha \rangle = \langle \beta, h_\alpha \rangle$, where $\langle \cdot, \cdot \rangle : h^* \times h \to k$ also denotes the evaluation pairing and $h_\alpha$ is the unique element of $[g_\alpha, g_{-\alpha}] \subseteq h$ satisfying $\langle \alpha, h_\alpha \rangle = 2$ (Theorem 6.10). More generally, we have

$$\langle \lambda, \alpha \rangle = \langle \lambda, h_\alpha \rangle \in k \quad (\lambda \in h^*, \alpha \in \Phi).$$

The various descriptions of the weight lattice $\Lambda \subseteq Q\Phi$ in §7.4.1 now take the following form:

$$\Lambda = \{ \lambda \in h^* \mid \langle \lambda, h_\alpha \rangle \in \mathbb{Z} \text{ for all } \alpha \in \Phi \}$$

$$= \{ \lambda \in h^* \mid \langle \lambda, h_{\alpha_i} \rangle \in \mathbb{Z} \text{ for } i = 1, \ldots, n \}$$

$$= \bigoplus_{i=1}^n \mathbb{Z}\lambda_i \quad \text{with} \quad \langle \lambda_j, h_{\alpha_i} \rangle = \delta_{i,j}.$$ 

(8.1)

The weights $\lambda_1, \ldots, \lambda_n$ are the fundamental weights for the given base $\Delta = \{\alpha_1, \ldots, \alpha_n\}$. The submonoid of dominant weights (§7.4.2) will continue to be relevant:

$$\Lambda_+ = \{ \lambda \in \Lambda \mid \langle \lambda, h_\alpha \rangle \in \mathbb{Z}_+ \text{ for all } \alpha \in \Phi_+ \}$$

$$= \{ \lambda \in h^* \mid \langle \lambda, h_{\alpha_i} \rangle \in \mathbb{Z}_+ \text{ for } i = 1, \ldots, n \}$$

$$= \bigoplus_{i=1}^n \mathbb{Z}_+\lambda_i.$$ 

(8.2)

8.1.2. Weyl Group

The groups $\text{Aut} \Phi$ and $W$, both introduced in Chapter 7 as subgroups of $\text{GL}(E)$, may in fact be regarded as subgroups of $\text{GL}(Q\Phi) \cong \text{GL}_n(Q)$ by restriction, and hence by scalar extension as subgroups of $\text{GL}(h^*) \cong \text{GL}_n(k)$. The reflection $s_\alpha \in W$ for a root $\alpha \in \Phi$ is given by

$$s_\alpha \mu = \mu - \langle \mu, h_\alpha \rangle \alpha \quad (\mu \in h^*).$$ 

(8.3)
Since \( \mathcal{W} \) preserves the bracket \( \langle \cdot , \cdot \rangle \) on \( \mathbb{E} \times (\mathbb{E} \setminus \{0\}) \) and the bracket is linear in the first variable, we have

\[
\langle w \lambda, h_\alpha \rangle = \langle \lambda, h_{w^{-1} \alpha} \rangle \quad (w \in \mathcal{W}, \lambda \in \mathfrak{h}^*, \alpha \in \Phi).
\]

**Proposition 8.1.** The Lie algebra \( \mathfrak{g} \) is simple if and only if \( \mathfrak{h}^* \) is an irreducible representation of \( \mathcal{W} \).

**Proof.** By Propositions 6.11 and 7.11, we know that \( \mathfrak{g} \) is simple if and only if \( \mathcal{E} \) is an irreducible representation of \( \mathcal{W} \). We need to show that this is equivalent to \( \mathfrak{h}^* \) being irreducible. In fact, both statements are equivalent to irreducibility of the representation \( \mathcal{Q} \Phi \). Indeed, since \( \mathcal{E} = \mathbb{R} \otimes_\mathcal{Q} \mathcal{Q} \Phi \), irreducibility of \( \mathcal{E} \) certainly forces \( \mathcal{Q} \Phi \) to be irreducible and likewise for \( \mathfrak{h}^* \) in place of \( \mathcal{E} \). Conversely, assume that \( V = \mathcal{Q} \Phi \) is irreducible and let \( d \in D(V) = \text{End}_\mathcal{Q} \mathcal{W}(V) \). Then, for each reflection \( s_\alpha \in \mathcal{W} \), we have \( ds_\alpha = s_\alpha d \) and so \( d \) stabilizes the subspace \( (\text{Id}_V - s_\alpha)(V) = \mathcal{Q} \alpha \). Therefore, all \( \alpha \in \Phi \) are eigenvectors for \( d \) with rational eigenvalues. All these eigenvalues are the same, because nonzero eigenspaces of \( d \) are \( \mathcal{W} \)-subrepresentations of \( V \) and hence they must be equal to \( V \). This shows that \( d \in \mathcal{Q} \text{Id}_V \). Thus, \( V \) is in fact absolutely irreducible (Proposition 1.36) and so both \( \mathcal{E} \) and \( \mathfrak{h}^* \) are irreducible. \( \square \)

The Weyl group \( \mathcal{W} \) also acts on \( \mathfrak{h} \) by duality (§3.3.3):

\[
\langle \lambda, w h \rangle = \langle w^{-1} \lambda, h \rangle \quad (\lambda \in \mathfrak{h}^*, w \in \mathcal{W}, h \in \mathfrak{h}).
\]

In particular, for the reflections \( s_\alpha \in \mathcal{W} \), we have \( ds_\alpha = s_\alpha d \) and so \( d \) stabilizes the subspace \( (\text{Id}_V - s_\alpha)(V) = \mathcal{Q} \alpha \). Therefore, all \( \alpha \in \Phi \) are eigenvectors for \( d \) with rational eigenvalues. All these eigenvalues are the same, because nonzero eigenspaces of \( d \) are \( \mathcal{W} \)-subrepresentations of \( V \) and hence they must be equal to \( V \). This shows that \( d \in \mathcal{Q} \text{Id}_V \). Thus, \( V \) is in fact absolutely irreducible (Proposition 1.36) and so both \( \mathcal{E} \) and \( \mathfrak{h}^* \) are irreducible.

**8.1.3. Triangular Decomposition**

The partial order \( \preceq \) on \( \mathcal{E} \) that was defined in §7.4.2 can be transported to \( \mathfrak{h}^* \) in such a way that the partial orders of \( \mathfrak{h}^* \) and \( \mathcal{E} \) agree on \( \mathcal{Q} \Phi \):

\[
\mu \preceq \lambda \iff \lambda - \mu = \mu_\lambda = \bigoplus_{i=1}^n \mathbb{Z}_+ \alpha_i \quad (\lambda, \mu \in \mathfrak{h}^*).
\]
This partial order manifestly depends on our choice of base $\Delta = \{\alpha_1, \ldots, \alpha_n\}$. In terms of $\preceq$, the sets $\Phi_{\pm}$ of positive/negative roots with respect to $\Delta$ may be written as

$$\Phi_+ = \{\alpha \in \Phi \mid \alpha > 0\} \quad \text{and} \quad \Phi_- = \{\alpha \in \Phi \mid \alpha < 0\}.$$  

Since $\Phi = \Phi_+ \sqcup \Phi_-$, the root space decomposition (6.12) takes the form

$$(8.8) \quad g = n_- \oplus h \oplus n_+ \quad \text{with} \quad n_\pm \overset{\text{def}}{=} \bigoplus_{\alpha \in \Phi_\pm} 0_\alpha.$$  

All three summands are Lie subalgebras of $g$, with $h$ being abelian and both $n_\pm$ nilpotent (Exercise 8.1.1). The decomposition (8.8) is called the triangular decomposition of $g$; for the classical Lie algebras, with $h$ chosen to consist of the diagonal matrices as in §§6.4.2, 6.4.3 and with $\Delta$ as exhibited in Examples 7.3–7.6, $n_+$ and $n_-$ are respectively the Lie subalgebras of strictly upper and lower triangular matrices in the Lie algebra in question.

The following Lie subalgebras of $g$, called the positive and negative Borel subalgebra (for the chosen base $\Delta$), will also play an important role in the following:

$$(8.9) \quad b_\pm \overset{\text{def}}{=} h \oplus n_\pm$$  

8.1.4. The Case of $\mathfrak{sl}_2$

We briefly review the special case $g = \mathfrak{sl}_2$. In subsequent sections, we will frequently refer to $\mathfrak{sl}_2$ in order to illustrate new material by relating it to our earlier findings in Section 5.7.

Recall that $\mathfrak{sl}_2 = \mathbb{k}f \oplus \mathbb{k}h \oplus \mathbb{k}e$ with $[h, f] = -2f$, $[h, e] = 2e$ and $[e, f] = h$. The Cartan subalgebra that we will work with is $h = \mathbb{k}h$. We will identify each $\lambda \in h^*$ with the value $\langle \lambda, h \rangle \in \mathbb{k}$ and hence use the identification $h^* = \mathbb{k}$. With this, the set of roots is $\Phi = \{\pm 2\}$ and the root lattice is $L = 2\mathbb{Z}$. The Weyl group $W \cong S_2$ operates on $h^* = \mathbb{k}$ by multiplication with $\pm 1$. The $W$-invariants in $h^*/L = \mathbb{k}/2\mathbb{Z}$ are $\mathbb{Z}/2\mathbb{Z}$ and so the weight lattice is $\Lambda = \mathbb{Z}$. Thus:

$$L \subset \Lambda \subset h^*$$

$$2\mathbb{Z} \subset \mathbb{Z} \subset \mathbb{k}$$

Choosing $\Delta = \{2\}$ for our base, as we will throughout, we have $\Phi_+ = \{2\}$ and $\Phi_- = \{-2\}$. The positive and negative parts of the triangular decomposition of $\mathfrak{sl}_2$ are $n_+ = \mathbb{k}e$ and $n_- = \mathbb{k}f$. Furthermore, $L_+ = 2\mathbb{Z}_+$, $\Lambda_+ = \mathbb{Z}_+$ and

$$\alpha_1 = 2, \quad h_{\alpha_1} = h \quad \text{and} \quad \lambda_1 = 1.$$
The partial order \( \preceq \) on \( \mathfrak{h}^* = \mathbb{K} \) is given by \( \mu \preceq \lambda \) if and only if \( \lambda - \mu \in 2\mathbb{Z}_+ \).

**Exercises for Section 8.1**

8.1.1 (The subalgebras \( \mathfrak{n}_\pm \) and \( \mathfrak{b}_\pm \)). Using the notation of §8.1.3, show:

(a) Both \( \mathfrak{n}_\pm \) and both \( \mathfrak{b}_\pm \) are Lie algebras of \( \mathfrak{g} \).

(b) Both \( \mathfrak{b}_\pm \) are solvable, and \( \mathfrak{n}_\pm \) is the nilpotent radical of \( \mathfrak{b}_\pm \).

(c) \( \mathfrak{n}_\pm \) is generated by the root spaces \( \mathfrak{g}_\alpha \) with \( \alpha \in \pm \Delta \). (Use Exercise 7.2.2.)

8.1.2 (Zariski dense subsets). Show that \( \Lambda_+ \) is dense in \( \mathfrak{h}^* \) for the Zariski topology (Section C.3). More generally, for any collection of infinite subsets \( I_1, \ldots, I_n \subseteq \mathbb{K} \), the subset \( \{ \sum k_i \lambda_i \mid k_i \in I_i \} \) is Zariski dense in \( \mathfrak{h}^* \).

8.2. Finite-Dimensional Representations

8.2.1. Weight Spaces

Since the Cartan subalgebra \( \mathfrak{h} \) is abelian, the enveloping algebra \( \mathcal{U}\mathfrak{h} \) coincides with the symmetric algebra \( \text{Sym} \mathfrak{h} \). All irreducible representations of \( \mathfrak{h} \) are 1-dimensional, having the form \( \mathbb{K}_\lambda = \mathbb{K} \) with \( h.1 = \langle \lambda, h \rangle \) for \( h \in \mathfrak{h} \), and we have bijections

\[
\begin{array}{ccc}
\text{Irr} \mathfrak{h} & \sim \rightarrow & \text{Hom}_{\text{Alg}}(\mathcal{U}\mathfrak{h}, \mathbb{K}) \\
\sim \rightarrow & \mathfrak{h}^* & \sim \rightarrow \\
\mathbb{K}_\lambda & \sim \rightarrow & \Lambda
\end{array}
\]

For an arbitrary \( M \in \text{Rep} \mathfrak{h} \), we will denote the homogeneous component \( M(\mathbb{K}_\lambda) \) (§1.4.2) by \( M_\lambda \). Thus,

\[
M_\lambda = \{ m \in M \mid h.m = \langle \lambda, h \rangle m \text{ for all } h \in \mathfrak{h} \}
\]

The socle of \( M \) is the sum of the various homogeneous components \( M_\lambda \), which is in fact a direct sum (Proposition 1.31): \( \text{soc} M = \bigoplus_{\lambda \in \mathfrak{h}^*} M_\lambda \).

We will be primarily interested in the case where \( M \) arises by restriction from some \( V \in \text{Rep} \mathfrak{g} \); so \( M = V \downarrow_{\mathcal{U}\mathfrak{h}} \). In this case, \( \text{soc}(V \downarrow_{\mathcal{U}\mathfrak{h}}) = \bigoplus_{\lambda \in \mathfrak{h}^*} V_\lambda \) is a subrepresentation of \( \mathfrak{g} \). This is a consequence of the following inclusion, which was stated earlier as (6.10) for \( V = \mathfrak{g}_{\text{ad}} \); the verification in general is similarly straightforward:

\[
\text{soc}(V_\alpha \downarrow \mathfrak{h}) \subseteq V_{\alpha + \lambda} \quad (\alpha, \lambda \in \mathfrak{h}^*).
\]

The nonzero components \( V_\lambda \) are called the **weight spaces** of \( V \) and the corresponding \( \lambda \in \mathfrak{h}^* \) the **weights** of \( V \). Nonzero elements \( v \in V \) belonging to some \( V_\lambda \) are called **weight vectors** and the dimension \( \dim_{\mathbb{K}} V_\lambda \), which may be infinite, is referred to as the **multiplicity** of the weight \( \lambda \) in \( V \).
8.2.2. Weights of Finite-Dimensional Representations

A general \( V \in \text{Rep}_g \) need not have any weights, even if \( V \) is irreducible (Exercise 8.2.1). However, this is not the case for finite-dimensional representations. The following proposition collects some basic facts about weights of finite-dimensional representations of \( g \). Part (b) hints at why \( \Lambda \) is called the “weight lattice.” We will prove shortly that the converse also holds: every element of \( \Lambda \) is in fact a weight of some \( V \in \text{Rep}_{\text{fin}} g \); see (8.19).

**Proposition 8.2.** Let \( V \in \text{Rep}_{\text{fin}} g \). Then:

(a) \( V = \bigoplus_{\lambda \in \Lambda} V_{\lambda} \).

(b) All weights of \( V \) belong to \( \Lambda \).

(c) If \( \lambda \) is a weight of \( V \), then the orbit \( W_{\lambda} \) consists of weights of \( V \), all having the same multiplicity: \( \dim_k V_{\lambda} = \dim_k V_{w\lambda} \) for all \( w \in W \).

**Proof.** (a) By preservation of Jordan decomposition (Proposition 6.8), all elements of \( h \in \mathfrak{b} \) act as diagonalizable operators \( h_{\lambda} \) on \( V \), and since \( \mathfrak{b} \) is abelian, all \( h_{\lambda} \) are simultaneously diagonalizable. This proves (a).

(b) Let \( \lambda \) be a weight of \( V \). We need to show that \( \langle \lambda, h_{\alpha} \rangle \in \mathbb{Z} \) for all \( \alpha \in \Phi \); see (8.1). For this, consider the \( \mathfrak{sl}_2 \)-triple \( s_{\alpha} = k f_{\alpha} \oplus k h_{\alpha} \oplus k e_{\alpha} \subseteq g \) (Theorem 6.10). By restriction to \( s_{\alpha} \), the representation \( V \) can be viewed as a finite-dimensional \( \mathfrak{sl}_2 \)-representation and \( \langle \lambda, h_{\alpha} \rangle \) is a weight of this representation. Hence, \( \langle \lambda, h_{\alpha} \rangle \in \mathbb{Z} \) by Theorem 5.39.

(c) We may assume that \( w = s_{\alpha} \) with \( \alpha \in \Phi \); so \( s_{\alpha} \lambda = \lambda - l_{\alpha} \), where we have put \( l := \langle \lambda, h_{\alpha} \rangle \in \mathbb{Z} \). As in the proof of (b), regard \( V_{s_{\alpha}} \) as a finite-dimensional \( \mathfrak{sl}_2 \)-representation. If \( l \geq 0 \), then consider the endomorphism \( (f_{\alpha})_{\lambda} \in \text{End}_k(V) \). By (8.10), \( (f_{\alpha})_{\lambda} V_{\lambda} \subseteq V_{\lambda - l_{\alpha}} \), because \( f_{\alpha} \in \mathfrak{g}_{-\alpha} \). Moreover, if \( 0 \neq \nu \in V_{\lambda} \), then \( \nu \) must have a nonzero component in some irreducible constituent \( V(m) \) of \( V_{s_{\alpha}} \), with \( m \geq l \). It follows from Proposition 5.37 that \( f_{\alpha} \nu \neq 0 \). Thus, \( (f_{\alpha})_{\lambda} \) embeds \( V_{\lambda} \) into \( V_{\lambda - l_{\alpha}} \). If \( l \leq 0 \), then one can argue similarly with \( (e_{\alpha})_{\lambda} \). In either case, \( \dim_k V_{\lambda} \leq \dim_k V_{\lambda - l_{\alpha}} = \dim_k V_{s_{\alpha} \lambda} \) and, by symmetry, equality must hold. \( \square \)

Part (a) of the proposition implies the following formulae, for any \( V, W \in \text{Rep}_{\text{fin}} g \) and \( \lambda \in \Lambda^* \),

\[
(V \oplus W)_\lambda = V_{\lambda} \oplus W_{\lambda} \quad \text{and} \quad (V \otimes W)_\lambda = \bigoplus_{\mu + \nu = \lambda} (V_\mu \otimes W_\nu).
\]

These relations have been observed earlier for \( \mathfrak{sl}_2 \) (proof of Lemma 5.44) and the reasoning given there applies in general. Finally, in the usual way, Proposition 8.2 extends to representations of \( g \) that are merely locally finite, because such representations are sums of finite-dimensional representations.
8.3. Highest Weight Representations

8.3.1. Maximal Vectors and Highest Weight Vectors

Let \( V \in \text{Rep} \mathfrak{g} \) be arbitrary, not necessarily finite dimensional. A vector \( 0 \neq v \in V \) is called **maximal** if \( v \) is a common eigenvector for the operators \( x_{\lambda} \) with \( x \in \mathfrak{b}_+ = \mathfrak{b} \oplus \mathfrak{n}_+ \). Being an eigenvector for \( \mathfrak{h} \), any maximal \( v \) belongs to some weight space \( V_{\lambda} \) (\( \lambda \in \mathfrak{h}^* \)) and (8.10) further implies \( n_+ \cdot v = 0 \), because \( g_{\alpha} \cdot v = 0 \) for some \( \alpha \in \Phi \). One calls \( V \) a **highest weight representation** (with highest weight \( \lambda \)) if \( V \) is generated by some maximal vector \( 0 \neq v \in V \) (with weight \( \lambda \)); so \( V = U \mathfrak{g} \cdot v \).

The reason for the “highest weight” nomenclature will become apparent shortly. For now, we remark that any nonzero homomorphic image of a highest weight representation is also a highest weight representation with the same highest weight: the image of a generating maximal vector will again be a generating maximal vector. In general, maximal vectors need not exist, but Lie’s Theorem guarantees their existence if \( V \) is finite dimensional. Alternatively, since \( V \) has only finitely many weights (Proposition 8.2), we may choose \( \lambda \) to be maximal among all weights of \( V \) for the partial order \( \preceq \). Then any \( 0 \neq v \in V_{\lambda} \) is a maximal vector, because (8.7) and (8.10) imply \( n_+ \cdot v = 0 \). In particular, any \( V \in \text{Irr}_{\text{fin}} \mathfrak{g} \) is a highest weight representation, being generated by any maximal vector \( 0 \neq v \in V \). The converse fails: highest weight representations are generally neither finite dimensional nor irreducible. However, we will see that finite-dimensional highest weight representations are in fact irreducible (Corollary 8.4). Thus, finite-dimensional highest weight representations make up the set \( \text{Irr}_{\text{fin}} \mathfrak{g} \), whose description is our main objective in this chapter.

**First Properties of Highest Weight Representations.** The proposition below lists some basic features of highest weight representations. In particular, part (b) shows that the generating maximal vector \( v \) of a highest weight representation \( V = U \mathfrak{g} \cdot v \) is unique up to scalar multiples and that its weight is the highest for \( \preceq \). Therefore,
Witt Theorem also tells us that the standard monomials \( v \) have the form

\[
\varphi_+(\mu) \overset{\text{def}}{=} \# \{ (i_\alpha)_{\alpha \in \Phi_+} \in \mathbb{Z}_{+}^{\Phi_+} \mid \mu = \sum_{\alpha \in \Phi_+} i_\alpha \alpha \} \quad (\mu \in \mathfrak{h}^*).
\]

Thus, \( \varphi_+(\mu) > 0 \) if and only if \( \mu \in L_+ \). Furthermore, since each \( \alpha \in \Phi_+ \) is a nonzero \( \mathbb{Z}_+ \)-linear combination of the simple roots \( \alpha_i \), which are linearly independent, it is easy to see that \( \varphi_+(\mu) < \infty \) for all \( \mu \in \mathfrak{h}^* \) and \( \varphi_+(0) = 1 \).

**Proposition 8.3.** Let \( V \in \text{Rep}_\mathfrak{g} \) be a highest weight representation with highest weight \( \lambda \in \mathfrak{h}^* \). Then:

1. \( V = \bigoplus_{\mu \in \mathfrak{h}^*} V_\mu \);
2. All weights \( \mu \) of \( V \) satisfy \( \mu \leq \lambda \) and \( \dim_{\mathbb{k}} V_\mu < \infty \), with \( \dim_{\mathbb{k}} V_\lambda = 1 \);
3. \( V \) has a unique largest proper subrepresentation. In particular, \( V \) is indecomposable.

**Proof.** Write \( V = U_\mathfrak{q}.v \) with \( 0 \neq v \in V_\lambda \) such that \( \mathfrak{n}_+.v = 0 \) and observe that the enveloping algebra \( U(\mathfrak{b}_+) \) acts by scalars on \( v \). Since \( U_\mathfrak{q} = U(\mathfrak{n}_-) \mathfrak{b}_+ \) by the Poincaré-Birkhoff-Witt Theorem, it follows that \( V = U(\mathfrak{n}_-).v \). Fixing an ordering of \( \Phi_+ \), say \( \Phi_+ = \{ \beta_1, \ldots, \beta_t \} \), and writing \( a_{-\beta_i} = \mathbb{k} f_i \), the Poincaré-Birkhoff-Witt Theorem also tells us that the standard monomials \( f^i = f_1^{i_1} f_2^{i_2} \ldots f_t^{i_t} \) with \( i = (i_1, \ldots, i_t) \in \mathbb{Z}_+^t \) form a \( \mathbb{k} \)-basis of \( U(\mathfrak{n}_-) \). Therefore, \( V = \sum_{\mathbf{i} \in \mathbb{Z}_+^t} \mathbb{k} f^\mathbf{i}.v \). Since \( f^\mathbf{i}.v \in V_{\lambda - \sum_{i_\alpha} \alpha_i} \) by (8.10), part (a) follows. We also see that all weights \( \mu \) of \( V \) have the form \( \mu = \lambda - \sum_{i_\alpha} \alpha_i \) with \( \mathbf{i} = (i_1, \ldots, i_t) \in \mathbb{Z}_+^t \), and the weight space \( V_\mu \) is spanned by the corresponding vectors \( f^\mathbf{i}.v \). Thus, \( \mu \leq \lambda \) and

\[
\dim_{\mathbb{k}} V_\mu \leq \varphi_+(\lambda - \mu).
\]

Since \( \varphi_+(\lambda - \mu) < \infty \) for all \( \mu \) and \( \varphi_+(0) = 1 \), part (b) is proved.

Finally, if \( W \) is any \( \mathfrak{g} \)-subrepresentation of \( V \), then \( W = \bigoplus_{\mu \in \mathfrak{h}^*} (W \cap V_\mu) \) by Proposition 1.31. If \( W \cap V_\lambda \neq 0 \), then the generating maximal vector of \( V \) must belong to \( W \), because \( \dim_{\mathbb{k}} V_\lambda = 1 \). Therefore, \( W = V \) in this case. Consequently, the sum of all subrepresentations \( W \subseteq V \) is a proper subrepresentation of \( V \), being contained in \( \bigoplus_{\mu \prec \lambda} V_\mu \subseteq V \). Indecomposability of \( V \) is now clear.

**Corollary 8.4.** Finite-dimensional highest weight representations coincide with finite-dimensional irreducible representations of \( \mathfrak{g} \).

**Proof.** We have already pointed out that \( \text{Irr}_{\text{fin}} \mathfrak{g} \) consists of highest weight representations. Conversely, if \( V \) is any finite-dimensional highest weight representation, then \( V \) is completely reducible by Weyl’s Theorem and indecomposable by Proposition 8.3, and hence \( V \) must be irreducible.
For any highest weight representation \( V \), we define the head\(^1\) of \( V \) to be the factor of \( V \) by its largest proper subrepresentation and we will use the notation
\[
\text{head} \, V.
\]

Thus, head \( V \) is an irreducible highest weight representation, not necessarily finite dimensional, having the same highest weight as \( V \).

**Construction of Maximal Vectors.** Recall that an element \( 0 \neq v \in V \) of an arbitrary \( V \in \text{Rep} \, g \) is maximal if \( v \in V_{\lambda} \) for some \( \lambda \in \mathfrak{h}^* \) and \( n_+ \cdot v = 0 \). The latter condition does of course depend on the choice of \( \Delta \); in fact, it is equivalent to \( g_{\alpha} \cdot v = 0 \) for all \( \alpha \in \Delta \), because \( n_+ \) is generated by the root spaces of simple roots (Exercise 8.1.1). The following technical lemma will be useful below.

**Lemma 8.5.** Let \( v \in V \) be maximal with weight \( \lambda \) and let \( \alpha \in \Delta \). Assume that \( \langle \lambda, h_{\alpha} \rangle + 1 \in \mathbb{Z}_+ \) and put \( v_{\alpha} = f_{\alpha}^{1+}(\lambda, h_{\alpha}) \cdot v \in V \), where \( 0 \neq f_{\alpha} \in g_{-\alpha} \). Then either \( v_{\alpha} = 0 \) or \( v_{\alpha} \) is maximal with weight \( s_{\alpha} \lambda - \alpha \).

**Proof.** Put \( l = \langle \lambda, h_{\alpha} \rangle \). Then \( v_{\alpha} = f_{\alpha}^{l+1} \cdot v \in V_{\lambda-(l+1)\alpha} \) by (8.10) and \( \lambda - (l + 1)\alpha = s_{\alpha} \lambda - \alpha \). We need to check that \( g_{\beta} \cdot v = 0 \) for all \( \beta \in \Delta \). For \( \beta = \alpha \), consider the \( \mathfrak{sl}_2 \)-triple \( s_{\alpha} = \mathbb{k} f_{\alpha} \oplus \mathbb{k} h_{\alpha} \oplus \mathbb{k} e_{\alpha} \subseteq g \) (Theorem 6.10) and use the identity \([e_{\alpha}, f_{\alpha}^{l+1}] = (l + 1) f_{\alpha}^{l}(h_{\alpha} - l) \) (Exercise 5.7.1). This gives
\[
e_{\alpha} \cdot v_{\alpha} = f_{\alpha}^{l+1} e_{\alpha} \cdot v + [e_{\alpha}, f_{\alpha}^{l+1}] \cdot v = 0 + (l + 1) f_{\alpha}^{l}(h_{\alpha} - l) \cdot v = 0.
\]
For \( \beta \neq \alpha \), we have \([e_{\beta}, f_{\alpha}] \in g_{\beta - \alpha} = 0 \), because \( \beta - \alpha \) is not a root, being neither positive nor negative. Therefore, \( e_{\beta} \cdot v_{\alpha} = f_{\alpha}^{l+1} e_{\beta} \cdot v = 0 \) again. This proves the lemma. \( \square \)

**8.3.2. Verma Modules**

We shall now construct, for any given \( \lambda \in \mathfrak{h}^* \), a certain highest weight representation with highest weight \( \lambda \) that is universal among all such representations.

Inflate \( k_\lambda \in \text{Irr} \, \mathfrak{b} \) (§8.2.1) to a representation of \( \mathfrak{b}_+ \) via the projection \( \mathfrak{b}_+ \rightarrow \mathfrak{h} \) along \( n_+ \). Thus, \( h.1 = \langle \lambda, h \rangle \) for \( h \in \mathfrak{b} \) and \( x.1 = 0 \) for \( x \in n_+ \). The **Verma module** associated to \( \lambda \) is the induced representation,
\[
M(\lambda) \overset{\text{def}}{=} \left. k_\lambda \right|_{U_0 = U_0}^\mathbb{U}_0 = U_0 \otimes_{U(\mathfrak{b}_+)} k_\lambda
\]

This is a highest weight representation with highest weight \( \lambda \): \( 1 \otimes 1 \in M(\lambda) \) is a generator that is maximal with weight \( \lambda \) by construction. We also define
\[
V(\lambda) \overset{\text{def}}{=} \text{head} \, M(\lambda)
\]

---

\(^1\)See also Exercise 1.4.1 and §2.1.4.
Thus, \( V(0) = 1 \) and always \( V(\lambda) \in \text{Irr}_g \), an irreducible highest weight representation with highest weight \( \lambda \). We will see shortly that, up to isomorphism, \( V(\lambda) \) is the unique such representation.

**Example 8.6** (Verma modules for \( sl_2 \)). Continuing with the notation of §8.1.4, let \( \lambda \in \mathfrak{h}^* = \mathbb{k} \). Then \( \mathbb{k}_\lambda = \mathbb{k} \) with \( h.1 = \lambda \) and \( e.1 = 0 \). Therefore,

\[
M(\lambda) = \bigoplus_{i \in \mathbb{Z}} \mathbb{k} f_i \quad \text{with} \quad f_i := f^i \otimes 1.
\]

With the aid of the commutation relations in Exercise 5.7.1, one checks that

\[
f . f_i = f_{i+1}, \quad h . f_i = (\lambda - 2i) f_i \quad \text{and} \quad e . f_i = i(\lambda - i + 1) f_{i-1}.
\]

The representation \( M(\lambda) \) was already studied in Exercise 5.7.3: if \( \lambda \not\in \mathbb{Z}_+ \), then \( M(\lambda) \) is irreducible, and so \( V(\lambda) = M(\lambda) \). However, for \( \lambda \in \mathbb{Z}_+ \), the Verma module \( M(\lambda) \) maps onto the familiar \((m + 1)\)-dimensional irreducible representation \( V(m) \) with \( m = \lambda \) (§5.7.2): \( f_i \mapsto \lambda(\lambda - 1) \ldots (\lambda - i + 1) h_j \) gives an epimorphism \( M(m) \to V(m) \); see (5.51). Thus, for \( \lambda = m \in \mathbb{Z}_+ \), the notation \( V(\lambda) \) is consistent with our earlier notation \( V(m) \).

The Verma module \( M(\lambda) \) is universal among all highest weight representation with highest weight \( \lambda \): these representations are exactly the nonzero homomorphic images of \( M(\lambda) \). More precisely:

**Proposition 8.7.**

(a) If \( V \) is any highest weight representation with highest weight \( \lambda \in \mathfrak{h}^* \), then \( \text{Hom}_{\mathfrak{b}}(M(\lambda), V) \cong \mathbb{k} \) and all nonzero homomorphisms are onto.

(b) \( M(\lambda)|_{\mathfrak{n}_{\text{reg}}} \cong \mathbb{U}(n_{\text{reg}}) \).

(c) \( \dim_\mathbb{k} M(\lambda)_\mu = \varphi_+(\lambda - \mu) \) for all \( \mu \in \mathfrak{h}^* \); see (8.12).

**Proof.** (a) By Frobenius reciprocity (Proposition 1.9),

\[
\text{Hom}_{\mathfrak{b}}(M(\lambda), V) = \text{Hom}_{\mathfrak{b}}(\mathbb{k}_\lambda, V_{\mathfrak{b}_+}(\mathbb{k}_\lambda, V)|_{\mathfrak{b}_+}).
\]

The image of any homomorphism \( \mathbb{k}_\lambda \to V_{\mathfrak{b}_+} \) in \( \text{Rep} \mathfrak{b}_+ \) must be contained in \( V_\mathfrak{t} \) and \( V_\lambda \cong \mathbb{k}_\lambda \) (Proposition 8.3). Therefore,

\[
\text{Hom}_{\mathfrak{b}}(\mathbb{k}_\lambda, V)|_{\mathfrak{b}_+}) = \text{Hom}_{\mathfrak{b}}(\mathbb{k}_\lambda, V_\mathfrak{t}) \cong \text{Hom}_{\mathfrak{b}}(\mathbb{k}_\lambda, \mathbb{k}_\lambda) \cong \mathbb{k}.
\]

This results in the isomorphism

\[
\text{Hom}_{\mathfrak{b}}(M(\lambda), V) \xrightarrow{\sim} V_\lambda \cong \mathbb{k}
\]

Since \( 1 \otimes 1 \) generates \( M(\lambda) \), we must have \( f(1 \otimes 1) \neq 0 \) if \( f \neq 0 \). Thus, \( V_\lambda \subseteq \text{Im} f \) in this case and so \( V = \mathbb{U} \mathfrak{b} V_\lambda \subseteq \text{Im} f \).
(b) By Corollary 5.25, we know that $U \cong U(n-) \otimes_k U(b_+)$ as $(U(n-), U(b_+))$-bimodules, with the left and right regular module structures of $U(n-)$ and $U(b_+)$, respectively. Therefore,

\[
M(\lambda) = U \otimes_{U(b_+)} k, \\
\cong (U(n-) \otimes_k U(b_+)) \otimes_{U(b_+)} k, \\
\cong U(n-) \otimes_k (U(b_+) \otimes_{U(b_+)} k), \\
\cong U(n-),
\]

with the regular module structure for $U(n-)$. 

(c) Using the notation employed in the proof of Proposition 8.3, part (b) tells us that the standard monomials $f_i \otimes 1$ form a $k$-basis of $M(\lambda)$ consisting of weight vectors. Thus, the inequality (8.13) is in fact an equality for $V = M(\lambda)$. □

**Corollary 8.8.** Let $V$ be a highest weight representation with highest weight $\lambda$. Then:

(a) $\text{End}_{U}\lambda(V) \cong k$.

(b) If $V$ is irreducible, then $V \cong \lambda(V)$.

**Proof.** (a) follows from part (a) of Proposition 8.7: there is an epimorphism $f: M(\lambda) \twoheadrightarrow V$ and $\circ f$ embeds $\text{End}_U(V)$ into $\text{Hom}_{U\lambda}(M(\lambda), V) \cong k$. If $V$ is irreducible, then $f$ factors through the canonical map $M(\lambda) \twoheadrightarrow \text{head} M(\lambda) = \lambda(V)$, giving the isomorphism $V(\lambda) \cong V$ in (b). □

### 8.3.3. Central Characters

Central characters of irreducible representations have already made an appearance in §5.6.4 in connection with the Nullstellensatz. As a consequence of Corollary 8.8, any highest weight representation $V$ also has a central character: the center $Z = Z(U)$ acts on $V$ by endomorphisms belonging to $\text{End}_U(V) = k$; so $\mathcal{Z}$ acts by scalars. Furthermore, this action depends only on the highest weight of $V$, say $\lambda$, because $V$ is an image of $M(\lambda)$ and so $\mathcal{Z}$ acts on $V$ as it does on $M(\lambda)$. Therefore, we have an algebra map,

\[
\chi_V = \chi_\lambda: \mathcal{Z}(U) \longrightarrow \text{End}_U(V) = k
\]

(8.15)

This map is called the **central character** of $V$ or associated to $\lambda \in \mathfrak{h}^*$.

**Example 8.9** (Central characters for $\mathfrak{sl}_2$). Recall from Example 8.6 that the Verma module $M(\lambda)$ for $\lambda \in \mathfrak{h}^* = k$ has the form $M(\lambda) = \bigoplus_{i \in \mathbb{Z}} k f_i$ with $f_i f_j = f_{i+j}$, $h f_i = (\lambda - 2i) f_i$ and $e f_i = i(\lambda - i + 1) f_{i-1}$. The center of $U(\mathfrak{sl}_2)$ is the polynomial
algebra $Z = \mathbb{Z}[c]$ with $c = 4fe + h(h + 2)$ (Proposition 5.47). Since $c.f_0 = \lambda(\lambda + 2)$, the central character $\chi_\lambda$ is given by $\chi_\lambda(c) = \lambda(\lambda + 2)$.

It will occasionally be convenient to consider the following shifted action of the Weyl group $W$ on $\mathfrak{h}^\ast$, where $\rho$ is the strongly dominant weight in (7.12), (7.21):

$$w \cdot \lambda \overset{\text{def}}{=} w(\lambda + \rho) - \rho \quad (w \in W, \lambda \in \mathfrak{h}^\ast).$$

**Proposition 8.10.** For any $\lambda \in \Lambda_+$ and $w \in W$, the Verma module $M(w \cdot (\lambda - \rho))$ embeds into $M(\lambda - \rho)$. In particular, $\chi_{w \cdot (\lambda - \rho)} = \chi_{\lambda - \rho}$.

**Proof.** The embedding is of course clear for $w = 1$. So assume that $w \neq 1$ and proceed by induction on the length $\ell(w)$ (Theorem 7.9). Write $w = s_\alpha w'$ with $\alpha \in \Delta$ and $\ell(w') = \ell(w) - 1$. By induction, $M(w' \cdot (\lambda - \rho))$ embeds into $M(\lambda - \rho)$. Thus, putting $\mu = w \cdot (\lambda - \rho)$ and $\mu' = w' \cdot (\lambda - \rho)$, it suffices to show that $M(\mu)$ embeds into $M(\mu')$. Since $s_\alpha \rho = \rho - \alpha$ by Lemma 7.8, we have

$$\mu = s_\alpha \cdot \mu' = s_\alpha \mu' - \alpha.$$

By Lemma 7.8, we also know that $w^{-1} \alpha = (w')^{-1} s_\alpha \alpha \in \Phi_-$, and hence $(w')^{-1} \alpha = -w^{-1} \alpha \in \Phi_+$. Therefore, our hypothesis $\lambda \in \Lambda_+$ implies that the following scalar is in fact a non-negative integer; see (8.2):

$$m := \langle \lambda, h_{(w')^{-1} \alpha} \rangle \overset{(8.4)}{=} \langle w' \lambda, h_{\alpha} \rangle = \langle \mu' + \rho, h_{\alpha} \rangle = \langle \mu', h_{\alpha} \rangle + 1.$$

Consider the vector $v := f_{\alpha}^m.(1 \otimes 1) \in M(\mu')$, where $1 \otimes 1$ is the standard highest weight vector of $M(\mu')$, a maximal vector of weight $\mu'$. Since $M(\mu')$ is torsion-free as $U(\mathfrak{n}_-)$-module, being isomorphic to $U(\mathfrak{n}_-)_{\text{reg}}$ (Proposition 8.7), we certainly have $v \neq 0$. Therefore, Lemma 8.5 tells us that $v$ is maximal with weight $s_\alpha \mu' - \alpha = \mu$. Consequently, the subrepresentation $V := U_{\lambda} v \subseteq M(\mu')$ is a highest weight representation with highest weight $\mu$, and hence $V$ is a homomorphic image of $M(\mu)$ (Proposition 8.7). Since $M(\mu) \downarrow U(\mathfrak{n}_-) \cong U(\mathfrak{n}_-)_{\text{reg}}$ and $V$ is torsion-free as $U(\mathfrak{n}_-)$-module, it follows that $V \cong M(\mu)$. This completes the proof of the embedding, and the character formula is an immediate consequence. \hfill \Box

### 8.3.4. Finite Length

We finish this section with another fundamental property of general highest weight representations: they all have finite length, that is, they have a composition series (§1.2.4). The proof given below relies on the following fact, stronger than what was shown in Proposition 8.10, which will only be proven later (Corollary 8.24):

Let $\lambda, \mu \in \mathfrak{h}^\ast$. Then $\chi_\lambda = \chi_\mu$ if and only if $\mathcal{W} \cdot \mu = \mathcal{W} \cdot \lambda$.

As the reader will see, no use of Proposition 8.11 below is being made in Sections 8.4–8.6; it will however be an essential ingredient in the proof of Weyl’s Character Formula (Section 8.7).
Proposition 8.11. Let \( V \) be a highest weight representation with highest weight \( \lambda \). Then \( V \) has finite length. The multiplicity of \( V(\lambda) \) as a composition factor of \( V \) is one; all other composition factors have the form \( V(\mu) \) with \( \mu \not\leq \lambda \) and \( \mu \in \mathcal{W} \cdot \lambda \).

**Proof.** Let \( M \) be an irreducible subquotient of \( V \). We will show that \( M \cong V(\mu) \) for some \( \mu \in \mathfrak{h}^* \) with \( \mu \leq \lambda \) and \( \mu \in \mathcal{W} \cdot \lambda \). To prove this, note that every weight \( \mu \) of \( M \) is also a weight of \( V \), because \( V|_{U_{\mathfrak{h}}} \) is completely reducible. Hence \( \mu \leq \lambda \) (Proposition 8.3) and so \( \lambda - \mu = \sum_{\alpha \in \Delta} z_{\alpha} \alpha \) with unique \( z_{\alpha} \in \mathbb{Z}_+ \). Choosing a weight \( \mu \) of \( M \) such that \( \sum_{\alpha \in \Delta} z_{\alpha} \) is minimal, we see that no \( \mu + \alpha \) with \( \alpha \in \Phi_+ \) is a weight of \( M \). Therefore, \( n_{\mu}.M_{\mu} = 0 \) by (8.10) and any \( 0 \neq m \in M_{\mu} \) is a maximal vector. Since \( M = U_{\mathfrak{g}}.m \) by irreducibility, \( M \) is an irreducible highest weight representation with highest weight \( \mu \), and so \( M \cong V(\mu) \) (Corollary 8.8).

Recall that the center \( \mathcal{Z}(U_{\mathfrak{g}}) \) acts on \( V \), and hence also on the subquotient \( M \), via the central character \( \chi_\lambda \) and on \( V(\mu) \) via \( \chi_\mu \). The isomorphism \( M \cong V(\mu) \) forces \( \chi_\lambda = \chi_\mu \), which in turn is equivalent to \( \mu \in \mathcal{W} \cdot \lambda \) by the fact mentioned above (Corollary 8.24). This proves our assertions about irreducible subquotients of \( V \).

In particular, since \( \mathcal{M} = \{ \mu \in \mathfrak{h}^* \mid \mu \leq \lambda \} \cap \mathcal{W} \cdot \lambda \) is a finite set and \( \dim_k V_\mu < \infty \) for each \( \mu \in \mathfrak{h}^* \) (Proposition 8.3), we may define \( \ell(V) = \sum_{\mu \in \mathcal{M}} \dim_k V_\mu \in \mathbb{Z}_+ \) and similarly for any subrepresentation of \( V \). Note that \( V \) is noetherian, being a cyclic module over the noetherian ring \( U_{\mathfrak{g}} \). Thus, we may choose a series of subrepresentations \( V = V_0 \supset V_1 \supset V_2 \supset \ldots \) with irreducible quotients \( V_i/V_{i+1} \). Since \( V_i/V_{i+1} \cong V(\mu_i) \) with \( \mu_i \in \mathcal{M} \) by the first paragraph of this proof, we see that \( \ell(V) > \ell(V_1) > \ell(V_2) > \ldots \). Therefore, the series must terminate at 0 after at most \( \ell(V) \) steps, which shows that length \( V \leq \ell(V) \).

Finally, the statement that \( V(\lambda) \) occurs exactly once as composition factor of \( V \) is clear, because \( \dim_k V_\lambda = 1 \) (Proposition 8.3), and the statement about the other composition factors has been verified in the first paragraph of this proof. \( \square \)

**Exercises for Section 8.3**

**8.3.1** ("Trick of Lesieur and Croisot"). Let \( R \) be a left noetherian domain. Show that any two \( 0 \neq x, y \in R \) have a nonzero common left multiple: \( Rx \cap Ry \neq 0 \). (Consider the chain \( L_0 \subseteq L_1 \subseteq \ldots \) with \( L_n = \sum_{i=0}^n Rx^i \).)

**8.3.2** (The socle of \( M(\lambda) \)). Show that \( M(\lambda) \) has a unique irreducible subrepresentation. (Use Propositions 8.7 and 8.11 and Exercise 8.3.1.)

**8.4. Finite-Dimensional Irreducible Representations**

Recall that \( \text{Irr}_{\text{fin}} \mathfrak{sl}_2 = \{ V(m) \mid m \in \Lambda_+ = \mathbb{Z}_+ \} \) (Theorem 5.39 and §8.1.4). Generalizing this fact, we will show in this section that, for an arbitrary (semisimple) \( \mathfrak{g} \), we have \( \text{Irr}_{\text{fin}} \mathfrak{g} = \{ V(\lambda) \mid \lambda \in \Lambda_+ \} \), where \( \Lambda_+ \subseteq \mathfrak{h}^* \) is the monoid of dominant weights.
8.4.1. The Main Result

Every $V \in \text{Irr}_{\text{fin}} \mathfrak{g}$ is isomorphic to some $V(\lambda)$. For, $V$ is a highest weight representation (Corollary 8.4), say with highest weight $\lambda$, and hence there is an epimorphism $M(\lambda) \to V$ (Proposition 8.7), whose kernel must be the unique largest proper subrepresentation of $M(\lambda)$. The following theorem determines the weights $\lambda$ occurring in this way.

**Theorem 8.12.** $V(\lambda)$ is finite dimensional if and only if $\lambda \in \Lambda_+$. In this case, the set of weights of $V(\lambda)$ is the union of the ‘$W$-orbits’ $W\lambda'$ with $\lambda' \in \Lambda_+$, $\lambda' \leq \lambda$.

**Proof.** Throughout this proof, we will write $V = V(\lambda)$ and let $\Pi \subseteq \mathfrak{h}^*$ denote the set of weights of $V$. So $\lambda \in \Pi$ and $\mu \leq \lambda$ for all $\mu \in \Pi$. Furthermore, $V = \bigoplus_{\mu \in \Pi} V_\mu$ and all $V_\mu$ are finite dimensional (Proposition 8.3). Thus, finite dimensionality of $V$ amounts to finiteness of $\Pi$.

First, assume that $V$ is finite dimensional. Then Proposition 8.2 tells us that $\lambda \in \Lambda$ and $W\lambda \subseteq \Pi$. Thus, all members of the orbit $W\lambda$ are $\leq \lambda$. On the other hand, $W\lambda$ intersects $\Lambda_+$ in exactly one point and this point is the highest, with respect to $\leq$, in the orbit (Proposition 7.16). Therefore, $\lambda$ must be the point in question; so $\lambda \in \Lambda_+$.

Conversely, assume that $\lambda \in \Lambda_+$. Then $\Pi \subseteq \Lambda$, because all $\mu \in \Pi$ satisfy $\mu \leq \lambda$ and so $\mu \equiv \lambda \mod \Lambda$. We will show that

$$\Pi = \{\mu \in \Lambda \mid w\mu \leq \lambda \text{ for all } w \in W\}$$

(8.17)

$$= \{w\lambda' \mid w \in W, \lambda' \in \Lambda_+, \lambda' \leq \lambda\}.$$  

The second equality, between the two sets on the right, follows from Proposition 7.16 and the same proposition also tells us that $\{\lambda' \in \Lambda_+ \mid \lambda' \leq \lambda\}$ is a finite set. Since $W$ is finite, the last set in (8.17) is finite as well and so (8.17) will imply that $V$ is finite dimensional, completing the proof of the theorem. Observe that (8.17) amounts to the following two assertions: (1) $\Pi$ is stable under $W$ and (2) $\{\lambda' \in \Lambda_+ \mid \lambda' \leq \lambda\} \subseteq \Pi$. Both assertions are consequences of the following

**Claim.** If $\mu \in \Pi$ and $\alpha \in \Delta$, then $\mu - i\alpha \in \Pi$ for all $i \in \mathbb{Z}$ between 0 and $\langle \mu, h_\alpha \rangle$.

Note that $\langle \mu, h_\alpha \rangle \in \mathbb{Z}$, because $\mu \in \Lambda$. Thus, the claim gives $s_\alpha \mu = \mu - \langle \mu, h_\alpha \rangle \alpha \in \Pi$. Since $W$ is generated by the simple reflections $s_\alpha$, we obtain (1). The claim also implies the following fact, which in turn plainly yields (2): if $\nu \in \Lambda_+$ and $\nu \leq \mu$ for some $\mu \in \Pi$, then $\nu \in \Pi$. To prove this, write $\mu = \nu + \sum_{\alpha \in \Delta} z_\alpha \alpha$ with $z_\alpha \in \mathbb{Z}_+$ and argue by induction on $s := \sum_{\alpha \in \Delta} z_\alpha$. The case $s = 0$ is of course clear, being equivalent to if $\mu = \nu$; so assume $s > 0$. Then $(\mu - \nu, \mu - \nu) > 0$ (Proposition 6.12) and so $(\mu - \nu, \alpha) > 0$ for some $\alpha \in \Delta$ with $z_\alpha > 0$. Therefore, $\langle \mu - \nu, h_\alpha \rangle > 0$ and so $\langle \mu, h_\alpha \rangle > \langle \nu, h_\alpha \rangle \geq 0$, where the $\geq$ holds because $\nu \in \Lambda_+$. The claim therefore gives $\mu - \alpha \in \Pi$, and replacing $\mu$ by $\mu - \alpha$ reduces $s$ by 1.
There still remains the task of proving the claim. We first show that \( V \downarrow_{s_\alpha} \) is locally finite for the \( sl_2 \)-triple \( s_\alpha = \mathbb{R} f_\alpha \oplus \mathbb{R} h_\alpha \oplus \mathbb{R} e_\alpha \), \( \subseteq g \) (Theorem 6.10). It is enough to show that \( V \) contains some nonzero finite-dimensional \( s_\alpha \)-subrepresentation, \( T \). For, then \( V = U g, T \) by irreducibility, and \( U g, T \) is the union of the finite-dimensional \( s_\alpha \)-subrepresentations \( U_n T \), where \( (U_n)_{n \geq 0} \) is the standard filtration of \( U g \). To construct \( T \), pick \( 0 \neq v \in V \). Since \( \langle \lambda, h_\alpha \rangle \in \mathbb{Z}_+ \) by assumption on \( \lambda \), the vector \( v_\alpha := f_\alpha^{(\langle \lambda, h_\alpha \rangle)+1} \cdot v \in V \) is either 0 or a maximal vector with weight \( s_\alpha \lambda - \alpha \) by Lemma 8.5. In the latter case, it would follow that \( V = U \cdot v_\alpha \) by irreducibility and so \( V \) would be a highest weight representation with highest weight \( s_\alpha \lambda - \alpha \), which is absurd, because \( s_\alpha \lambda - \alpha \not\subseteq \lambda \). Therefore, we must have \( v_\alpha = 0 \), and hence \( U(sl_\alpha) \cdot v_\alpha = \sum_{i=0}^{\langle \lambda, h_\alpha \rangle} k f_\alpha^i \cdot v \) is the desired finite-dimensional \( sl_\alpha \)-subrepresentation of \( V \downarrow_{s_\alpha} \). It follows that \( V \downarrow_{s_\alpha} \) is a direct sum of finite-dimensional irreducible subrepresentations (Theorem 5.39). We may now copy the argument in the proof of Proposition 8.2(c). In detail, if \( m \geq 0 \), then any \( 0 \neq v \in V \) must have a nonzero component in some irreducible constituent \( V(m') \) of \( V \downarrow_{s_\alpha} \) with \( m' \geq m \). It follows from Proposition 5.37 that \( f_\alpha^i \cdot v \neq 0 \) for all \( 0 \leq i \leq m \) and (8.10) gives the desired conclusion \( \mu - i \alpha \in \Pi \). If \( m \leq 0 \), then one can argue similarly, using \( e_\alpha^{-m} \) in place of \( f_\alpha^m \). This proves the claim, and hence the theorem is proved. \( \square \)

Theorem 8.12 shows that \( \Lambda_+ \) parametrizes the finite-dimensional irreducible representation of \( g \) up to isomorphism:

\[
\begin{array}{ccc}
\Lambda_+ & \sim & \text{Irr}_{\text{fin}} g \\
\downarrow & \downarrow & \downarrow \\
\lambda & \mapsto & V(\lambda)
\end{array}
\]  
(8.18)

Surjectivity holds by the remarks before the statement of the theorem; the map is injective, because \( \lambda \) is determined as the highest weight of \( V(\lambda) \). Theorem 8.12 also shows that \( \Lambda_+ \) consists of weights of finite-dimensional representations of \( g \). Since the weights of any finite-dimensional representation form a \( \mathcal{W} \)-stable subset of \( \Lambda \) (Proposition 8.2) and \( \mathcal{W}(\Lambda_+) = \Lambda \) (Proposition 7.16), we obtain that

\[
\Lambda = \{ \text{weights of finite-dimensional representations of } g \}.
\]  
(8.19)

### 8.4.2. Fundamental Representations of \( sl_{n+1} \)

The representations \( V(\lambda_i) \) for the fundamental weights \( \lambda_1, \ldots, \lambda_n \) are called the **fundamental representations** of \( g \). In this subsection, we determine these representations for \( g = sl_{n+1} \).

Let us start with a few reminders (§§6.4.2, 7.1.4). We work with the Cartan subalgebra \( h \) consisting of all diagonal matrices in \( sl_{n+1} \) and we view \( h \subseteq b_{n+1} \), the Lie algebra of all diagonal matrices in \( gl_{n+1} \). Let \( (e_i)_{i=1}^{n+1} \) denote the dual basis of \( b_{n+1}^* \)}
for the standard basis \((e_i)_{i=1}^{n+1}\) of \(\mathfrak{g}_{n+1}\) and put \(E = \{\sum_{i=1}^{n+1} x_i e_i \mid \sum_{i=1}^{n} x_i = 0\} \subseteq \mathfrak{h}_n^\ast\). The restriction map \(\mathfrak{g}_{n+1} \to \mathfrak{h}_n^\ast\) yields an isomorphism \(E \cong \mathfrak{h}_n^\ast\), which we will view as an identification. The set of roots for \(\mathfrak{h}\) is then 
\[ \Phi = \{e_i - e_j \mid 1 \leq i \neq j \leq n+1\} \subseteq E; \]
a base for \(\mathfrak{h}\) is provided by \(\Delta = \{\alpha_i \mid 1 \leq i \leq n\}\) with \(\alpha_i = e_i - e_{i+1}\) (Example 7.3); the Weyl group \(W\) is the symmetric group \(S_{n+1}\), acting by permuting \(\{e_i\}_{i=1}^{n+1}\) in the standard way; and the fundamental weights for the base \(\Delta\) are given by \(\lambda_i = \mu_1 + \cdots + \mu_i\) with \(\mu_1 = e_1 - e_2\) (Exercise 8.4.3). Thus, \(\Lambda^\ast := \mathfrak{h}_n^\ast \mathfrak{g}\) yields an isomorphism

In order to find irreducible representations for \(\mathfrak{sl}_{n+1}\) other than \(\mathfrak{sl}_2\), we follow the approach taken earlier for \(\mathfrak{sl}_2\) and start with the defining representation \(V = \mathbb{C}^{n+1}\), which is easily seen to be irreducible (§6.4.2). Hence, \(V\) must be a highest weight representation with highest weight belonging to \(\Lambda^\ast\). In order to identify this weight, let \((e_k)_{k=1}^{n+1}\) be the standard basis of \(V\), with \(e_{i,j} \cdot e_k = e_i \delta_{j,k}\). For \(h \in \mathfrak{h}\), this gives

\[ h \cdot e_k = (e_k, h)e_k = (\mu_k, h)e_k. \]
Thus, each \(e_k\) is a weight vector with weight \(\mu_k\). Since \(\mu_k = \mu_{k-1} - \alpha_k\), we have \(\mu_1 = \lambda_1 > \mu_2 > \cdots\) So the highest weight is \(\lambda_1\), and hence \(V = V(\lambda_1)\).

Next, we have a look at the exterior powers \(\Lambda^i V\) with \(i \leq n + 1\). A basis for \(\Lambda^i V\) is provided by the elements \(e_K := e_{k_1} \wedge e_{k_2} \wedge \cdots \wedge e_{k_i}\) with \(K = (1 \leq k_1 < k_2 < \cdots < k_i \leq n + 1)\). It is straightforward to check that, once again, \(e_K\) is a weight vector, with weight \(\mu_K := \mu_{k_1} + \cdots + \mu_{k_i}\). In particular, \(\mu_{(1,2,\ldots,n+1)} = 0\) and so \(\Lambda^{n+1} V = \mathbb{C} e_{(1,2,\ldots,n+1)} = 1\). For \(i \leq n\), we have \(\mu_{(1,2,\ldots,i)} = \lambda_i\) and \(\mu_{(1,2,\ldots,i)} - \mu_K = \sum \delta_{i,j} - \delta_{k_i} \in \Phi^+\) for \(K \neq (1,2,\ldots,i)\). Thus, \(\lambda_i \geq \mu_K\) for all \(K\). Finally, it is not hard to check that \(e_{(1,2,\ldots,i)}\) generates \(\Lambda^i V\) (Exercise 8.4.3). Thus, \(\Lambda^i V\) is a finite-dimensional highest weight representation with highest weight \(\lambda_i\):

\[ \Lambda^i V = V(\lambda_i) \quad (i = 1, \ldots, n). \]

### 8.4.3. Weight Diagrams

A picture showing the set \(\Pi = \Pi(V)\) of weights of given \(V \in \text{Rep}_\text{fin} \mathfrak{g}\) is called a weight diagram of \(V\). Ideally, the diagram also indicates the multiplicity of each weight. For \(V = V(\lambda)\) with \(\lambda \in \Lambda^\ast\), we know by Theorem 8.12 that \(\Pi\) is the union of the \(W\)-orbits of the elements \(\lambda' \in \Lambda^\ast\) such that \(\lambda' \leq \lambda\). The lemma below gives some further properties of \(\Pi\) in this case, but we need some terminology. If \(\lambda, \mu, \nu \in \Lambda\) and \(\nu - \mu \in \mathbb{Z}_+ \lambda\), then the \(\lambda\)-string between \(\mu\) and \(\nu\) consists of the elements \(\mu, \mu + \lambda, \ldots, \nu - \lambda, \nu\) of \(\Lambda\). An element \(\mu \in \Pi\) is called \(\Phi\)-extremal if, for any \(\alpha \in \Phi\), not both of \(\mu \pm \alpha\) can belong to \(\Pi\).

**Lemma 8.13.** Let \(\lambda \in \Lambda^\ast\) and let \(\Pi\) denote the set of weights of \(V(\lambda)\). Then:

(a) \(\Pi\) is \(\Phi\)-convex: if \(\mu, \nu \in \Pi\) and \(\nu - \mu \in \mathbb{Z}_+ \alpha\) for some \(\alpha \in \Phi\), then the entire \(\alpha\)-string between \(\mu\) and \(\nu\) belongs to \(\Pi\).

(b) The \(\Phi\)-extremal elements of \(\Pi\) are exactly the elements of the orbit \(W\lambda\).
Proof. Put \( V = V(\lambda) \).

(a) Consider the \( sl_3 \)-triple \( s_\alpha = \mathbb{R}f_\alpha \oplus \mathbb{R}h_\alpha \oplus \mathbb{R}e_\alpha \subseteq g \), with \( f_\alpha \in g_{-\alpha} \) and \( e_\alpha \in g_{\alpha} \), and write \( \nu = \mu + k\alpha \) \( (k \in \mathbb{Z}) \). Thus, \( \langle \nu, h_\alpha \rangle = \langle \mu, h_\alpha \rangle + 2k \). Assume first that \( |\langle \nu, h_\alpha \rangle| \geq |\langle \mu, h_\alpha \rangle| \) and let \( 0 \neq v \in V_\nu \). Then \( v \) has a nonzero component in some irreducible constituent \( V(m) \) of \( V \supseteq s_\alpha \) and Proposition 5.37 further gives that \( f_\alpha^l v \neq 0 \) for \( 0 \leq l \leq k \). Since \( f_\alpha^l v \in g_{\alpha-l\alpha}, V_\nu \subseteq V_{\nu-l\alpha} \), we obtain that \( v, v-\alpha, \ldots, v-k\alpha = \mu \) all are weights of \( V \). If \( |\langle \nu, h_\alpha \rangle| \leq |\langle \mu, h_\alpha \rangle| \), then consider \( 0 \neq v \in V_\mu \) and argue similarly to conclude that \( 0 \neq e_\alpha^l v \in V_{\mu+l\alpha} \) for \( 0 \leq l \leq k \).

(b) The set of \( \Phi \)-extremal elements of \( \Pi \) is stable under the operation of \( \mathcal{W} \), because both \( \Phi \) and \( \Pi \) are \( \mathcal{W} \)-stable. Also, \( \lambda \) is certainly \( \Phi \)-extremal, being the highest weight of \( V \) for any \( \alpha \in \Phi \), one of \( \lambda + \alpha > \lambda \), and hence it cannot belong to \( \Pi \). It follows that the orbit \( \mathcal{W}\lambda \) consists of \( \Phi \)-extremal elements of \( \Pi \). Now let \( \mu \in \Pi \) be \( \Phi \)-extremal. We wish to show that \( \mu \in \mathcal{W}\lambda \). Replacing \( \mu \) by an element in its \( \mathcal{W} \)-orbit if necessary, we may assume that \( \mu \in \Lambda_+ \) (Proposition 7.16). Thus, \( \langle \mu, h_\alpha \rangle \in \mathbb{Z}_+ \) for all \( \alpha \in \Phi_+ \). If \( \mu + \alpha \in \Pi \), then \( \mu \) would be in the interior of the \( \alpha \)-string between the weights \( \mu + \alpha \) and \( s_\alpha(\mu + \alpha) = \mu - (\langle \mu, h_\alpha \rangle + 1)\alpha \), which lies entirely in \( \Pi \) by (a). But this contradicts our assumption that \( \mu \) is \( \Phi \)-extremal; so \( \mu + \alpha \not\in \Pi \) for all \( \alpha \in \Phi_+ \). Thus, if \( 0 \neq v \in V_\mu \), then \( \mathfrak{n}_+, v = 0 \) and we also have \( V = \mathfrak{u}_g, v \) by irreducibility. Therefore, \( \mu \) is the highest weight: \( \mu = \lambda \). This finishes the proof.

As an illustration of the foregoing, the weight diagram of the irreducible \( sl_3 \)-representation \( V(\lambda) \) with \( \lambda = 3\lambda_1 + 5\lambda_2 \) is displayed in Figure 8.1. The multiplicities range from 1 for the weights in the outer shell to 4 for the innermost shell as indicated in the diagram; these multiplicities will be justified later (Example 8.33). The gray region is the fundamental Weyl chamber \( W(\Delta) \), with walls, for the base \( \Delta = \{\alpha_1, \alpha_2\} \), as in Figure 7.2. The dots in this region are exactly the weights \( \lambda' \in \Lambda_+ \) such that \( \lambda' \leq \lambda \) as in Theorem 8.12. The union of the \( \mathcal{W} \)-orbits of these \( \lambda' \) yields the entire diagram; the red dots in particular form the orbit \( \mathcal{W}\lambda \), the extremal points of the diagram. For another weight diagram, in dimension three, see Figure 8.3.

**Exercises for Section 8.4**

**8.4.1** (Kernels of finite-dimensional irreducible representations). Show that the intersection \( \bigcap_{\lambda \in \Lambda_+} \ker_V V(\lambda) \) is the zero-ideal of \( \mathfrak{u}_g \) (Use Proposition 5.28.)

**8.4.2** (The highest root). (a) Show that if \( g \) is simple, then there is a root \( \alpha \in \Phi_+ \) such that \( \alpha \geq \alpha \) for all \( \alpha \in \Phi \). Moreover, \( \alpha \in \Lambda_+ \) and \( g_{\text{ad}} = V(\alpha) \).

(b) For \( g = sl_{n+1} \), show that \( \alpha = e_1 - e_{n+1} = \sum_{i=1}^n \alpha_i = \lambda_1 + \lambda_n \) (§8.4.2).

**8.4.3** (A detail check). Let \( V = \bigoplus_{i=1}^{n+1} \mathbb{R}e_i \) be the defining representation of \( sl_{n+1} \) (§8.4.2). Check that \( \lambda' \in \text{Rep} sl_{n+1} \) is generated by \( e_1 \land e_2 \land \cdots \land e_i \).
8.4.4 (0 as a weight). Let \( \lambda \in \Lambda_+ \) and let \( \Pi \) denote the set of weights of \( V(\lambda) \). Show that \( 0 \in \Pi \) if and only if \( \lambda \in L_+ \).

8.4.5 (Duals of finite-dimensional irreducible representations). Let \( \lambda \in \Lambda_+ \) and let \( w_0 \in \mathcal{W} \) be the longest element (Exercise 7.2.6). Show:

(a) \( w_0^* \lambda \) is the lowest weight of \( V(\lambda) \): all other weights \( \mu \) satisfy \( \mu \preceq w_0^* \lambda \).
(b) \( V(\lambda)^* \cong V(-w_0 \lambda) \).

8.5. The Representation Ring

Recall that \( \mathcal{R}(\mathfrak{g}) \), by definition, is the abelian group with generators \([V]\) for \( V \in \text{Rep}_{\text{fin}} \mathfrak{g} \) and a relation \([V] = [U] + [W]\) for each short exact sequence \( 0 \to U \to V \to W \to 0 \) in \( \text{Rep}_{\text{fin}} \mathfrak{g} \). Multiplication in \( \mathcal{R}(\mathfrak{g}) \) comes from the tensor product of representations: \([V] \cdot [W] = [V \otimes W]\) for \( V, W \in \text{Rep}_{\text{fin}} \mathfrak{g} \). All this is true for any Lie algebra (§5.5.8), but more can be said in the semisimple case. In this section, we generalize our earlier description of \( \mathcal{R}(\mathfrak{sl}_2) \) (§5.7.7).

8.5.1. Group Structure

In view of (8.18) the (group) isomorphism of Proposition 1.46 now takes the following form, with \( m_\nu(\lambda) := \mu(V(\lambda), V) \) denoting the multiplicity of the irreducible
representation \( V(\lambda) \) as a composition factor of \( V \):

\[
\mathcal{R}(\mathfrak{g}) \xrightarrow{\psi} \mathbb{Z}^{\oplus \Lambda_+}
\]

(8.20)

Thus, the classes \([V(\lambda)]\) with \( \lambda \in \Lambda_+ \) form a \( \mathbb{Z} \)-basis of \( \mathcal{R}(\mathfrak{g}) \). Furthermore, we have the following consequence of Weyl’s Theorem.

**Lemma 8.14.** Let \( V, W \in \text{Rep}_{\text{fin}} \mathfrak{g} \). Then \( V \cong W \) if and only if \([V] = [W]\) in \( \mathcal{R}(\mathfrak{g}) \).

**Proof.** To prove the non-trivial direction, that \([V]\) determines \( V \) up to isomorphism, observe that (8.20) says the multiplicities \( m_V(\lambda) \) are determined by \([V]\). Furthermore, by Weyl’s Theorem, these multiplicities determine \( V \) up to isomorphism:

\[
V \cong \bigoplus_{\lambda \in \Lambda_+} (\lambda) \oplus m_V(\lambda).
\]

\( \square \)

### 8.5.2. Ring Structure

Letting \( \mathbb{Z} \Lambda \) denote the integral group ring of \( \Lambda \) and \((\mathbb{Z} \Lambda)^W\) its subring of \( W \)-invariants (§7.4.4), we define the **formal character** of \( V \in \text{Rep}_{\text{fin}} \mathfrak{g} \) by

\[
\text{ch}_V \overset{\text{def}}{=} \sum_{\lambda \in \Lambda} (\dim_k V_\lambda) x^\lambda \in (\mathbb{Z} \Lambda)^W
\]

The statement that \( \text{ch}_V \) is \( W \)-invariant is a reformulation of part (c) of Proposition 8.2. Moreover, (8.11) is equivalent to the following relations in \((\mathbb{Z} \Lambda)^W\):

\[
\begin{align*}
\text{ch}(V \oplus W) &= \text{ch} V + \text{ch} W, \\
\text{ch}(V \otimes W) &= \text{ch} V \cdot \text{ch} W.
\end{align*}
\]

The formal characters of the finite-dimensional irreducible representations will be of particular interest to us; they will be denoted by

\[
\text{ch}_\lambda := \text{ch} V(\lambda) \quad (\lambda \in \Lambda_+).
\]

Thus, \( \text{ch}_0 = x^0 = 1 \). Weyl’s Character Formula (Section 8.7) will determine the characters \( \text{ch}_\lambda \) in general.

**Theorem 8.15.** Formal characters give a ring isomorphism

\[
\text{ch}: \mathcal{R}(\mathfrak{g}) \cong (\mathbb{Z} \Lambda)^W, \quad [V] \mapsto \text{ch} V.
\]

If \( \{\lambda_i\}_1^n \) are fundamental weights as in (8.1), then \((\mathbb{Z} \Lambda)^W = \mathbb{Z}[\text{ch}_{\lambda_1}, \ldots, \text{ch}_{\lambda_n}] \), a polynomial ring in \( n \) variables over \( \mathbb{Z} \).

**Proof.** Clearly, \( \text{ch}(V) \) depends only on the isomorphism type of \( V \). Moreover, if \( 0 \to U \to V \to W \to 0 \) is a short exact sequence in \( \text{Rep}_{\text{fin}} \mathfrak{g} \), then \( V \cong U \oplus W \) by complete reducibility (Weyl’s Theorem) and so \( \text{ch}(V) = \text{ch}(U) + \text{ch}(V) \) by the direct
sum formula in (8.21). Therefore, $\text{ch}$ yields a well-defined group homomorphism on $R(g)$. The second formula in (8.21) says that it is in fact a ring homomorphism.

To prove bijectivity, recall from Proposition 8.3 that the highest weight $\lambda$ of $V(\lambda)$ has multiplicity 1 and any other weight $\mu$ satisfies $\mu \preceq \lambda$. Thus, for $\lambda \in \Lambda_+$,

$$\text{ch}_\lambda = x^\lambda + \sum_{\mu \preceq \lambda} z_{\lambda, \mu} x^\mu \in (\mathbb{Z}\Lambda)^W.$$  

This is exactly the form required in Theorem 7.17; so this result gives that the formal characters $\text{ch}_{\lambda_i}$ $(i = 1, \ldots, n)$ generate the ring $(\mathbb{Z}\Lambda)^W$ and are algebraically independent over $\mathbb{Z}$. Since all $\text{ch}_{\lambda_i}$ belong to the image of the ring map $\text{ch}$, it also follows that $\text{ch}$ is surjective.

For injectivity, recall that $\{[V(\lambda)]\}_{\lambda \in \Lambda_+}$ is a $\mathbb{Z}$-basis of $R(g)$; so we need to make sure that the collection $(\text{ch}_{\lambda_i})_{i \in \Lambda_+}$ is $\mathbb{Z}$-independent. But this is immediate from the form of $\text{ch}_1$ as displayed above; see the paragraph after (7.25) in the proof of Theorem 7.17. This completes the proof. \hfill $\square$

The theorem recovers the earlier isomorphism $R(\mathfrak{sl}_2) \cong (\mathbb{Z}\Lambda)^W \cong \mathbb{Z}[t + t^{-1}]$ (Proposition 5.46). Indeed, with $\lambda_1 = 1$ ($\S 8.1.4$) and $t = x^1$, the group ring $\mathbb{Z}\Lambda$ becomes the Laurent polynomial ring $\mathbb{Z}[t^\pm]$ and $t + t^{-1} = \text{ch}_{\lambda_1}$ is the formal character of the defining representation $V(1)$ of $\mathfrak{sl}_2$. The following very useful corollary, which is an immediate consequence of Theorem 8.15 and Lemma 8.14, was also noted earlier for $\mathfrak{sl}_2$ (Lemma 5.44).

**Corollary 8.16.** If $V, W \in \text{Rep}_{\text{fin}} g$, then $V \cong W$ if and only if $\text{ch} V = \text{ch} W$.

### 8.5.3. Fundamental Characters and Representation Ring of $\mathfrak{sl}_{n+1}$

We will now determine the characters $\text{ch}_{\lambda_i}$ $(i = 1, \ldots, n)$ of the fundamental representations for $\mathfrak{sl}_{n+1}$. Recall that $\lambda_i = \mu_1 + \cdots + \mu_i$ with $\mu_k = e_k - \frac{1}{n+1} \sum_{j=1}^{n+1} e_j$ ($\S 8.4.2$). Thus, $\sum_{k=1}^{n+1} \mu_k = 0$ and $\Lambda = \bigoplus_{i=1}^{n+1} \mathbb{Z}\lambda_i = \bigoplus_{i=1}^{n+1} \mathbb{Z}\mu_i$. Putting $x_i = x^{\mu_i}$, we have $x_1 x_2 \cdots x_{n+1} = 1$ and

$$\mathbb{Z}\Lambda = \mathbb{Z}[x_1^\pm, \ldots, x_n^\pm] = \mathbb{Z}[x_1, \ldots, x_n, x_{n+1}].$$

Recall further that $V(\lambda_i) = \Lambda^i V = \bigoplus_{K \subseteq [n+1], |K| = i} \mathbb{C} e_K$ and each $e_K$ is a weight vector with weight $\mu_K = \sum_{k \in K} \mu_k$ ($\S 8.4.2$). Therefore, $\text{ch}_{\lambda_i}$ is the $i$th elementary symmetric polynomial in $x_1, \ldots, x_{n+1}$:

$$\text{ch}_{\lambda_i} = \sum_{K \subseteq [n+1], |K| = i} x^{\mu_K} = \sum_{K \subseteq [n+1], |K| = i} \prod_{k \in K} x_k = e_i(x_1, \ldots, x_{n+1}).$$

The Weyl group $W = S_{n+1}$ acts on $\mathfrak{h}^* \cong E \subseteq \mathfrak{h}^*_{n+1}$ by permuting the basis $(e_i)_{i=1}^{n+1}$ of $\mathfrak{h}^*_{n+1}$ in the usual fashion; $\mathfrak{h}^*_{n+1}$ is the standard permutation representation and $E$ is the standard representation of $S_{n+1}$ over $\mathbb{R}$ ($\S 3.2.4$). Thus, $W$ acts on $\Lambda$ by $w \cdot \mu_i = \mu_{wi}$ and on $\mathbb{Z}\Lambda$ by $w \cdot x_i = x_{wi}$. The $i$th elementary symmetric polynomial
8.5.1 (Duality). (a) Show that the map \( .^\ast : \mathcal{R}(\mathfrak{g}) \to \mathcal{R}(\mathfrak{g}) \) that is defined by \([V] \mapsto [V^\ast]\) for \(V \in \text{Rep}_{\text{fin}} \mathfrak{g}\) is an involution, that is, a ring map such that \( .^\ast = \text{Id} \).

(b) Let \( .^\ast \) denote the standard involution (3.28) of the group ring \(\mathbb{Z}\Lambda\); so \((x^\lambda)^\ast = x^{-\lambda}\) for \(\lambda \in \Lambda\). Show that \(\text{ch}(V^\ast) = (\text{ch} V)^\ast\) for \(V \in \text{Rep}_{\text{fin}} \mathfrak{g}\).

8.5.2 (The adjoint representation of \(\mathfrak{sl}_{n+1}\)). For \(\mathfrak{g} = \mathfrak{sl}_{n+1}\), show that \(\text{ch}_{\text{ad}} = \text{ch}_{\text{ad}}^1 \cdot \text{ch}_{\text{ad}}^{n-1}\). Conclude that \(V(\Lambda_1) \otimes V(\Lambda_n) \cong \mathfrak{g}_{\text{ad}} \oplus \mathbb{1}\).

8.5.3 (Decomposing tensor products). Let \(\lambda = \sum_{i=1}^{n} z_i \lambda_i \in \Lambda_+ \) (\(z_i \in \mathbb{Z}_+\)). Show that \(V(\lambda)\) is an irreducible constituent of \(V(\lambda_1) \otimes \cdots \otimes V(\lambda_n)\), with multiplicity 1, and all other irreducible constituents have the form \(V(\mu)\) with \(\mu \in \Lambda_+, \mu \preceq \lambda\).

8.6. The Center of the Enveloping Algebra

In this section, we will write \(U = U_0\) and \(\mathcal{Z} = \mathcal{Z}U\), the center of \(U\). Unless explicitly suspended, the notation and hypotheses laid out at the beginning of this chapter remain in effect.

For any Lie algebra \(\mathfrak{g}\), the center \(\mathcal{Z}\) coincides with the algebra of \(\mathfrak{g}\)-invariants under the adjoint action of \(\mathfrak{g}\) on \(U\) (5.38):

\[\mathcal{Z} = U^0.\]

Our goal in this section is to give another invariant theoretic description of \(\mathcal{Z}\) that is specific to a semisimple Lie algebra \(\mathfrak{g}\). This description will in particular show that \(\mathcal{Z}\) is a polynomial algebra over \(k\) in rank \(\mathfrak{g}\) many variables (Corollary 8.25). The symmetrization isomorphism \(\omega: \text{Sym}_{\mathfrak{g}} \xrightarrow{\cong} U\) (Proposition 5.27) will play an important role. Recall that \(\omega\) is defined for any \(\mathfrak{g}\) (over a field of characteristic 0) and is equivariant for the adjoint \(\mathfrak{g}\)-actions. Hence, \(\omega\) restricts to an isomorphism \((\text{Sym}_{\mathfrak{g}})^0 \xrightarrow{\cong} U^0 = \mathcal{Z}\), which is however generally not an algebra map. A remarkable theorem of Duflo [63] constructs, for any finite-dimensional Lie algebra \(\mathfrak{g}\), a linear transformation \(D: \text{Sym}_{\mathfrak{g}} \to \text{Sym}_{\mathfrak{g}}\) such that \(\omega \circ D\) gives an algebra isomorphism

\[(\text{Sym}_{\mathfrak{g}})^0 \xrightarrow{\cong} \mathcal{Z}.\]

This isomorphism, called the Duflo isomorphism, is beyond the scope of this book. The earlier special case, for semisimple \(\mathfrak{g}\), that we are about to construct in this section is due to Harish-Chandra [96].
8.6.1. Invariant Polynomial Functions

In this subsection, \( g \in \text{Lie}_k \) need not be semisimple and \( k \) need not be algebraically closed, but we assume \( \dim_k g < \infty \) and we continue to assume \( \operatorname{char} k = 0 \). Throughout, let \( V \in \text{Rep}_{\text{fin}} g \) and let \( O(V) = \text{Sym}^* V^\ast \) denote the algebra of polynomial functions \( V \to k \) (Appendix C.3). Then \( O(V) \in g_{\text{Alg}} \) (5.5.5) via the dual representation,

\[
x_{V^*} = (-x_V)^\ast \quad (x \in g).
\]

Thus, we have graded derivations \( x_{O(V)} \in \text{Der} O(V) \) extending the \( g \)-action (8.22) on \( V^* \subseteq O(V) \). We will write \( x_{O(V)}(f) \) for \( f \in O(V) \).

**Polarization.** The following isomorphism for the \( d \)-th homogeneous component \( O^d(V) = \text{Sym}^d V^* \) was discussed earlier (Lemma 3.39):

\[
O^d(V) \leftrightarrow \{ \text{symmetric multilinear maps } V^d \to k \}
\]

\[
(x.g)(v_1, \ldots, v_d) = -\sum_{i=1}^d g(v_1, \ldots, x.v_i, \ldots, v_d) \quad (x \in g, v_i \in V).
\]

Note that the action of \( g \) commutes with the place permutation action of \( S_d \); so \( x.g \) will be symmetric if \( g \) is. If \( g \in \text{MultLin}(V^d, k) \) is not necessarily symmetric, then \( \overline{g} = \frac{1}{d!} \sum_{s \in S_d} g \circ s \) is a symmetric multilinear map \( V^d \to k \) satisfying \( \overline{g}(v, v, \ldots, v) = g(v, v, \ldots, v) \) for all \( v \in V \). Moreover, \( x.\overline{g} = \overline{x.g} \) for \( x \in g \).

**The Adjoint Representation.** The following lemma constructs invariant polynomial functions in the special case of the adjoint representation, \( V = g_{\text{ad}} \). We will see in §8.6.4 that the functions in the lemma span the space \( O(g)^g \) of all invariant polynomial functions on \( g = g_{\text{ad}} \).

**Lemma 8.17.** Let \( V \in \text{Rep}_{\text{fin}} g \). For any \( d \in \mathbb{Z}_+ \) and any \( x \in g \), consider the power \( x_V^d \in \text{End}_k(V) \) and put \( x_{V,d}(x) = \text{trace}(x_V^d) \). Then \( x_{V,d} \in O^d(g)^g \).
Proof. Observe that \( \chi_{V,d}(x) = g(x, \ldots, x) \) with \( g \in \text{MultLin}(\mathfrak{g}^d, \mathbb{K}) \) being defined by \( g(x_1, \ldots, x_d) = \text{trace} \left( (x_1)_V(x_2)_V \cdots (x_d)_V \right) \). It follows that \( \chi_{V,d} \in \mathcal{O}_d(\mathfrak{g}) \). In order to prove \( \mathfrak{g} \)-invariance, we calculate

\[
(x.g)(v_1, \ldots, v_d) = - \sum_{i=1}^{d} \text{trace} \left( (x_1)_V \cdots (x_{i-1})_V [x, x_i]_V (x_{i+1})_V \cdots (x_d)_V \right)
\]

\[
= - \text{trace} \left( (x_1)_V \cdots (x_d)_V \right) + \text{trace} \left( (x_1)_V \cdots (x_d)_V [x, \cdot]_V \right)
\]

\[
= 0.
\]

Here, the second equality holds because the first sum above telescopes, and the last equality holds because the trace of a product is invariant under cyclic permutations of the factors. \( \square \)

8.6.2. Elementary Automorphisms of \( \mathfrak{g} \)

It will turn out to be advantageous to express \( \mathfrak{g} \)-invariants as invariants of a certain group. To define this group, observe that if \( d \) is a nilpotent element of an arbitrary \( A \in \text{Alg}_\mathbb{K} \), then we obtain a unit of \( A \) by defining \( e^d := \sum_{n \geq 0} \frac{d^n}{n!} \in A^\times \); the inverse is given by \( (e^d)^{-1} = e^{-d} \).

Let us now revert to the standing hypotheses of this chapter; so \( \mathfrak{g} \in \text{Lie}_\mathbb{K} \) is semisimple and \( \mathbb{K} \) is algebraically closed of characteristic 0. Consider the algebra \( A = \text{End}_\mathbb{K} \mathfrak{g} \) and let \( d \in \text{Der} \mathfrak{g} \subseteq A \) be nilpotent. Then the Leibniz formula implies that \( e^d \in \text{Aut} \mathfrak{g} \), the group of Lie algebra automorphism of \( \mathfrak{g} \) (Exercise 8.6.2). Moreover, all nilpotent derivations \( d \in \text{Der} \mathfrak{g} \) have the form \( d = \text{ad} x \) for ad-nilpotent elements \( x \in \mathfrak{g} \) (Corollary 6.4). The resulting automorphisms \( e^\text{ad} x \in \text{Aut} \mathfrak{g} \) are called elementary. We put

\[
\text{Aut}_e \mathfrak{g} \overset{\text{def}}{=} \left\langle e^\text{ad} x \mid x \in \mathfrak{g} \text{ ad-nilpotent} \right\rangle \leq \text{Aut} \mathfrak{g}
\]

As was mentioned earlier (§6.3.1), any two Cartan subalgebras of \( \mathfrak{g} \) are conjugate under the group \( \text{Aut}_e \mathfrak{g} \). We will refer to this fact, which we will not prove, as the Conjugacy Theorem; see Bourbaki [28] or Humphreys [101].

Proposition 8.18. Let \( G = \text{Aut}_e \mathfrak{g} \) and \( H = \text{stab}_G \mathfrak{h} \). Then:

(a) The image of the restriction map \( H \to \text{GL}(\mathfrak{h}), g \mapsto g|_\mathfrak{h} \), contains \( \mathcal{W} \).

(b) \( \bigcup_{g \in G} g \mathfrak{h} \) is a Zariski dense subset of \( \mathfrak{g} \).

Proof. (a) It suffices to show that, for any \( \alpha \in \Phi \), there exists \( g_\alpha \in \text{Aut}_e \mathfrak{g} \) such that \( g_\alpha|_\mathfrak{h} = s_\alpha \). Consider the \( \mathfrak{s}_\mathfrak{l}_2 \)-triple \( s_\alpha = \mathbb{K} f_\alpha \oplus \mathbb{K} h_\alpha \oplus \mathbb{K} e_\alpha \subseteq \mathfrak{g} \), with \( f_\alpha \in \mathfrak{g}_{-\alpha} \), \( e_\alpha \in \mathfrak{g}_\alpha \) and \( h_\alpha = [e_\alpha, f_\alpha] \in \mathfrak{h} \). Since both \( e_\alpha \) and \( f_\alpha \) are ad-nilpotent elements of
Finally, consider the linear operator $g\alpha$ in $\mathfrak{g}$ (Exercise 6.3.3), we may put

$$g\alpha := e^{ad e\alpha} e^{ad f\alpha} e^{ad e\alpha} \in \text{Aut}_e \mathfrak{g}. $$

We need to show that $g\alpha h = h - \langle \alpha, h \rangle h\alpha$ for $h \in \mathfrak{h}$; see (8.5). This is obvious if $\langle \alpha, h \rangle = 0$, because $e^{ad e\alpha} h = e^{ad f\alpha} h = h = x_a h$ in this case. Hence, we may assume that $h = h\alpha$ and we need to check that $g\alpha h\alpha = -h\alpha$. But the relation $[e\alpha, h\alpha] = -2e\alpha$ gives $e^{ad e\alpha} h\alpha = h\alpha - 2e\alpha$. Next, $[f\alpha, h\alpha - 2e\alpha] = 2f\alpha - 2h\alpha$ and $[f\alpha, 2f\alpha - 2h\alpha] = -4f\alpha$, which implies $e^{ad f\alpha}(h\alpha - 2e\alpha) = h\alpha - 2e\alpha + 2f\alpha - 2h\alpha - 2f\alpha = -2e\alpha - h\alpha$. Finally, $[e\alpha - 2e\alpha - h\alpha] = 2e\alpha$ and so $e^{ad e\alpha}(-2e\alpha - h\alpha) = -2e\alpha - h\alpha + 2e\alpha = -h\alpha$. Thus, $g\alpha h\alpha = -h\alpha$ as desired.

(b) For any $x \in \mathfrak{g}$, let $c(x; t) = \sum_i c_i(x) t^i \in \mathbb{k}[t]$ denote characteristic polynomial of $ad x$. If $x = x_s + x_n$ is the abstract Jordan decomposition of $x$ (Proposition 6.8), then $c(x; t) = c(x_s; t)$, because $ad x_s = (ad x)_s$ is the semisimple part of the Jordan decomposition of $ad x$. The coefficient functions have the form $c_i(x) = (-1)^{dim \beta_i} g(x_1, x_2, \ldots, x_{dim \beta_i})$, where $g$ is the multilinear function $g(x_1, x_2, \ldots, x_{dim \beta_i}) = \text{trace}(ad x_1 \wedge ad x_2 \wedge \cdots \wedge ad x_{dim \beta_i})$ (Lemma 3.33). It follows that $c_i : \mathfrak{g} \to \mathbb{k}$ is a polynomial function, homogeneous of degree $dim \mathfrak{g} - i$ (§8.6.1). Let $\mathfrak{g}^0(x)$ denote the generalized 0-eigenspace of $ad x$; this is the same as the ordinary 0-eigenspace of $ad x_s$, that is, the centralizer of $x_s$. Then $\min\{i \mid c_i(x) \neq 0\} = \dim_k \mathfrak{g}^0(x) = \dim_k C_{\mathfrak{g}}(x_s)$. Put $r := \min\{i \mid c_i \neq 0\}$ and

$$\mathcal{R} := \{x \in \mathfrak{g} \mid c_r(x) \neq 0\} = \{x \in \mathfrak{g} \mid \dim C_{\mathfrak{g}}(x_s) = r\}. $$

Thus, $\mathcal{R}$ is a nonempty Zariski open subset of $\mathfrak{g}$, and hence $\mathcal{R}$ is Zariski dense (Exercise C.3.3). Therefore, it suffices to show that each $x \in \mathcal{R}$ is contained in $g\mathfrak{h}$ for some $g \in G$. But $x_s \in \mathcal{R}$ and, since $x_s$ is contained in some Cartan subalgebra of $\mathfrak{g}$ (§6.3.1), we have $x_s \in g\mathfrak{h}$ for some $g \in G$ by the Conjugacy Theorem. For any $h \in \mathfrak{h}$, the centralizer $C_{\mathfrak{g}}(h)$ contains $\mathfrak{h}$ and we also know that there are elements $h \in \mathfrak{h}$ such that $C_{\mathfrak{g}}(h) = \mathfrak{h}$ (Exercise 6.3.4). Therefore, we must have $C_{\mathfrak{g}}(x_s) = g\mathfrak{h}$ (and, furthermore, $r = \dim \mathfrak{h} = \text{rank}\, \mathfrak{g}$). In particular, $x \in g\mathfrak{h}$, finishing the proof. \hfill $\Box$

### 8.6.3. $\mathfrak{g}$-Invariants and $\text{Aut}_e \mathfrak{g}$-$\mathfrak{g}$-Invariants

Let $x \in \mathfrak{g}$ be ad-nilpotent. Then, for any $V \in \text{Rep}_{\text{fin}} \mathfrak{g}$, the operator $x_V \in \text{End}_k V$ is nilpotent by preservation of Jordan decomposition (Proposition 6.8). Thus, we may consider the linear operator $e^{x_V} = \sum_{n \geq 0} x_V^n / n! \in \text{GL}(V)$. The operator $e^{x_V} \in \text{GL}(V)$ is clearly also well-defined if the representation $V$ is merely locally finite.

**Lemma 8.19.** Let $V \in \text{Rep} \mathfrak{g}$ be locally finite and let $v \in V$. Then $v \in V^\mathfrak{g}$ if and only if $e^{x_V} v = v$ for all ad-nilpotent $x \in \mathfrak{g}$.

---

2The elements of $\mathcal{R}$ are called *regular*. 
Proof. We may assume that $V$ is finite dimensional. Let $x \in \mathfrak{g}$ be ad-nilpotent. If $x.v = 0$ then $x^nv = 0$ for all $n \geq 0$ and so $e^{xv} = v$. For the converse, note that $0 = e^{xv} - v = (\text{Id}_V + \phi)(x.v)$ where $\phi = \sum_{n \geq 2} x^{n-1} \frac{v}{n!} \in \text{End}_x V$ is nilpotent. Therefore, $\text{Id}_V + \phi$ is invertible, whence $x.v = 0$ for all ad-nilpotent $x \in \mathfrak{g}$. Since $\mathfrak{g}$ is generated by ad-nilpotent elements (Exercise 6.3.3), it follows that $v \in V^\mathfrak{g}$. □

We now concentrate on the case where $V = O(\mathfrak{g})$ is the $\mathfrak{g}$-algebra of a polynomial functions on $\mathfrak{g} = \mathfrak{g}_{\text{ad}}$. Since the $\mathfrak{g}$-action on $O(\mathfrak{g})$ stabilizes all homogeneous components, $O(\mathfrak{g}) \in \text{Rep} \mathfrak{g}$ is locally finite. The following proposition expresses the subalgebra $O(\mathfrak{g})^\mathfrak{g}$ of invariant polynomial functions in terms of elementary automorphisms of $\mathfrak{g}$.

**Proposition 8.20.** The group $G = \text{Aut}_\mathbb{C} \mathfrak{g}$ acts by automorphisms on the algebra $O(\mathfrak{g})$ in such a way that the subalgebras $\mathfrak{g}$-invariants and $G$-invariants coincide:

$$O(\mathfrak{g})^\mathfrak{g} = O(\mathfrak{g})^G = \{ f \in O(\mathfrak{g}) \mid g.f = f \text{ for all } g \in G \}.$$  

The algebra $O(\mathfrak{g})^G$ consists of all polynomial functions on $\mathfrak{g}$ that are constant on $G$-orbits in $\mathfrak{g}$.

**Proof.** The standard group action $\text{Aut} \mathfrak{g} \subset \mathfrak{g}$ gives rise to an action $\text{Aut} \mathfrak{g} \subset \mathfrak{g}^*$ by duality (§3.3.3): $g^* = (g^{-1})^*$ for $g \in \text{Aut} \mathfrak{g}$. This action in turn yields an action of $\text{Aut} \mathfrak{g}$ by graded algebra automorphisms on $O(\mathfrak{g}) = \text{Sym} \mathfrak{g}^*$ by functoriality of $\text{Sym}$:

$$\delta_{O(\mathfrak{g})} = \text{Sym}(g^{-1})^* \quad (g \in \text{Aut} \mathfrak{g}).$$

A polynomial function $f \in O(\mathfrak{g})$ is fixed by the automorphism $\delta_{O(\mathfrak{g})}$ if and only if $f$ is constant on the $(g)$-orbit in $\mathfrak{g}$. We will write $g.f = \delta_{O(\mathfrak{g})}(f)$.

All this does of course also apply to the subgroup $G \leq \text{Aut} \mathfrak{g}$. In particular, if $x \in \mathfrak{g}$ is ad-nilpotent, then $(e^{\text{ad}x})_{O(\mathfrak{g})} = \text{Sym}(e^{-\text{ad}x})^* = \text{Sym} e^{(-\text{ad}x)^*}$, where the second equality holds because the duality functor $\cdot^*$ is $k$-linear on Hom-spaces and commutes with powers. Furthermore, $(-\text{ad}x)^* = x_{\text{ad}}^*$ by (8.22) and $\text{Sym} e^{x_{\text{ad}}^*} = e^{x_{\text{ad}}}^*$, where $x_{\text{ad}} \in \text{Der} O(\mathfrak{g})$ is the graded (and locally nilpotent) derivation of $O(\mathfrak{g})$ coming from the representation $\mathfrak{g} = \mathfrak{g}_{\text{ad}}$ as in §8.6.1. The last equality can be checked on the generating subspace $\mathfrak{g}^\mathfrak{g} \subseteq \text{Sym} \mathfrak{g}^* = O(\mathfrak{g})$, where it is obvious. Thus,

$$(e^{\text{ad}x})_{O(\mathfrak{g})} = e^{x_{\text{ad}}}.$$  

Now, $O(\mathfrak{g})^\mathfrak{g} = \{ f \in O(\mathfrak{g}) \mid e^{x_{\text{ad}}}(f) = f \text{ for all ad-nilpotent } x \in \mathfrak{g} \}$ by Lemma 8.19, and this set is equal to $O(\mathfrak{g})^G$ by the above equality. This proves the proposition. □

**8.6.4. Invariants in the Symmetric Algebra**

We now have a closer look at the adjoint action of $\mathfrak{g}$ on the symmetric algebra $\text{Sym} \mathfrak{g}$ as described in §5.5.6 for arbitrary Lie algebras: the adjoint representation gives rise to an action of $\mathfrak{g}$ by derivations on $\text{Sym} \mathfrak{g} = \text{Sym} \mathfrak{g}_{\text{ad}}$, making $\text{Sym} \mathfrak{g}$ a graded
\(g\)-algebra. Our goal in this subsection is to give a description, for semisimple \(g\), of the invariant subalgebra,

\[
(Sym g)^0 = \{ f \in Sym g \mid x.f = 0 \text{ for all } x \in g \}.
\]

The description will be in terms of \(Sym h\), where \(h \subseteq g\) is our fixed Cartan subalgebra. More specifically, the action (8.5) of the Weyl group \(W\) on \(h\) extends uniquely to an action \(W \subset Sym h\) by \(k\)-algebra automorphisms. We will show that \((Sym g)^0\) is isomorphic to the subalgebra of all \(W\)-invariants in \(Sym h\),

\[
(Sym h)^W = \{ f \in Sym h \mid w.f = f \text{ for all } w \in W \}.
\]

To construct this isomorphism, consider the triangular decomposition \(g = n_- \oplus h \oplus n_+\) for the given base \(\Delta\). By functoriality of \(Sym\), the projection map \(g \to h\) along the subspace \(n_- \oplus n_+\) gives rise to an epimorphism of graded \(k\)-algebras,

\[
\psi : Sym g = (n_- \oplus n_+) \oplus Sym h \twoheadrightarrow Sym h.
\]

Here, \(Ker \psi = (n_- \oplus n_+)\) is the ideal of \(Sym g\) that is generated by \(n_- \oplus n_+\).

**Theorem 8.21.** \(\psi|_{(Sym g)^0} : (Sym g)^0 \cong (Sym h)^W\).

**Proof.** We proceed in two steps.

**Step 1: Reformulation in terms of polynomial functions.** The restriction map \(g^* \to h^*\) gives a graded epimorphism of the algebras of polynomial functions on \(g\) and \(h\),

\[
\Psi : O(g) = Sym g^* \to O(h) = Sym h^*, \quad f \mapsto f|_h.
\]

Furthermore, the standard action \(W \subset h^*\) gives rise to an action \(W \subset O(h) = Sym h^*\) by graded \(k\)-algebra automorphisms. The purpose of this step is to show that the theorem is equivalent to the statement that \(\Psi\) restricts to an isomorphism of invariant algebras,

\[
(8.26) \quad \Psi|_{O(g)^0} : O(g)^0 \cong O(h)^W.
\]

In order to justify this claim, we will construct an isomorphism of \(g\)-algebras, \(\kappa : Sym g \to Sym g\), and an isomorphism of \(W\)-algebras, \(\kappa_h : Sym h \to O(h)\), such that \(\kappa_h \circ \psi = \Psi \circ \kappa\). The maps \(\kappa\) and \(\kappa_h\) will then restrict to isomorphisms of invariant subalgebras \((Sym g)^0 \cong O(g)^0\) and \((Sym h)^W \cong O(h)^W\), respectively, proving the equivalence of (8.26) and the isomorphism in the theorem.

In detail, recall from (6.3) that the Killing form \(B\) of \(g\) yields a \(\text{Rep}_g\)-isomorphism \(g_{ad} \cong g_{ad}\), \(x \mapsto B(x, \cdot)\). This map lifts to a \(\text{Alg}_g\)-isomorphism, \(\kappa : Sym g_{ad} \to Sym g_{ad} = O(g)\). Similarly, (8.6) gives a \(\text{Rep}_k\)-\(W\)-isomorphism \(h \cong h^*, \quad h \mapsto B(h, \cdot)|_{h}\), which lifts to a \(\text{Alg}_h\)-isomorphism, \(\kappa_h : Sym h \cong Sym h^* = O(h)\). Finally, \(n_- \oplus n_+ = h^\perp\) for the Killing form \(B\) by (6.11). Therefore, for any \(x \in g\) and \(h \in h\), we compute

\[
\langle(\Psi \circ \kappa)(x), h\rangle = \langle \kappa(x), h\rangle = B(x, h) = B(\psi(x), h) = \langle(\kappa_h \circ \psi)(x), h\rangle.
\]
This shows that \((\Psi \circ \kappa)(x) = (\kappa \circ \psi)(x)\), and since \(\text{Sym}_g\) is generated by \(g\), it follows that \(\kappa \circ \psi = \Psi \circ \kappa\) as desired.

**Step 2: The Chevalley Restriction Theorem.** Now we tackle the isomorphism (8.26). By Proposition 8.20, \(O(g)^0 = O(g)^G\) is the algebra of all polynomial functions on \(g\) that are constant on orbits of the group \(G\) = Aut\(_e\) \(g\) in \(g\). Moreover, clearly, \(O(g)^G \subseteq O(g)^H\), where \(H = \text{stab}_G(h)\) as in Proposition 8.18. By part (a) of this result, \(\mathcal{W}\) is contained in the image of \(H\) under restriction to \(h\). Thus, \(\Psi(O(g)^G) \subseteq \Psi(O(g)^H) \subseteq O(h)^W\). Part (b) of Proposition 8.18 states that the union of all \(G\)-orbits of elements of \(h\) is a Zariski dense subset of \(g\). If \(f \in O(g)^G\) vanishes on \(h\), then \(f\) also vanishes on all \(G\)-orbits of elements of \(h\), and hence \(f = 0\). Thus, \(\Psi\) is injective on \(O(g)^0 = O(g)^G\) and \(\Psi(O(g)^0) \subseteq O(h)^W\).

We still need to prove surjectivity: \(\Psi(O(g)^0) = O(h)^W\). For this, we use the invariant polynomial functions \(\chi_{\mathcal{V},d} \in O^d(g)^0\) from Lemma 8.17. It will suffice to show that, for every \(d \in \mathbb{Z}_+\), the restrictions \(\Psi(\chi_{\mathcal{V},d}) = \chi_{\mathcal{V},d}|_h\) with \(\mathcal{V} \in \text{Rep}_{\text{fin}} g\) span the vector space \(O^d(h)^W\). But \(O^d(h)^W\) is the image of \(O^d(h)\) under the averaging operator \(f \mapsto \frac{1}{|\mathcal{W}|} \sum_{w \in \mathcal{W}} w.f\), were \(\sigma_f = \sum_{f' \in \mathcal{W}} f'^\star\) is the \(\mathcal{W}\)-orbit sum of \(f\) (Proposition 3.16). Furthermore, since \(\Lambda_+\) is a dense subset of \(\mathfrak{h}^*\) for the Zariski topology (Exercise 8.1.2), the functions \(\lambda^d\) with \(\lambda \in \Lambda_+\) generate the vector space \(O^d(h) = \text{Sym}^d \mathfrak{h}^*\) (Proposition 3.37). Therefore, the orbit sums \(\sigma_{\lambda^d}\) \((\lambda \in \Lambda_+)\) span the vector space \(O^d(h)^W\) and our goal is to show that each \(\sigma_{\lambda^d}\) linear combinations of the functions \(\chi_{\mathcal{V},d}|_h\) for \(\mathcal{V} \in \text{Rep}_{\text{fin}} g\). Take \(V(\lambda)\) and recall that the set of weights of \(V(\lambda)\) is the union of the orbits \(\mathcal{W} \mu\) with \(\mu \in \Lambda_+\), \(\mu \leq \lambda\) and that multiplicities \(m(\mu) = \dim_x V(\lambda) \mu\) are constant on each orbit \(\mathcal{W} \mu\) (Theorem 8.12 and Proposition 8.2). Therefore, for any \(h \in \mathfrak{h}\),

\[
\chi_{\mathcal{V}(\lambda),d} (h) = \text{trace} (h^d_{\mathcal{V}(\lambda)}) = \sum_{\mu \in \Lambda_+ \atop \mu \leq \lambda} m(\mu) \sum_{\mu' \in \mathcal{W} \mu} \langle \mu', h \rangle^d
\]

and so, for suitable positive integers \(n(\mu)\),

\[
\chi_{\mathcal{V}(\lambda),d} = \sum_{\mu \in \Lambda_+ \atop \mu \leq \lambda} m(\mu) \sum_{\mu' \in \mathcal{W} \mu} (\mu')^d = \sum_{\mu \in \Lambda_+ \atop \mu \leq \lambda} n(\mu) \sigma_{\mu^d}.
\]

Arguing by induction on the size of the set \(\{ \mu \in \Lambda_+ \atop \mu \leq \lambda\}\), we may assume that all \(\sigma_{\mu^d}\) with \(\mu < \lambda\) in the last sum belong to the linear span of the functions \(\chi_{\mathcal{V},d}|_h\) with \(\mathcal{V} \in \text{Rep}_{\text{fin}} g\). Then the above equation shows that \(\sigma_{\lambda^d}\) does as well, which completes the proof of the theorem.

The proof of Theorem 8.21 also shows that a \(\mathbb{k}\)-linear generating set of the algebra \(O(g)^0\) of invariant polynomial functions on \(g\) is provided by the functions \(\chi_{\mathcal{V},d}(x) = \text{trace}(x^d_{\mathcal{V}})\) with \(\mathcal{V} \in \text{Rep}_{\text{fin}} g\) and \(d \in \mathbb{Z}_+\). The Conjugacy Theorem
only enters the proof of Theorem 8.21 via part (b) of Proposition 8.18, which is equivalent to injectivity of the map \( \psi \mid_{\text{Sym} \mathfrak{h}} \).

### 8.6.5. The Harish-Chandra Homomorphism

We now repeat some of the constructions in §8.6.4, but working in \( U = U_{\mathfrak{h}} \) rather than \( \text{Sym} \mathfrak{g} \). Again, consider the triangular decomposition \( \mathfrak{g} = \mathfrak{n}_- \oplus \mathfrak{h} \oplus \mathfrak{n}_+ \) and fix an ordered basis of \( \mathfrak{g} \) by starting with a basis of \( \mathfrak{n}_- \) in some order, then adding a basis of \( \mathfrak{h} \), and finally a basis of \( \mathfrak{n}_+ \). By the Poincaré-Birkhoff-Witt Theorem, the standard monomials in this basis form a basis of \( U \); the monomials involving only factors from \( \mathfrak{h} \) are a basis for the subalgebra \( U_{\mathfrak{h}} \subseteq U \), while the other monomials span the subspace \( \mathfrak{n}_- U + U \mathfrak{n}_+ \) of \( U \). Therefore, \( U = (\mathfrak{n}_- U + U \mathfrak{n}_+) \oplus U_{\mathfrak{h}} \) and \( U_{\mathfrak{h}} = \text{Sym} \mathfrak{h} \), because \( \mathfrak{h} \) is commutative. Letting \( \psi \) denote the projection of \( U \) onto \( U_{\mathfrak{h}} \) along the subspace \( \mathfrak{n}_- U + U \mathfrak{n}_+ \), we obtain a \( \mathbb{k} \)-linear epimorphism, analogous to \( \psi \) in (8.25) but not an algebra map:

\[
\begin{align*}
\varphi: \quad U &= (\mathfrak{n}_- U + U \mathfrak{n}_+) \oplus U_{\mathfrak{h}} \quad \xrightarrow{\text{proj.}} \quad U_{\mathfrak{h}} = \text{Sym} \mathfrak{h}.
\end{align*}
\]

Interestingly, the restriction of \( \varphi \) to \( \mathcal{Z} = \mathcal{Z}U \) actually is an algebra map. In fact, part (a) of the lemma below shows that this even holds for the restriction of \( \varphi \) to the subalgebra \( U^{\mathfrak{b}} \subseteq U \) consisting of the invariants for the adjoint action of \( \mathfrak{h} \) on \( U \); the algebra map in (a) is called the Harish-Chandra homomorphism (for the given base \( \Delta \)). The algebra \( U^{\mathfrak{b}} \) is identical to the centralizer of \( \mathfrak{h} \) in \( U \) by (5.37); so \( U^{\mathfrak{b}} \) certainly contains \( \mathcal{Z} = U^{\mathfrak{b}} \) as well as \( U_{\mathfrak{h}} \). Part (b) of the lemma explains the connection to central characters (8.15). Using the standard identification \( \mathfrak{h} = \mathfrak{b}^{*} \), we view \( \text{Sym} \mathfrak{h} \) as the algebra \( O(\mathfrak{b}^{*}) \) of polynomial functions on \( \mathfrak{b}^{*} \) and write \( \varphi(u)(\lambda) \) for the evaluation of \( \varphi(u) \in \text{Sym} \mathfrak{h} \) at \( \lambda \in \mathfrak{b}^{*} \).

**Lemma 8.22.**

(a) \( \varphi\mid_{U^{\mathfrak{b}}} \) is an epimorphism of algebras \( U^{\mathfrak{b}} \to \text{Sym} \mathfrak{h} \).

(b) If \( \lambda \in \mathfrak{h}^{*} \) and \( z \in \mathcal{Z} \), then \( \chi_{\lambda}(z) = \varphi(z)(\lambda) \).

**Proof.** (a) Since \( U^{\mathfrak{b}} \) contains \( U_{\mathfrak{h}} \), the restricted projection \( \varphi\mid_{U^{\mathfrak{b}}} \) is trivially surjective and \( U^{\mathfrak{b}} = I \oplus U_{\mathfrak{h}} \), where we have put \( I := U^{\mathfrak{b}} \cap \text{Ker} \varphi = U^{\mathfrak{b}} \cap (\mathfrak{n}_- U + U \mathfrak{n}_+) \). The issue is to show that \( I \) is an ideal of the algebra \( U^{\mathfrak{b}} \). To this end, we will prove the following equalities:

\[
I = U^{\mathfrak{b}} \cap \mathfrak{n}_- U = U^{\mathfrak{b}} \cap U \mathfrak{n}_+ .
\]

The first equality will show that \( I \) is a right ideal of \( U^{\mathfrak{b}} \) and the second will show that \( I \) is a left ideal. To prove the above equalities, write \( U = U_{\text{ad}} \) as \( U = \bigoplus_{\lambda \in \mathfrak{h}^{*}} U_{\lambda} \) with \( U_{0} = U^{\mathfrak{b}} \); this is possible, because \( U_{\text{ad}} \in \text{Rep} \mathfrak{g} \) is locally finite (§8.2.2). Since both \( \mathfrak{n}_- U \) and \( U \mathfrak{n}_+ \) are stable under the adjoint action of \( \mathfrak{h} \), it follows that \( I = (U^{\mathfrak{b}} \cap \mathfrak{n}_- U) + (U^{\mathfrak{b}} \cap U \mathfrak{n}_+) \) and so it suffices to show that \( U^{\mathfrak{b}} \cap \mathfrak{n}_- U = U^{\mathfrak{b}} \cap U \mathfrak{n}_+ \).
Fix a basis $h_1, \ldots, h_n$ of $\mathfrak{h}$, label the elements of $\Phi_+$ as $\alpha_1, \ldots, \alpha_r$, and choose generators for all root spaces, say $g_{\pm \alpha_i} = \pm x_{\pm \alpha_i}$. Then the standard monomials
\[ u = x_{-1}^{m_1} \cdots x_{-t}^{m_t} h_1^{k_1} \cdots h_n^{k_n} x_1^{m_1} \cdots x_t^{m_t} \quad (m_i, k_j \in \mathbb{Z}_+). \]
form a $\mathbb{k}$-basis of $U$ by the Poincaré-Birkhoff-Witt Theorem. Each such monomial $u$ is a weight vector of weight $|u|_+ - |u|_-$, where we have put $|u|_\pm = \sum_{i=1}^t m_{\pm i} \alpha_i$. The subspace $U_{n_\pm}$ is the $\mathbb{k}$-span of all monomials $u$ with at least one of $m_i \neq 0$ for $1 \leq i \leq t$ or, equivalently, $|u|_+ \neq 0$. Similarly, $n_- U$ is spanned by the monomials with $|u|_- \neq 0$. Since $u \in U^b$ if and only if $|u|_+ = |u|_-$, it follows that $U^b \cap n_- U = U^b \cap U_{n_+}$ as desired.

(b) Let $V = U \cdot v$ be a highest weight representation with highest weight vector $v \in V \lambda$. Continuing with the notation of (a) above, write a given $z \in \mathcal{Z}$ as $z = \varphi(z) + u$ with $u \in I$. Since $I \subseteq U_{n_\pm}$, we have $u \cdot v = 0$. Therefore,
\[ \chi_\lambda(z) \cdot v = z \cdot v = \varphi(z) \cdot v = \varphi(z)(\lambda) \cdot v, \]
where the last equality follows from $v$ having weight $\lambda$. This proves (b).

8.6.6. The Harish-Chandra Isomorphism

The Poincaré-Birkhoff-Witt isomorphism $\phi: \text{Sym} \mathfrak{g} \rightarrow \text{gr} U$ is an isomorphism of graded $\mathfrak{g}$-algebras by (5.39). Therefore, in conjunction with Theorem 8.21, $\phi$ yields an isomorphism of invariant algebras, $(\text{Sym} \mathfrak{h})^W \cong (\text{Sym} \mathfrak{g})^0 \rightarrow (\text{gr} U)^0$. In this subsection, we will show that, in fact, $(\text{Sym} \mathfrak{h})^W \cong U^0 = \mathcal{Z}$ as $\mathbb{k}$-algebras.

For a given $\lambda \in \mathfrak{h}^*$, consider the linear map $\mathfrak{h} \rightarrow \text{Sym} \mathfrak{h}$, $h \mapsto h - \langle \lambda, h \rangle$. By the universal property of the symmetric algebra (1.8), this map extends uniquely to an algebra endomorphism of $\text{Sym} \mathfrak{h}$, which we will denote by $\tau_\lambda$. Since $\tau_{-\lambda}$ is inverse to $\tau_\lambda$, we do in fact obtain an automorphism in $\text{Alg}_\mathbb{k}$.

$\tau_\lambda: \quad \text{Sym} \mathfrak{h} \longrightarrow \text{Sym} \mathfrak{h}$

\[ \begin{array}{c}
\psi \\
\downarrow \\
\psi \\
\hline
h \\
\mapsto h - \langle \lambda, h \rangle \\
(h \in \mathfrak{h})
\end{array} \]

Viewing $\text{Sym} \mathfrak{h}$ as the algebra of polynomial functions on $\mathfrak{h}^*$, the automorphism $\tau_\lambda$ corresponds to shifting the origin of $\mathfrak{h}^*$ by $\lambda$: $\tau_\lambda(p)(\mu) = p(\mu - \lambda)$ for $p \in \text{Sym} \mathfrak{h} = O(\mathfrak{h}^*)$ and $\mu \in \mathfrak{h}^*$. Below, we will use this with $\lambda = \rho$ (7.12).

The following theorem, the denouement of the developments in this section, is due to Harish-Chandra [96]. One can show (Exercise 8.6.4) that the isomorphism in the theorem is independent of the choice of the base $\Delta$, despite the fact that the ingredients in its construction, $\tau_\rho$, and the Harish-Chandra homomorphism $\varphi$, each depend on $\Delta$.

$^3$Automorphisms of this form are called \textit{winding automorphisms}; see also Exercises 3.3.11 and 10.1.6.
Theorem 8.23. \( (\tau_p \circ \varphi)|_{\mathcal{J}} \) is an isomorphism of \( \mathbb{F}\)-algebras, \( \mathcal{J} \rightarrow (\text{Sym } \mathfrak{h})^W \).

Proof. We already know that \( (\tau_p \circ \varphi)|_{\mathcal{J}} : U^h \rightarrow \text{Sym } \mathfrak{h} \) is a surjective algebra map (Lemma 8.22). We need to show that the restriction of this map to \( \mathcal{J} = U^h \subseteq U^h \) is a bijection,

\[
\phi : (\tau_p \circ \varphi)|_{\mathcal{J}} : \mathcal{J} \rightarrow A := (\text{Sym } \mathfrak{h})^W.
\]

Step 1: \( f(\mathcal{J}) \subseteq A \). To prove this, let \( \lambda \in \Lambda_+ \) and \( w \in W \). By Proposition 8.10, \( \chi_{W \cdot (\lambda - \rho)}(\lambda) = \chi_{\lambda - \rho} \) or, equivalently, \( \varphi(z)(\lambda) = \varphi(z)(\lambda) \) for all \( z \in \mathcal{J} \) by Lemma 8.22 and the definition of the shifted action of \( W \) (8.16). Since \( \lambda \in \Lambda_+ \) was arbitrary, this says that the polynomial functions \( f(z) \circ w \) and \( f(z) \) coincide on the subset \( \Lambda_+ \subseteq \mathfrak{h}^* \), and since \( \Lambda_+ \) is dense in \( \mathfrak{h}^* \) for the Zariski topology (Exercise 8.1.2), it follows that the two functions are identical: \( f(z) \circ w = f(z) \). Therefore, \( f(\mathcal{J}) \subseteq (\text{Sym } \mathfrak{h})^W = A \).

Step 2: filtrations. Consider the standard filtration \( (U_n) \) of \( U \) and the filtration of \( \text{Sym } \mathfrak{h} \) coming from the grading: \( \text{Sym}_n \mathfrak{h} = \bigoplus_{k \leq n} \text{Sym}^k \mathfrak{h} \). These induce filtrations of \( \mathcal{J} \) and \( A \), respectively:

\[
\mathcal{J}_n := \mathcal{J} \cap U_n \quad \text{and} \quad A_n := A \cap \text{Sym}_n \mathfrak{h}.
\]

Recall that the symmetrization \( \omega : \text{Sym } \mathfrak{g} \rightarrow U \) is a \( \text{Rep } \mathfrak{g} \)-isomorphism for the adjoint \( g \)-actions and it maps \( \text{Sym}_n \mathfrak{g} = \bigoplus_{k \leq n} \text{Sym}^k \mathfrak{g} \) onto \( U_n \) (Proposition 5.27). Hence, we obtain an isomorphism of \( \mathfrak{g} \)-invariants,

\[
\omega' : = \omega^{-1}|_{\mathcal{J}^1} : U^0 \rightarrow B := (\text{Sym } \mathfrak{g})^0,
\]

satisfying \( \omega'(\mathcal{J}_n) = B_n := B \cap \text{Sym}_n \mathfrak{g} \). In addition, Theorem 8.21 furnishes the isomorphism \( \psi' : \psi|_{\mathcal{J}} : B \rightarrow A \), which comes from the graded algebra epimorphism \( \psi : \text{Sym } \mathfrak{g} \rightarrow \text{Sym } \mathfrak{h} \) (8.25); so \( \psi'(B_n) = A_n \). Hence, we obtain the \( \mathbb{F}\)-linear isomorphism \( f' : \mathcal{J} \rightarrow A \) satisfying \( f'(\mathcal{J}_n) = A_n \).

Claim. \( f'(z) \equiv f(z) \mod A_{n-1} \) for all \( z \in \mathcal{J}_n \).

We postpone the proof and proceed to derive bijectivity of \( f \) from the claim. Let \( 0 \neq z \in \mathcal{J} \) and choose \( n \) minimal with \( z \in \mathcal{J}_n \). Then \( f'(z) \notin A_{n-1} \) and so \( f(z) \notin A_{n-1} \) by the claim, proving that \( f \) is injective. The claim also yields the monomials \( u_{m,k} = x_{-1}^m \cdots x_{-l}^m \cdot k_1 \cdots k_n \cdot h_1^j \cdots h_n x_1^m \cdots x_t^m \). Then a \( \mathbb{F}\)-basis of \( U \) is given by the monomials

\[
u_{m,k} = x_{-1}^m \cdots x_{-l} \cdot h_1^k \cdots h_n x_1^m \cdots x_t^m \quad (m_i, k_j \in \mathbb{Z}_+).\]
Putting $|\mathbf{m}| = \sum_i m_i$ and $|\mathbf{k}| = \sum_j k_j$, we may write a given $z \in Z_n$ uniquely as

$$z = \sum_{m,k \in \mathbb{K}} \zeta_{m,k} u_{m,k} \quad \text{with } \zeta_{m,k} \in \mathbb{K}.$$ 

By definition, the Harish-Chandra homomorphism $\varphi$ annihilates all monomials $u_{m,k}$ with $|\mathbf{m}| \neq 0$ and sends all $u_{0,k} \in \mathfrak{U} = \text{Sym} \mathfrak{h}$ to themselves. Since the automorphism $\tau_\rho$ of $\text{Sym} \mathfrak{h}$ clearly satisfies $\tau_\rho(u) \equiv u \mod \text{Sym}_{n-1} \mathfrak{h}$ for $u \in \text{Sym}_n \mathfrak{h}$, we obtain

$$f(z) \equiv \sum_{k : |\mathbf{k}| = n} \zeta_{0,k} u_{0,k} \mod \text{Sym}_{n-1} \mathfrak{h}.$$ 

For $f' = \psi \circ \omega^{-1}|_{\mathcal{X}}$, we know by (5.41) that $\omega^{-1}(u_{m,k}) \equiv s_{m,k} \mod \text{Sym}_{n-1} \mathfrak{g}$, where $s_{m,k}$ denotes the monomial $u_{m,k}$ but calculated in $\text{Sym} \mathfrak{g}$. Therefore,

$$\omega^{-1}(z) \equiv \sum_{m,k : |\mathbf{m}| + |\mathbf{k}| \leq n} \zeta_{m,k} s_{m,k} \mod \text{Sym}_{n-1} \mathfrak{g}.$$ 

Finally, since $\psi(s_{m,k}) = 0$ if $|\mathbf{m}| \neq 0$, and $\psi(s_{0,k}) = s_{0,k} = u_{0,k} \in \mathfrak{U} = \text{Sym} \mathfrak{h}$, we obtain

$$f'(z) \equiv \sum_{k : |\mathbf{k}| = n} \zeta_{0,k} u_{0,k} \mod \text{Sym}_{n-1} \mathfrak{h}.$$ 

Comparison shows that $f'(z) \equiv f(z) \mod \text{Sym}_{n-1} \mathfrak{h}$, and since we also know that $f'(z) - f(z) \in \Lambda$, the claim follows. This completes the proof of the theorem. \qed

**Corollary 8.24.** Let $\lambda, \mu \in \mathfrak{h}^*$. Then $\chi_\lambda = \chi_\mu$ if and only if $\mathfrak{W} \cdot \mu = \mathfrak{W} \cdot \lambda$.

**Proof.** As we have already observed in Step 1 of the above the proof, Lemma 8.22 gives the formula $\chi_{\lambda - \rho}(z) = f(z)(\lambda)$ for any $\lambda \in \mathfrak{h}^*$ and $z \in \mathcal{X}$, where $f$ is the Harish-Chandra isomorphism of Theorem 8.23. Thus, in view of the definition of the shifted $\cdot$ action of $\mathfrak{W}$, the corollary asserts that

$$f(z)(\lambda) = f(z)(\mu) \quad \text{for all } z \in \mathcal{X} \quad \iff \quad \mathfrak{W} \mu = \mathfrak{W} \lambda.$$

The direction $\Leftarrow$ states the polynomial functions $f(z) \in \text{Sym} \mathfrak{h} = \mathcal{O}(\mathfrak{h}^*)$ are constant on all $\mathfrak{W}$-orbits in $\mathfrak{h}^*$, which in turns says that $f(\mathcal{X}) \subseteq \mathcal{O}(\mathfrak{h}^*)^\mathfrak{W}$. We know this to be true; in fact, equality holds. Thus, or the reverse direction, it suffices to show that if $\lambda, \mu$ belong to different $\mathfrak{W}$-orbits in $\mathfrak{h}^*$, then there is some polynomial function $p \in \mathcal{O}(\mathfrak{h}^*)^\mathfrak{W}$ with $p(\lambda) \neq p(\mu)$. To see this, pick some polynomial function $p_0 \in \mathcal{O}(\mathfrak{h}^*)^\mathfrak{W}$ such that $p_0(\lambda) \neq 0$ but $p_0$ vanishes on $(\mathfrak{W} \cdot \{\lambda\}) \setminus \{\lambda\}$—this is a finite, and hence Zariski closed, subset of $\mathfrak{h}^*$; so $p_0$ certainly exists. Then $p = \sum_{w \in \mathfrak{W}} w \cdot p_0$ satisfies $p(\lambda) \neq 0$ but $p(\mu) = 0$ and $p$ belongs to $\mathcal{O}(\mathfrak{h}^*)^\mathfrak{W}$ as desired. This proves the corollary. \qed

Corollary 8.24 was already used in the proof Proposition 8.11; so this proposition is now fully established. Observe that only the fact that the Harish-Chandra
isomorphism has image \((\text{Sym } \mathfrak{h})^W = O(\mathfrak{h}^*)^W\) entered the proof of the corollary, whereas injectivity of the Harish-Chandra isomorphism and of the map \(\psi|_{(\text{Sym } \mathfrak{g})^X}\) were irrelevant. Thus, Corollary 8.24 does not depend on the Conjugacy Theorem.

### 8.6.7. Outlook: The Shephard-Todd-Chevalley Theorem

The Harish-Chandra isomorphism (Theorem 8.23) attains additional significance in light of the Shephard-Todd-Chevalley Theorem [184], [42], a classical result of invariant theory. To state this result, let \(V \in \text{Vect}_k\) be finite dimensional. A pseudo-reflection of \(V\) is an automorphism \(s \in \text{GL}(V)\) such that \(\text{Id}_V - s \in \text{End}_k(V)\) has rank 1; pseudo-reflections of order 2 are called reflections. A subgroup \(G \leq \text{GL}(V)\) is called a pseudo-reflection group if \(G\) is generated by pseudo-reflections. Any subgroup \(G \leq \text{GL}(V)\) acts by \(k\)-algebra automorphisms on the symmetric algebra \(\text{Sym } V\), and hence we may consider the invariant subalgebra \((\text{Sym } V)^G \subseteq \text{Sym } V\).

**Shephard-Todd-Chevalley Theorem.** Let \(V \in \text{Vect}_k\) with \(\dim_k V = n < \infty\) and let \(G\) be a finite subgroup of \(\text{GL}(V)\). Then \(G\) is a pseudo-reflection group if and only if the invariant algebra \((\text{Sym } V)^G\) is generated by (necessarily \(n\)) algebraically independent elements.

The theorem does not require \(k\) to be algebraically closed or \(\text{char } k = 0\), but it is essential that \(\text{char } k \nmid |G|\). For a proof, the reader may consult [23, chap. V §5, Théorème 4] or [35, Theorem 6.4.12], for example. The monograph [55] offers a discussion of more recent developments concerning the Shephard-Todd-Chevalley Theorem.

In the setting of the present chapter, all this applies with \(V = \mathfrak{h}\) and \(G = \mathcal{W}\), because \(\mathcal{W}\) is generated by the reflections \(s_\alpha\) \((\alpha \in \Phi)\). Thus, the Shephard-Todd-Chevalley Theorem yields that \((\text{Sym } \mathfrak{h})^\mathcal{W}\) is a polynomial algebra in \(n = \text{rank } \mathfrak{g}\) variables over \(k\), and hence so is \(\mathcal{Z}\) by Theorem 8.23. Thus:

**Corollary 8.25.** \(\mathcal{Z}\) is a polynomial algebra over \(k\) in \(n\) variables.

**Example 8.26** (\(\mathcal{Z}\) for \(\mathfrak{sl}_{n+1}\)). As in §6.4.2 and elsewhere, we work with the Cartan subalgebra \(\mathfrak{h} = \mathfrak{d}_{n+1} \cap \mathfrak{sl}_{n+1}\). By (8.5) and (6.30) the simple reflections \(s_\alpha \in \mathcal{W}\) are given by \(s_\alpha = s|_{\mathfrak{h}}\), where \(s_i \in \text{GL}(\mathfrak{d}_{n+1})\) interchanges the two matrices \(e_{i,i}\) and \(e_{i+1,i+1}\) and leaves all other \(e_{j,j}\) fixed; so \(\mathcal{W} \cong S_{n+1}\) with \(s_\alpha\) corresponding to the transposition \((i,i+1)\). Let us write \(x_i\) for the canonical image of \(e_{i,i}\) in \(\text{Sym } \mathfrak{d}_{n+1}\). Then \(\text{Sym } \mathfrak{d}_{n+1} = k[x_1, \ldots, x_{n+1}]\) and Weyl group \(\mathcal{W} \cong S_{n+1}\) operates by permuting the variables \(x_i\) in the usual fashion. Moreover, since \(\mathfrak{d}_{n+1} = \mathfrak{h} \oplus \mathfrak{k}\) with \(1 = e_{1,1} + \cdots + e_{n+1,n+1}\), the identity matrix, we may write \(\text{Sym } \mathfrak{d}_{n+1} = (\text{Sym } \mathfrak{h})[t]\), with \(t = x_1 + \cdots + x_{n+1}\). Since \(\text{Sym } \mathfrak{h}\) is a \(\mathcal{W}\)-stable subalgebra of \(\text{Sym } \mathfrak{d}_{n+1} = k[x_1, \ldots, x_{n+1}]\) and \(t\) is \(\mathcal{W}\)-invariant, the Fundamental Theorem of \(S_{n+1}\)-invariants gives

\[k[x_1, \ldots, x_{n+1}]^{S_{n+1}} = (\text{Sym } \mathfrak{h})^\mathcal{W}[t] = k[e_1, \ldots, e_{n+1}]\]
where the $e_i = e_i(x_1, \ldots, x_{n+1})$ are the elementary symmetric polynomials. Sending $t = e_1 \mapsto 0$ and invoking Theorem 8.23, we obtain the isomorphism

$$\mathbb{k}[e_2, \ldots, e_{n+1}] \cong (\operatorname{Sym} b)^W \cong \mathcal{Z},$$

which exhibits $\mathcal{Z}$ as a polynomial algebra in $n$ variables.

### 8.6.8. Outlook: Primitive Ideals

We conclude this long section by stating a ring-theoretic result that generalizes our findings for $\mathfrak{sl}_2$ in §§5.7.9, 5.7.10. We then derive a version of the Dixmier-Mœglin equivalence (§5.6.6) that is tailored for semisimple Lie algebras from this result and comment on the relationship between the topological spaces $\text{Prim} U$ and $\text{MaxSpec} \mathcal{Z}$.

Recall that an ideal $I$ of $U = U_\mathfrak{g}$ is said to be completely prime if $U/I$ is a domain.

**Theorem 8.27.** (a) Let $I$ be an ideal of $U$. Then $\mathcal{Z}(U/I) \cong \mathcal{Z}/I \cap \mathcal{Z}$ via the canonical map $U \twoheadrightarrow U/I$. If $I \neq 0$, then $I \cap \mathcal{Z} \neq 0$. Moreover, if $a$ is an ideal of $\mathcal{Z}$ such that $a \supseteq I \cap \mathcal{Z}$, then $(I + aU) \cap \mathcal{Z} = a$.

(b) Let $\mathfrak{m} \in \text{MaxSpec} \mathcal{Z}$. Then the ideal $\mathfrak{m}U$ of $U$ is completely prime and $\text{Spec}(U/\mathfrak{m}U)$ is finite.

**Remarks on the proof.** Part (b) is beyond the scope of this book. For the fact that $U/\mathfrak{m}U$ is a domain, see Dixmier [60, 8.4.4] and for the finiteness statement, see Jantzen [112, 7.3]. However, (a) is relatively simple. Both $U$ and $A := U/I$, viewed as representations of $\mathfrak{g}$ via the adjoint action, are locally finite and hence completely reducible by Weyl’s Theorem. Therefore, the canonical map $U \twoheadrightarrow A$ restricts to a surjection of $\mathfrak{g}$-invariants $U^{\mathfrak{ad}}_\mathfrak{ad} \twoheadrightarrow A^{\mathfrak{ad}}_\mathfrak{ad}$. Since $\mathcal{Z} = U^{\mathfrak{ad}}_\mathfrak{ad}$ and $\mathcal{Z} A = A^{\mathfrak{ad}}_\mathfrak{ad}$ by (5.38), this says that the canonical map $\mathcal{Z} \twoheadrightarrow \mathcal{Z} A$ is surjective. In order to prove that $I \cap \mathcal{Z} \neq 0$ for $I \neq 0$, it suffices to show that $I V \neq 0$ for some $V \in \text{Irr}_{\text{fin}} \mathfrak{g}$; then the argument in the proof of Proposition 5.47 applies verbatim to yield the desired nonzero element of $I \cap \mathcal{Z}$. To find such a representation $V$, see Exercise 8.4.1. The last assertion in (a) states that if $a$ is an ideal of $\mathcal{Z} A$, then the ideal $a A$ of $A$ satisfies $a A \cap \mathcal{Z} A = a$. To prove this, note that $A^{\mathfrak{ad}} = \mathcal{Z} A \oplus A'$, where $A'$ is the sum of the homogeneous components other than $\mathcal{Z} A$, and $(\mathcal{Z} A) A' \subseteq A'$. Therefore, $a A \subseteq a \oplus A'$ and hence $a \subseteq a A \cap \mathcal{Z} A \subseteq a$. 

**Corollary 8.28.** The following are equivalent, for any $P \in \text{Spec} U$:

1. $P$ is primitive;
2. $\mathcal{Z}(U/P) \cong \mathbb{k}$;
3. $\mathcal{Z}/P \cap \mathcal{Z} \cong \mathbb{k}$;
4. $P$ is locally closed.
Proof. The implications (iv) ⇒ (i) ⇒ (ii) follow from the Nullstellensatz for enveloping algebras (§§5.6.1, 5.6.6); for (ii), observe that if \( P = \ker V \) for \( V \in \operatorname{Irr} g \), then \( Z(U/P) \) embeds into \( D(V) = \mathbb{k} \). The implication (ii) ⇒ (iii) is trivial, and (iii) ⇒ (iv) is a consequence of Theorem 8.27: \( P \) corresponds to one of the finitely many primes of the algebra \( U/(P \cap Z)U \), and hence \( P \) is evidently locally closed. □

Finally, let us consider the topological spaces \( X = \operatorname{Prim} U \) and \( Z = \operatorname{MaxSpec} \mathcal{Z} \), each equipped with the Jacobson-Zariski topology (§1.3.4). For any finite-dimensional Lie algebra, we have the continuous map (5.45):

\[
\pi : X = \operatorname{Prim} U \longrightarrow Z = \operatorname{MaxSpec} \mathcal{Z}
\]

More can be said for semisimple Lie algebras. First, since \( Z \) is a polynomial algebra over \( \mathbb{k} \) in \( n \) variables (Corollary 8.25), the topological space \( Z \) is just affine \( n \)-space, \( \mathbb{k}^n \). Moreover:

**Corollary 8.29.** The map \( \pi \) is a closed surjection with finite fibers, each of which has a unique minimal element.

**Proof.** The fact that \( \pi \) is surjective with finite fibers follows from Theorem 8.27 and Corollary 8.28: \( mU \) is a prime ideal of \( U \) for any \( m \in Z \), even completely prime, with only finitely many primes above it; these are all primitive and make up the fiber \( \pi^{-1}(m) \). Clearly, \( mU \) is the unique minimal element of this fiber.

To show that \( \pi \) is closed, let \( C \subseteq X \) be a closed subset; so \( C = \mathcal{V}(I) \) with \( I = \mathcal{I}(C) = \bigcap_{P \in C} P \) (Exercise 1.3.1). Put \( D = \pi C \) and

\[
a = \mathcal{I}(D) = \bigcap_{P \in C} (P \cap \mathcal{Z}) = I \cap \mathcal{Z}.
\]

Our goal is to show that \( D = \mathcal{V}(a) \), with \( C \) being clear. For the reverse inclusion, let \( m \in \mathcal{V}(a) \). Then \( I + mU \) is a proper ideal of \( U \) by Theorem 8.27, and hence it is contained in some maximal ideal, say \( P \). Thus, \( P \) is primitive, a member of \( \mathcal{V}(I) \), and \( \pi P = P \cap \mathcal{Z} = m \). Therefore, \( m \in D \). □

Being the image of the continuous closed surjection \( \pi \), the topological space \( Z \) carries the quotient topology for \( \pi \). Even more can be said: the map \( \pi \) is also open, each fiber of \( \pi \) also has a unique maximal element, and “most” fibers consist of just one element. We had seen this earlier for \( g = \mathfrak{sl}_2 \) (Theorem 5.48). In general, the “exceptional” set \( Z^+ = \{ m \in Z \mid \#\pi^{-1}(m) \geq 2 \} \) is a locally finite union of algebraic hypersurfaces in \( Z = \mathbb{k}^n \). See the survey article [20] or the monograph [112] for more on this topic. For an illustration, Figure 8.2 exhibits the (\( \mathbb{k} \)-points of the) exceptional curves in the plane \( Z = \mathbb{k}^2 \) that make up \( Z^+ \) for \( g = \mathfrak{sl}_3 \). The picture is due to Borho [19], [20]. With \( x := e_2 \) and \( y := e_3 \) as in Example 8.26, the curves are given by \( x = -\frac{1}{3}(t^2 + nt + n^2) \) and \( y = \frac{1}{2}((2t + n)(t - n)(t + 2n)) \), where
$t \in \mathbb{k}$ is a parameter and $n \in \mathbb{N}$. The intersections of the curves mark the points $m \in \mathbb{Z}$ with $\# \pi^{-1}(m) \geq 3$.

**Figure 8.2.** The exceptional curves for $\mathfrak{sl}_3$ in $\mathbb{Z} = \mathbb{k}^2$

### Exercises for Section 8.6

**8.6.1** (Regular elements of $\mathfrak{sl}_n$). Show that $x \in \mathfrak{sl}_n$ is regular, as in the proof of Proposition 8.18, if and only if all $n$ eigenvalues of $x$ have multiplicity 1.

**8.6.2** (Exponentials). Let $A \in \text{Alg}_{\mathbb{k}}$. For any nilpotent element $a \in A$, define $e^a := \sum_{n \geq 0} \frac{a^n}{n!} \in A$. Show:

(a) If $a, b \in A$ are commuting nilpotent elements, then $e^a e^b = e^{a+b}$. In particular, $e^a$ has inverse $e^{-a}$, and hence $e^a \in A^\times$.

(b) Now let $A = \text{End}_k(\mathfrak{g})$, where $\mathfrak{g}$ is a general $\mathbb{k}$-algebra (§5.1.5). If $d \in A$ is a nilpotent derivation of $\mathfrak{g}$, then $e^d \in \text{Aut} \mathfrak{g} \subseteq A$, that is, $e^d(xy) = e^d(x)e^d(y)$ for all $x, y \in \mathfrak{g}$. (Use the Leibniz formula in Exercise 5.1.5.)
8.6.3 (Elementary automorphisms). In this exercise, \( \mathfrak{g} \) can be an arbitrary Lie algebra. The group of elementary automorphisms is defined as in §8.6.2:

\[
\text{Aut}_e \mathfrak{g} = \langle e^{ad_x} \mid x \in \mathfrak{g} \text{ ad-nilpotent} \rangle \leq \text{Aut} \mathfrak{g}
\]

Show that \( \text{Aut}_e \mathfrak{g} \) is a normal subgroup of \( \text{Aut} \mathfrak{g} \).

8.6.4 (Harish-Chandra isomorphism). Show that the map \( (\tau_\rho \circ \varphi)\big|_\mathfrak{g} \) is independent of the choice of the base \( \Delta \) of \( \Phi \). (Use Theorem 7.9.)

8.6.5 (The centralizer of \( \mathfrak{h} \) in \( U \)). Write \( U_{ad} = \bigoplus_{\lambda \in \mathcal{L}} U_\lambda \) with \( U_0 = U^b \), the centralizer of \( \mathfrak{h} \) in \( U \), as in the proof of Lemma 8.22 and Exercise 8.2.2. Show:

(a) Label the elements of \( \Phi \) as \( \alpha_1, \ldots, \alpha_t \), and choose generators for all root spaces, say \( g_{\alpha_i} = \mathbb{k}x_i \). Show that the algebra \( U^b \) is generated by \( \mathfrak{h} \) together with the standard monomials \( u_m = x_1^{m_1} \cdots x_t^{m_t} \) (\( m_i \in \mathbb{Z}_+ \)) such that \( m = (m_1, \ldots, m_t) \) is minimal member of the set \( \{ m \in \mathbb{Z}_+^t : \sum_{i=1}^t m_i \alpha_i = 0 \} \) for the partial order \( \leq \) on \( \mathbb{Z}_+^t \) that is given by \( m \leq k \) iff \( m_i \leq k_i \) for all \( i \).

(b) \( U^b \) is a (left and right) noetherian affine subalgebra of \( U \). (Show that every nonempty subset of \( \mathbb{Z}_+^t \) has only finitely many minimal elements for \( \leq \). Also, with \( U_n^b = U^b \cap U_n \) denoting the filtration of \( U^b \) that is induced by the standard filtration of \( U \), show that \( \text{gr} U^b = \bigoplus_{n \geq 0} U_n^b/U^b_{n-1} \) is affine and use Lemma 5.23.)

(c) Each \( U_\lambda \) is a finitely generated (left and right) \( U^b \)-module via multiplication.

8.6.6 (An example for the Shephard-Todd-Chevalley Theorem). Let \( V = \mathbb{k}^2 \) and let \( G \leq \text{GL}(V) \) be the subgroup that is generated by the matrices \( a = \left( \begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix} \right) \) and \( b = \frac{1}{\sqrt{2}} \left( \begin{smallmatrix} 1 & 1 \\ 1 & -1 \end{smallmatrix} \right) \). Using the identification \( \text{Sym} V = \mathbb{k}[x, y] \) via the standard basis of \( V \), prove:

(a) Both \( a \) and \( b \) are reflections and \( G \cong \mathcal{D}_8 \), the dihedral group of order 16.

(b) \( H((\text{Sym} V)^G, t) = \frac{1}{(1-t^2)(1-t^4)} \). (Use Molien’s Theorem.)

(c) \( s = x^2 + y^2 \) and \( t = x^2 y^2(x^2 - y^2)^2 \) are algebraically independent invariants. Conclude from (b) that \( (\text{Sym} V)^G = \mathbb{k}[s, t] \).

8.7. Weyl’s Character Formula

In this section, we will give a formula for the characters of finite-dimensional irreducible representations of \( \mathfrak{g} \). Recall that any such representation is isomorphic to a highest weight representation \( V(\lambda) \) for a unique \( \lambda \in \Lambda_\ast \) (8.18). Weyl’s Character Formula expresses the formal character \( \text{ch}_\lambda = \text{ch} V(\lambda) \in (\mathbb{Z}\Lambda)^W \) as a quotient of two anti-invariants (§7.4.5) of the form

\[
a_\mu = \sum_{w \in W} \sgn(w) x^{w \mu} \in (\mathbb{Z}\Lambda)^\pm \quad (\mu \in \Lambda).
\]
We already know that $a_\rho$ divides $a_{\lambda+\rho}$ in $\mathbb{Z}\Lambda$ and that the “denominator” $a_\rho$ has the following form (Proposition 7.19):

(8.29) \[ a_\rho = \prod_{\alpha \in \Phi_+} (x^{\frac{2}{\alpha}} - x^{\frac{-2}{\alpha}}) = x^\rho \prod_{\alpha \in \Phi_+} (1 - x^{-\alpha}) = x^{-\rho} \prod_{\alpha \in \Phi_+} (x^\alpha - 1). \]

The proof that the quotient $\frac{a_{\lambda+\rho}}{a_\rho}$ evaluates to $\text{ch}_\lambda$ will be given in §8.7.3 after first discussing an application and some examples of Weyl’s Character Formula.

8.7.1. Dimensions

As an application of Weyl’s Character Formula, we now give a formula for the dimensions of the finite-dimensional irreducible representations of $g$. Observe that the dimension of any $V \in \text{Rep}_\text{fin} g$ is identical to the sum of the coefficients of the formal character $\text{ch}_V$, because $V$ is the direct sum of its weight spaces (Proposition 8.2). Thus, $\dim_k V = \varepsilon(\text{ch}_V)$ where $\varepsilon$ is the counit of the group ring $\mathbb{Z}\Lambda$ ($\S3.3.2$):

\[
\varepsilon: \mathbb{Z}\Lambda \longrightarrow \mathbb{Z}\\
\sum_{\mu \in \Lambda} z_\mu x^\mu \longmapsto \sum_{\mu \in \Lambda} z_\mu
\]

This map is a ring homomorphism that is invariant under the action of $W$ on $\mathbb{Z}\Lambda$, because $\varepsilon(w.x^\mu) = \varepsilon(x^{\mu}) = 1 = \varepsilon(x^\mu)$ for all $\mu \in \Lambda$.

Let $(\ldots, \ldots): \mathfrak{h}^* \times \mathfrak{h}^* \rightarrow \mathbb{K}$ be the bilinear form (6.19) coming from the Killing form on $\mathfrak{h}$. Recall that $(\ldots, \ldots)$ is symmetric and $W$-invariant (Exercise 7.1.1) and that $\langle \lambda, \alpha \rangle = \langle \lambda, h_\alpha \rangle = \frac{\gamma(a, \alpha)}{(\alpha, \alpha)}$ for $\lambda \in \mathfrak{h}^*$ and $\alpha \in \Phi$ ($\S6.3.4$).

**Theorem 8.30.**

\[
\dim_k V(\lambda) = \prod_{\alpha \in \Phi_+} \frac{\langle \lambda + \rho, \alpha \rangle}{\langle \rho, \alpha \rangle} = \prod_{\alpha \in \Phi_+} \frac{(\lambda + \rho, \alpha)}{(\rho, \alpha)} \quad (\lambda \in \Lambda_+).
\]

**Proof.** Note that $\varepsilon(a_\mu) = 0$ for all $\mu \in \Lambda$, because $\varepsilon(a_\mu) = \varepsilon(w.a_\mu) = \text{sgn}(w)\varepsilon(a_\mu)$ for all $w \in W$. Thus, while it is tempting to derive the theorem from Weyl’s Character Formula by simply applying $\varepsilon$, this would merely give $0 \cdot \dim_k V(\lambda) = 0$. So more care is required. Note also that the two expressions for $\dim_k V(\lambda)$ in the theorem agree term by term. We will focus on the version with $(\ldots, \ldots)$.

Recall that $(\ldots, \ldots)$ is $\mathbb{Q}$-valued on $\mathbb{Q}\Phi \times \mathbb{Q}\Phi$ (Proposition 6.12). Hence, for any $\mu \in \Lambda$, we may define

\[
\Psi_\mu(\lambda) := e^{(\mu, \lambda)} = \sum_{l \geq 0} \frac{(\mu, \lambda)^l}{l!} t^l \in \mathbb{Q}[t] \quad (\lambda \in \Lambda).
\]
Since \( \Psi_\mu(0) = 1 \) and \( \Psi_\mu(\lambda + \lambda') = e^{(\mu, \lambda) + (\mu, \lambda')} = \Psi_\mu(\lambda)\Psi_\mu(\lambda') \), the map \( \Psi_\mu \) extends uniquely to a ring homomorphism \( \Psi_\mu : \mathbb{Z}\Lambda \to \mathbb{Q}[t] \) and we may even replace \( \Lambda \) by \( \mathbb{Q}\Phi \) throughout. The composite of \( \Psi_\mu \) with the homomorphism \( \mathbb{Q}[t] \to \mathbb{Q} \) that is given by evaluation at \( t = 0 \) is the counit \( \varepsilon \) of \( \mathbb{Z}\Lambda \). Furthermore, since \((\cdot, \cdot)\) is symmetric and \( W \)-invariant and \( \text{sgn}(w) = \text{sgn}(w^{-1}) \) for all \( w \in W \), we obtain

\[
\Psi_\mu(a_\lambda) = \sum_{w \in W} \text{sgn}(w)e^{(\mu, w\lambda)} = \sum_{w \in W} \text{sgn}(w^{-1})e^{(\mu, w\lambda)} = \sum_{w \in W} \text{sgn}(w)e^{(\lambda, w\mu)} = \Psi_\lambda(a_\mu) \quad (\mu, \lambda \in \Lambda).
\]

Using this equality and the denominator formula (8.29) we compute

\[
\Psi_\rho(a_\mu) = \sum_{w \in W} \text{sgn}(w)e^{(\mu, w\lambda)} = \sum_{w \in W} \text{sgn}(w^{-1})e^{(\mu, w\lambda)} = \sum_{w \in W} \text{sgn}(w)e^{(\lambda, w\mu)} = \Psi_\lambda(a_\mu) \quad (\mu, \lambda \in \Lambda).
\]

Dividing both sides of this equality by \( t^{\dim \mathcal{V}} \) and then setting \( t = 0 \) gives the equality

\[
\left( \prod_{\alpha \in \Phi_+} (\rho, \alpha) \right)^{\dim \mathcal{V}} V(\lambda) = \prod_{\alpha \in \Phi_+} (\lambda + \rho, \alpha).
\]

Finally, \( (\rho, \alpha) > 0 \) for all \( \alpha \in \Phi_+ \), since \( \langle \rho, \alpha \rangle = 1 \) for \( i = 1, \ldots, n \) (Lemma 7.8). The dimension formula follows. \( \square \)

### 8.7.2. Characters and Dimensions for \( \mathfrak{sl}_{n+1} \)

**The Case of \( \mathfrak{sl}_2 \)**

In order to relate our earlier findings for \( \mathfrak{sl}_2 \) (Section 5.7) to Weyl’s Character Formula and Theorem 8.30, we use the setup from §8.1.4: \( \mathfrak{h}^* \) and \( \mathfrak{z} \) are identified via \( \mu \leftrightarrow \langle \mu, h \rangle \); the Weyl group \( W \equiv S_2 \) operates on \( \mathfrak{h}^* \) by multiplication with \( \pm 1 \); the set of roots is \( \Phi = \{ \pm 2 \} \); and the weight lattice is \( \Lambda = \mathbb{Z} \). Working with the base \( \Delta = \{ 2 \} \), we have \( \Phi_+ = \{ 2 \} \) and \( \Lambda_+ = \mathbb{Z}_+ \). Therefore, \( \rho = 1 \).

The denominator formula (8.29) becomes \( a_\rho = x - x^{-1} \) and the anti-invariant (8.28) for an arbitrary \( \mu \in \Lambda = \mathbb{Z} \) is given by \( a_\mu = x^{\mu} - x^{-\mu} \). Therefore, for \( \lambda = m \in \Lambda_+ = \mathbb{Z}_+ \), Weyl’s Character Formula gives the formula of Example 5.43 with \( t = x \):

\[
\text{ch} \mathcal{V}(m) = \frac{x^{m+1} - x^{-m-1}}{x - x^{-1}}.
\]

For the dimension, note that \( \langle \lambda, 2 \rangle = \langle \lambda, h \rangle = m \) and \( \langle \rho, 2 \rangle = \langle \rho, h \rangle = 1 \). So Theorem 8.30 gives the familiar formula \( \dim_z \mathcal{V}(m) = \frac{m+1}{2} = m + 1 \).
Dominant Weights for $sl_{n+1}$ and Partitions

Turning to $sl_{n+1}$, we continue with the setup of §§6.4.2, 7.1.4 and 8.4.2. In particular, we work with $\Delta = \{\alpha_1, \ldots, \alpha_n\}$, where $\alpha_i = e_i - e_{i+1}$. Then $\Phi_+ = \{e_i - e_j \mid 1 \leq i < j \leq n + 1\}$ and the fundamental weights are given by

$$\lambda_j = \mu_1 + \cdots + \mu_i$$

with $\mu_k = \varepsilon_k - \frac{1}{\varepsilon_k + 1} \sum_{j=1}^{n+1} \varepsilon_j$. Both $(\lambda_j)_i$ and $(\mu_j)_i$ are bases of the weight lattice $\Lambda$. So an arbitrary weight $\lambda \in \Lambda$ can be written in the two ways, with unique $m_k$, $l_i \in \mathbb{Z}$:

$$\lambda = m_1 \lambda_1 + \cdots + m_n \lambda_n = l_1 \mu_1 + \cdots + l_n \mu_n \quad \text{and} \quad l_i = \sum_{j \geq i} m_j.$$

Note that $\lambda$ is dominant (i.e., all $m_i \geq 0$) if and only if $l_1 \geq l_2 \geq \cdots \geq 0$. In other words, $\lambda \in \Lambda_+$ if and only if $(l_1, \ldots, l_n)$ is a partition (§4.3.1). Since this partition and the dominant weight $\lambda$ determine each other, we will allow ourselves to denote the partition by $\lambda$ as well; the Young diagram of $\lambda$ is displayed above. In particular,

$$\rho = (n, n-1, \ldots, 2, 1).$$

Dimensions for $sl_{n+1}$

Since $\langle \cdot, h_{e_i - e_j} \rangle = \langle \cdot, h_{e_i} - e_j \rangle = \langle \cdot, e_i - e_j \rangle$ by (6.28) and $\mu_k|_b = \varepsilon_k|_b$, we obtain from (8.30) the following equalities, where we put $l_{n+1} = 0$:

$$\langle \lambda, e_i - e_j \rangle = \sum_{k \geq 0} m_k \langle \lambda_k, e_i - e_j \rangle = \sum_{k \geq 0} m_k \sum_{l \leq k} \langle e_i, e_{i,l} - e_{j,l} \rangle$$

$$= \sum_{k \geq 0} (\delta_{i,l} - \delta_{j,l}) \sum_{k \geq l} m_k = \sum_{k \geq l} m_k - \sum_{k \geq j} m_k$$

$$= m_l + \cdots + m_{j-1} = l_i - l_j.$$

In particular, $\langle \rho, e_i - e_j \rangle = j - i$ by (8.31). Therefore, the dimension formula (Theorem 8.30), takes the following form:

$$\dim V(\lambda) = \prod_{1 \leq i < j \leq n+1} \frac{m_i + \cdots + m_{j-1} + j - i}{j - i}$$

(8.32)

Example 8.31 ($n = 2$). The product (8.32) has three factors: $(i, j) = (1, 2), (2, 3)$ or $(1, 3)$. Hence, $\dim V(\lambda) = \frac{1}{2}(m_1 + 1)(m_2 + 1)(m_1 + m_2 + 2)$. 


Example 8.32 (Fundamental representations). We already know that \( V(\lambda_k) = \Lambda^k V \), where \( V = \mathbb{k}^{n+1} \) is the defining representation of \( \mathfrak{s}l_{n+1} \) (§8.4.2); so we must have \( \dim_k V(\lambda_k) = \binom{n+1}{k} \). Indeed, with \( m_i = \delta_{i,k} \), formula (8.32) evaluates to

\[
\dim_k V(\lambda_k) = \prod_{1 \leq i < j \leq n+1}^{k} \frac{1 + j - i}{j - i} = \prod_{i=1}^{k} \prod_{j=k+1}^{n+1} \frac{1 + j - i}{j - i} = \prod_{i=1}^{k} \frac{n + 2 - i}{k + 1 - i} = \binom{n+1}{k}.
\]

Characters of \( \mathfrak{s}l_{n+1} \)

Put \( x_i = x_i^\mu \in \mathbb{Z} \Lambda \) as in §8.5.3. Thus, \( x_1 x_2 \ldots x_{n+1} = 1 \) and the Weyl group \( \mathcal{W} = S_{n+1} \) acts on \( \mathbb{Z} \Lambda = \mathbb{Z}[x_1, \ldots, x_{n+1}] = \mathbb{Z}[x_1, \ldots, x_{n}, x_{n+1}] \) by \( w \cdot x_i = x_{w_i} \).

With \( l_{n+1} = 0 \), we have

\[
x^\lambda = x_1^{l_1} x_2^{l_2} \ldots x_{n+1}^{l_{n+1}} = x_1^{l_1} x_2^{l_2} \ldots x_{n+1}^{l_{n+1}} \quad (\lambda \in \Lambda).
\]

Defining \( A_\lambda := (x_i^j)_{i,j=1,\ldots,n+1} \), the anti-invariants (8.28) take the following form:

\[
a_\lambda = \sum_{w \in S_{n+1}} \text{sgn}(w) x_1^{l_1} x_2^{l_2} \ldots x_{w(n+1)}^{l_{n+1}} = \det A_\lambda.
\]

In particular, \( A_\rho \) is the Vandermonde matrix:

\[
A_\rho = \begin{pmatrix}
x_1^n & x_1^{n-1} & \cdots & x_1 & 1 \\
x_2^n & x_2^{n-1} & \cdots & x_2 & 1 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
x_{n+1}^n & x_{n+1}^{n-1} & \cdots & x_{n+1} & 1
\end{pmatrix}.
\]

The denominator formula (8.29) gives the familiar product expansion of the Vandermonde determinant:

\[
a_\rho = \prod_{1 \leq i < j \leq n+1} (x_i^{\frac{1}{2}} x_j^{-\frac{1}{2}} - x_i^{-\frac{1}{2}} x_j^{\frac{1}{2}}) = \prod_{1 \leq i < j \leq n+1} (x_i - x_j) \prod_{1 \leq i < j \leq n+1} (x_i x_j)^{-\frac{1}{2}} = \prod_{1 \leq i < j \leq n+1} (x_i - x_j)(x_1 x_2 \ldots x_{n+1})^{-\frac{1}{2}} = \prod_{1 \leq i < j \leq n+1} (x_i - x_j).
\]
To summarize, Weyl’s Character Formula for $\mathfrak{sl}_{n+1}$ and $\lambda = l_1 \mu_1 + \cdots + l_n \mu_n \in \Lambda_+$ reads as follows:

$$\text{(8.34)} \quad \chi_4 = \frac{\det A_{\lambda+\rho}}{\prod_{1 \leq i < j \leq n+1} (x_i - x_j)} , \quad A_{\lambda+\rho} = \begin{pmatrix} x_1^{l_1+n} & x_1^{l_1+n-1} & \cdots & x_1^{l_1+1} & 1 \\ x_2^{l_2+n} & x_2^{l_2+n-1} & \cdots & x_2^{l_2+1} & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ x_n^{l_n+n} & x_n^{l_n+n-1} & \cdots & x_n^{l_n+1} & 1 \\ x_{n+1}^{l_{n+1}} & x_{n+1}^{l_{n+1}-1} & \cdots & x_{n+1}^{l_{n+1}+1} & 1 \end{pmatrix}.$$ 

**Example 8.33** (An irreducible representation of $\mathfrak{sl}_3$). Let $n = 2$ and consider the weight $\lambda = 3 \mu_1 + 5 \mu_2 = 8 \mu_1 + 5 \mu_2$ as in Figure 8.1. Then

$$A_{\lambda+\rho} = \begin{pmatrix} x_{10}^1 & x_1^6 & 1 \\ x_2^6 & x_2^6 & 1 \\ x_3^6 & x_3^6 & 1 \end{pmatrix}$$

and (8.34) becomes

$$\chi_4 = \frac{x_1^{10}(x_2^6 - x_3^6) - x_2^{10}(x_1^6 - x_3^6) + x_3^{10}(x_1^6 - x_2^6)}{(x_1 - x_2)(x_1 - x_3)(x_2 - x_3)}.$$ 

Using the relation $x_1 x_2 x_3 = 1$ and putting $y_1 = x^4_{1} = x^{\mu_1} = x_1$ and $y_2 = x^4_{2} = x^{\mu_1+\mu_2} = x_1 x_2$, the above expression for $\chi_4$ can be expanded into the following Laurent polynomial in $y_1$ and $y_2$, which explains the multiplicities in Figure 8.1. The underlined monomials, starting with $X^4 = y_1^3 y_2^5$, correspond to the points in the gray region of Figure 8.1; each of them is followed by the monomials in the same orbit under $\mathcal{W} = S_3$.

$$\chi_4 = \frac{y_1^3 y_2^5 + y_1^5 y_2^3 + y_1^5 y_2^7 + y_1^8 y_2^{-5} + y_1^{-3} y_2^8 + y_1^{-8} y_2^3}{(x_1 - x_2)(x_1 - x_3)(x_2 - x_3)}.$$
Example 8.34 (An irreducible representation of \( \mathfrak{sl}_4 \)). Now let us take \( n = 3 \) and 
\[ \lambda = \lambda_1 + 2\lambda_2 = 3\mu_1 + 2\mu_2. \]
Then
\[ A_{\lambda} = \begin{pmatrix} x_1^6 & x_1^4 & x_1 \ 0 & x_2^4 & x_2 \ 0 & 0 & x_3^4 \end{pmatrix} \]
and (8.34) gives the following formula, with highest term \( x_1^4 \) underlined:
\[ \text{ch}_A = x_1^3 x_2^2 + x_1^2 x_2 + x_1 x_3^2 + x_1 x_3 + x_1^3 x_4 + x_1 x_4^2 + x_2^2 x_4 + x_2 x_3^2 + x_2 x_3 + x_3^3 x_4 + x_3^2 x_4^2 + x_3 x_4 \]
\[ + x_1^2 x_2 x_3 + x_1 x_2 x_3^2 + x_1 x_2 x_4 + x_1 x_2 x_3 x_4 + x_1 x_2 x_4^2 + x_1 x_2 x_3 x_4^2 + x_1 x_2 x_4^2 + x_1 x_2 x_3 x_4 \]
\[ + x_1 x_2 x_3 x_4 + x_1 x_2 x_3 x_4 + x_1 x_2 x_3 x_4 + x_1 x_2 x_3 x_4 + x_1 x_2 x_3 x_4 + x_1 x_2 x_3 x_4 \]
\[ + 2 x_1^2 x_2 x_3 x_4 + 2 x_1 x_2 x_3 x_4^2 + 2 x_1 x_2 x_3 x_4^2 + 2 x_1 x_2 x_3 x_4^2 + 2 x_1 x_2 x_3 x_4^2 \]
\[ + 2 x_1 x_2 x_3 x_4^2 + 2 x_1 x_2 x_3 x_4 + 2 x_1 x_2 x_3 x_4^2 + 2 x_1 x_2 x_3 x_4^2 + 2 x_1 x_2 x_3 x_4 \]
\[ + 3 x_1 + 3 x_2 + 3 x_3 + 3 x_4. \]
Of course, \( \text{ch}_A \) can alternatively be expressed as a Laurent polynomial in \( y_i = x_i^{\mu_i} = \prod_{j \leq i} x_j \) for \( i \leq 3 \), but we will build the weight diagram of \( V(\lambda) \) directly from the above formula in terms of \( x_i = x_i^{\mu_i} \); see the picture (7.6) of \( A_3 \). The result is displayed in Figure 8.3. The Weyl group \( \mathcal{W} = S_4 \) acts on \( \mathbb{R}^3 \) by permuting \( \mu_1, \ldots, \mu_4 \); these weights happen to be the four weights of multiplicity 3. The twelve weights of multiplicity 2 also form one \( \mathcal{W} \)-orbit, but there are two \( \mathcal{W} \)-orbits of weights of multiplicity 1. The orbit of \( \lambda \), rendered in bright green, gives the extremal elements of the weight diagram as in §8.4.3.

8.7.3. Proof of Weyl’s Character Formula

Formal Characters Revisited

We will have to use formal characters in a slightly more general setting than in §8.5.2, which only covered finite-dimensional representations. Specifically, we will say that \( V \in \text{Rep}_\mathfrak{g} \) admits a formal character if \( V_{\lambda} \) is completely reducible with finite multiplicities: \( V = \bigoplus_{\mu \in h^*} V_\mu \) and \( \text{dim}_\mathbb{C} V_\mu < \infty \) for all \( \mu \in h^* \). All highest weight representations, finite dimensional or not, admit a formal character (Proposition 8.3) as do all subrepresentations and all images of representations admitting a formal character. For any such \( V \), we formally write
\[ \text{ch} V \overset{\text{def}}{=} \sum_{\mu \in h^*} (\text{dim}_\mathbb{C} V_\mu) x^\mu \in \mathbb{Z}^{h^*} \]
Here, $\mathbb{Z}^{b^*}$ denotes the $\mathbb{Z}$-module of all functions $f: b^* \to \mathbb{Z}$ and $x^\mu \in \mathbb{Z}^{b^*}$ is the function given by $x^\mu(v) = \delta_{\mu,v}$. We write a function $f \in \mathbb{Z}^{b^*}$ as a formal sum $f = \sum_{\mu \in b^*} f_\mu x^\mu$; so $f_\mu = f(\mu)$. The set $\text{Supp } f = \{ \mu \in b^* \mid f_\mu \neq 0 \}$ is called the support of $f$. The group ring $\mathbb{Z}[b^*]$ of the group $b^*$ arises as the collection of all finite sums of the above form, that is, the collection of functions $f$ having finite support. Alternatively, $\mathbb{Z}[b^*]$ is the $\mathbb{Z}$-submodule of $\mathbb{Z}^{b^*}$ that is generated by the functions $x^\mu$.

**Example 8.35** (Formal character of Verma modules). For the Verma module $M(\lambda)$, we know by Proposition 8.7 that $\dim_k M(\lambda)_\mu = \varphi_+(\lambda - \mu)$, where $\varphi_+$ is the vector partition function (8.12). Thus, $\varphi_+(\lambda - \mu) \neq 0$ if and only if $\lambda - \mu \in L_+$. Writing $\nu = \lambda - \mu$, we obtain

$$\text{ch } M(\lambda) = \sum_{\mu \in b^*} \varphi_+(\lambda - \mu) x^\mu = \sum_{\nu \in L_+} \varphi_+(\nu) x^{\lambda - \nu}.$$ 

Observe that, for any short exact sequence $0 \to U \to V \to W \to 0$ of representations in $\text{Rep} \mathfrak{g}$ that all admit formal characters, the equality

$$(8.35) \quad \text{ch } V = \text{ch } U + \text{ch } W$$

holds in $\mathbb{Z}^{b^*}$. Indeed, since the restriction of the sequence to $\text{Rep} \mathfrak{h}$ splits, we have $\dim_k V_\mu = \dim_k U_\mu + \dim_k W_\mu$ for all $\mu \in b^*$.

**Multiplication of Functions**

Below, we will have to consider products of certain functions on $b^*$, generally not with finite support. The multiplication of the group ring $\mathbb{Z}[b^*]$ does not extend to
all of \( \mathbb{Z}^{b^*} \). Instead, we focus on the \( \mathbb{Z} \)-submodule of \( \mathbb{Z} \langle b^* \rangle \subseteq \mathbb{Z}^{b^*} \) that is generated by all functions \( f \in \mathbb{Z}^{b^*} \) whose support is bounded above for the partial order \( \preceq \), that is, there is some \( \lambda \in b^* \) such that \( \mu \preceq \lambda \) for all \( \mu \in \text{Supp } f \). In other words, \( f \in \mathbb{Z} \langle b^* \rangle \) if and only if \( \text{Supp } f \) is a finite union of subsets of \( b^* \), each of which has an upper bound in \( b^* \) for \( \preceq \). Evidently, the functions \( x^\mu \) all belong to \( \mathbb{Z} \langle b^* \rangle \) and so \( \mathbb{Z} \langle b^* \rangle \subseteq \mathbb{Z} \langle b^* \rangle \). The multiplication of \( \mathbb{Z} \langle b^* \rangle \) extends to \( \mathbb{Z} \langle b^* \rangle \) for \( \preceq \). We may define

\[
fg = \sum_{\mu \in \mathbb{Z}^{b^*}} \left( \sum_{\mu' + \mu = \nu} f_{\mu'} g_{\mu} \right) x^\nu \quad (f, g \in \mathbb{Z} \langle b^* \rangle).
\]

To see that this makes sense, one needs to observe that the sums \( \sum_{\mu' + \mu = \nu} f_{\mu'} g_{\mu} \) are in fact finite sums (Exercise 8.7.2). This multiplication makes \( \mathbb{Z} \langle b^* \rangle \) a commutative ring containing \( \mathbb{Z}[b^*] \) as a subring, with \( 1 = x^0 \) as the identity element.

For any \( \lambda \in L_+ \), the geometric series \( 1 + x^{-1} + x^{-2} + \ldots \) belongs to \( \mathbb{Z} \langle b^* \rangle \), having support bounded above by 0 in \( b^* \); in fact, the series is a unit in \( \mathbb{Z} \langle b^* \rangle \):

\[
(1 + x^{-1} + x^{-2} + \ldots)(1 - x^{-1}) = 1.
\]

If \( V \in \text{Rep } g \) is a highest weight representation, then \( \text{ch } V \in \mathbb{Z} \langle b^* \rangle \), the highest weight of \( V \) being an upper bound for the support of \( \text{ch } V \) (Example 8.3). Moreover, if \( W \) is a subrepresentation or a homomorphic image of a representation \( V \in \text{Rep } g \) that admits a formal character \( \text{ch } V \in \mathbb{Z} \langle b^* \rangle \), then we also have \( \text{ch } W \in \mathbb{Z} \langle b^* \rangle \), because \( \text{Supp } \text{ch } W \subseteq \text{Supp } \text{ch } V \) by (8.35).

The following lemma gives an alternative expression for \( \text{ch } M(\lambda) \) to the one in Example 8.35 and it also shows that the denominator \( a_\rho \) of Weyl’s Character Formula is a unit of \( \mathbb{Z} \langle b^* \rangle \). We put \( k := \text{ch } M(\lambda) \in \mathbb{Z} \langle b^* \rangle \); so Example 8.35 gives

\[
k = \sum_{v \in L_+} \varphi_+(v) x^{-v}.
\]

**Lemma 8.36.** In \( \mathbb{Z} \langle b^* \rangle \), we have \( a_\rho \cdot k \cdot x^{-\rho} = 1 \) and \( \text{ch } M(\lambda) = a_\rho^{-1} x^{\lambda + \rho} \) \( (\lambda \in b^* \).)

**Proof.** Since \( L_+ = \sum_{\alpha \in \Phi_+} \mathbb{Z}_+ \alpha \), it is clear that \( k \) can be written as the product

\[
k = \prod_{\alpha \in \Phi_+} (1 + x^{-\alpha} + x^{-2\alpha} + \ldots),
\]

which is inverse in \( \mathbb{Z} \langle b^* \rangle \) to \( \prod_{\alpha \in \Phi_+} (1 - x^{-\alpha}) \), as we have remarked above. Now recall from (8.29) that \( a_\rho = x^\rho \prod_{\alpha \in \Phi_+} (1 - x^{-\alpha}) \) to obtain the equality \( a_\rho \cdot k \cdot x^{-\rho} = x^0 = 1 \). Writing the formula in Example 8.35 as \( \text{ch } M(\lambda) = k x^\lambda \), we also obtain \( \text{ch } M(\lambda) = a_\rho^{-1} x^{\lambda + \rho} = a_\rho^{-1} x^{\lambda + \rho} \).

**The Proof**

We are now ready to give the proof of Weyl’s Character Formula. To start, let \( \lambda \in b^* \) be arbitrary and put \( \text{ch } \mathcal{L} = \text{ch } V(\lambda) \in \mathbb{Z} \langle b^* \rangle \), extending our earlier notation for the case \( \lambda \in \Lambda_+ \). Furthermore, put

\[
\mathcal{M} = \mathcal{M}(\lambda) := \{ \mu \in b^* \mid \mu \preceq \lambda \text{ and } W \cdot \mu = W \cdot \lambda \}.
\]
By Proposition 8.11, the Verma module $M(\lambda)$ has finite length, with composition factors of the form $V(\mu)$ ($\mu \in \mathcal{M}$) and with $V(\lambda)$ having multiplicity 1. Therefore, (8.35) gives $\text{ch} M(\lambda) = \sum_{\mu \in \mathcal{M}} z_\mu \text{ch}_\mu$ for suitable $z_\mu \in \mathbb{Z}_+$ with $z_\lambda = 1$. A similar expression holds for each $\text{ch} M(\mu)$ with $\mu \in \mathcal{M}$ and we also have $\mathcal{M}(\mu) \subseteq \mathcal{M}$ for all these $\mu$. Thus, we obtain a system of equations with coefficients $z_{\mu,\nu} \in \mathbb{Z}_+$:

$$
\text{ch} M(\mu) = \sum_{\nu \in \mathcal{M}} z_{\mu,\nu} \text{ch}_\nu \quad (\mu \in \mathcal{M}).
$$

Here, $z_{\nu,\nu} = 1$ and $z_{\mu,\nu} = 0$ if $\nu \not\equiv \mu$. Choose some total order $\leq$ on the finite set $\mathcal{M}$, that is compatible with the partial order $\leq$ in the sense that $\mu \leq \mu'$ implies $\mu \leq \mu'$ for $\mu, \mu' \in \mathcal{M}$. Then the coefficient matrix $(z_{\mu,\nu})$ of the above system is untriangular, and hence it has a untriangular inverse over $\mathbb{Z}$. Therefore, for suitable $c_\mu \in \mathbb{Z}$ ($\mu \in \mathcal{M}$) with $c_\lambda = 1$, we have $\text{ch}_1 = \sum_{\mu \in \mathcal{M}} c_\mu \text{ch} M(\mu)$ and hence

$$
a_\rho \text{ch}_1 = \sum_{\mu \in \mathcal{M}} c_\mu x^{\mu + \rho}.
$$

Now we specialize the foregoing to $\lambda \in \Lambda_+$. Then $\text{ch}_1 \in (\mathbb{Z}\Lambda)^{\mathcal{W}}$ (§8.5.2) and $a_\rho \text{ch}_1 \in (\mathbb{Z}\Lambda)^{^\times}$, the module of anti-invariants for $\mathcal{W}$ in the group ring $\mathbb{Z}\Lambda$. A $\mathbb{Z}$-basis of $(\mathbb{Z}\Lambda)^{^\times}$ is given by the elements $a_{\lambda'} = \sum_{w \in \mathcal{W}} \text{sgn}(w) x^{w \lambda'}$, where $\lambda'$ runs over the set $\rho + \Lambda_+$ of strongly dominant weights (Proposition 7.18). Since $\mathcal{M} \subseteq \mathcal{W} \cdot \lambda = \mathcal{W}(\lambda + \rho) - \rho$ and the strongly dominant weight $\lambda + \rho$ has trivial isotropy group in $\mathcal{W}$ (Proposition 7.16), we may write the above expression for $a_\rho \text{ch}_1$ in the following form, with $c_w \in \mathbb{Z}$ and $c_1 = 1$:

$$
a_\rho \text{ch}_1 = \sum_{w \in \mathcal{W}} c_w x^{w(\lambda + \rho)}.
$$

Therefore, we must have $a_\rho \text{ch}_1 = c a_{\lambda + \rho}$ for some $c \in \mathbb{Z}$; in fact, $c = 1$ because $c_1 = 1$. This completes the proof of Weyl’s Character Formula. \hfill \Box

### Exercises for Section 8.7

**8.7.1** (Some irreducible representations of $\mathfrak{so}_5$ (type $B_2$)). Recall that the root system $B_2$ consists of the vectors $\pm e_1, \pm e_2$ and $\pm e_1 \pm e_2$; the Weyl group is $\mathcal{W} = \{ \pm 1 \}^2 \times S_2$; a base is $\Delta = \{ \alpha_1 = e_1 - e_2, \alpha_2 = e_2 \}$; and the fundamental weights for $\Delta$ are $\lambda_1 = \frac{1}{2}(e_1 - e_2)$ (Exercise 7.4.1). Show:

(a) $\dim_k V(\lambda) = \frac{1}{6}(m_1 + 1)(m_2 + 1)(m_1 + m_2 + 2)(2m_1 + m_2 + 3)$ for $\lambda = m_1 \lambda_1 + m_2 \lambda_2 \in \Lambda_+$ ($m_i \in \mathbb{Z}_+$).

(b) $\text{ch}_{\lambda_1} = x^0 + \sigma_{\lambda_1}$ (not an orbit sum) and $\text{ch}_{\lambda_2} = \sigma_{\lambda_2}$, where $\sigma_{\lambda_i}$ denotes the $\mathcal{W}$-orbit sum of $x^{\lambda_i}$ (as in Theorem 7.17).

(c) $V(\lambda_1)$ is the defining representation $V = k^5$ of $\mathfrak{so}_5$ (notation as in §6.4.3).
8.7.2 (Multiplication in $\mathbb{Z}(\mathfrak{h}^*)$). Show that the product $fg = \sum_{\nu \in \mathfrak{h}^*} (\sum_{\mu' + \mu = \nu} f_{\mu'} g_{\mu}) x^\nu$ for $f = \sum_{\mu \in \mathfrak{h}^*} f_{\mu} x^\mu, g = \sum_{\mu \in \mathfrak{h}^*} g_{\mu} x^\mu \in \mathbb{Z}(\mathfrak{h}^*)$ is well-defined (notation as in §8.7.3).

8.7.3 (Dimension polynomial). Prove:

(a) There is a unique polynomial function $D \in O(\mathfrak{h}^*) = \text{Sym}(\mathfrak{h})$ such that $D(\lambda) = \dim_k V(\lambda)$ for $\lambda \in \Lambda_+$; it has degree $\deg D = \# \Phi_+$. 

(b) $D(a \rho) = (a + 1)^{\# \Phi^*}$ for $a \in k$. 

(c) For any given $N \in \mathbb{N}$, there are at most finitely many $\lambda \in \Lambda_+$ with $D(\lambda) \leq N$.

(d) $D(w \cdot (\lambda - \rho)) = \text{sgn}(w) D(\lambda - \rho)$ for $\lambda \in \mathfrak{h}^*$ and $w \in W$. Consequently, the function $\lambda \mapsto D(\lambda - \rho)^2$ belongs to $(\text{Sym} \mathfrak{h})^W$. 

(e) There is a unique element $z \in \mathcal{Z}(U_k)$ such that $\chi_\lambda(z) = D(\lambda)^2$ for $\lambda \in \mathfrak{h}^*$. (Use Theorem 8.23.) What is this element for $\mathfrak{s}\mathfrak{l}_2$?

8.7.4 (Dimension of invariants). Put $d = a_\rho$, as in Proposition 7.18 and let $\cdot^*$ denote the standard involution (3.28) of the group ring $\mathbb{Z} \Lambda$; so $(x^\lambda)^* = x^{-\lambda}$ for $\lambda \in \Lambda$. Define $\mathbb{Z}$-linear maps $\delta : \mathbb{Z} \Lambda \to \mathbb{Z}, x^\lambda \mapsto \delta_{0, \lambda}$, and $I : \mathbb{Z} \Lambda \to \mathbb{Q}, f \mapsto \frac{1}{|W|} \delta(dd^* f)$. For $\lambda, \mu \in \Lambda_+$ and $V \in \text{Rep}_{\text{fin}} \mathfrak{g}$, prove:

(a) $dd^* = \prod_{\alpha \in \Phi}(1 - x^\alpha) = (-1)^{\# \Phi^*} d^2$.

(b) $I(\text{ch}_\lambda) = \delta_{0, \lambda}$. (Use Weyl’s Character Formula and (8.28).)

(c) $I(\text{ch} V) = \dim_k V^\lambda$. (Use Weyl’s Theorem (Section 6.2) and (b).)

(d) $I(\text{ch}_\lambda^* \text{ch}_\mu) = \delta_{\lambda, \mu}$. (Use Exercise 8.5.1 and (c).) With $D(\lambda) = \dim_k V(\lambda)$ as in Exercise 8.7.3, $\dim_k V = \sum_{\lambda \in \Lambda_+} I(\text{ch}_\lambda V) D(\lambda)$.

8.8. Schur Functors and Representations of $\mathfrak{s}\mathfrak{l}(V)$

In this short section, we pick up the thread of Schur-Weyl duality from Section 4.7. With the aid of the Lie theoretic material now at our disposal, particularly the information we have amassed for the special linear Lie algebra $\mathfrak{s}\mathfrak{l}(V)$, we will be able to answer some questions on representations of the general linear group $\text{GL}(V)$ that were left open earlier. The reader wishing to delve deeper into the connections between representations of $\text{GL}(V)$ and other “classical” groups and their associated Lie algebras is once again referred to the monographs by Weyl [205], Goodman and Wallach [88], [89] and Procesi [168]. In addition, [90] by Green is an excellent source for representations of $\text{GL}(V)$ over an infinite base field of arbitrary characteristic.

Throughout this section, we fix $0 \neq V \in \text{Vec}_k$ with $\dim V = d < \infty$. The base field $k$ is understood to be algebraically closed with char $k = 0$, as in the rest of this chapter, and it will frequently be omitted from our notation, as in Chapter 4.
8.8. Schur Functors and Representations of \(\mathfrak{sl}(V)\)  

8.8.1. The Action of \(\mathfrak{gl}(V)\) and \(\mathfrak{sl}(V)\) on \(V^\otimes n\)

Fix a positive integer \(n\). Recall that the place permutation action \(S_n \subset V^\otimes n\) and the diagonal action \(GL(V) \subset V^\otimes n\) commute with each other and that the resulting algebra map \(\mathbb{k}[GL(V)] \to \text{End}_{S_n}(V^\otimes n)\) is surjective by Schur’s Double Centralizer Theorem. We use the standard identification \(\text{End}_k(V^\otimes n) = \text{End}_k(V^\otimes n)\); see (B.17).

As we have remarked in the proof of Schur’s Double Centralizer Theorem, this identification respects the \(S_n\)-actions given by place permutations, and hence it gives an identification \((\text{End}_k(V)^{\otimes n})_{S_n} = (\text{End}_k(V^\otimes n))_{S_n} = \text{End}_{S_n}(V^\otimes n)\).

The standard action (5.31) of the Lie algebra \(\mathfrak{gl}(V)\) on \(V^\otimes n\) can be written in the following form:

\[
x_{V^\otimes n} = \sum_i x_i \quad \text{with} \quad x_i = \text{Id}_V \otimes \ldots \otimes x \otimes \ldots \otimes \text{Id}_V \quad (x \in \mathfrak{gl}(V)).
\]

For a given \(x \in \mathfrak{gl}(V)\), the elements \(x_i \in \text{End}_k(V^\otimes n)\) commute pairwise and \(s \cdot x_i = x_{sl}\) for \(s \in S_n\). In particular, the \(\mathfrak{gl}(V)\)-action on \(V^\otimes n\) commutes with the place permutation action of \(S_n\), giving rise to an algebra map \(U(\mathfrak{gl}(V)) \to \text{End}_{S_n}(V^\otimes n)\), \(u \mapsto u_{V^\otimes n}\). We claim that this map is also surjective. To prove this, note that, for each \(x \in \mathfrak{gl}(V)\) and \(k \in \mathbb{N}\), the composition power \(x^k\) belongs to \(\text{End}_k(V) = \mathfrak{gl}(V)\) and it acts by \((x^k)_{V^\otimes n} = \sum_i x_i^k = p_k\), the \(k\)th power sum of \(x_1, \ldots, x_n\). Thus, all \(p_k\) belong to the image of the map \(U(\mathfrak{gl}(V)) \to \text{End}_{S_n}(V^\otimes n)\). It now follows from the Newton identities (3.56) that the elementary symmetric polynomials \(e_k\) all belong to the image as well. In particular, \(e_n = x_1 x_2 \cdots x_n = x \otimes x \otimes \cdots \otimes x\) belongs to the image for each \(x \in \mathfrak{gl}(V)\). Since these endomorphisms span \(\text{End}_{S_n}(V^\otimes n)\) (Proposition 3.37), our claim is proved.

Finally, \(\mathfrak{gl}(V) = \mathbb{k} \text{Id}_V \oplus \mathfrak{sl}(V)\) with \(\text{Id}_V\) acting on \(V^\otimes n\) by multiplication with \(n\). Thus, to summarize, the standard actions of \(GL(V)\) and of \(\mathfrak{gl}(V)\) and \(\mathfrak{sl}(V)\) on \(V^\otimes n\) give rise to algebra maps with the same image in \(\text{End}_k(V^\otimes n)\):

\[
\begin{array}{ccc}
U(\mathfrak{sl}(V)) & \longrightarrow & \text{End}_{S_n}(V^\otimes n) \\
\uparrow & & \uparrow \\
U(\mathfrak{gl}(V)) & \longrightarrow & \mathbb{k}[GL(V)]
\end{array}
\]

Consequently, a subspace of \(V^\otimes n\) is stable under the action of \(\mathfrak{gl}(V)\) or \(\mathfrak{sl}(V)\) if and only if it is stable under \(GL(V)\).

8.8.2. Schur Functors Revisited

We now return to the setting of §4.7.3. Thus, we fix a partition \(\lambda = (l_1, l_2, \ldots) \in \mathcal{P}\) having at most \(d = \dim V\) nonzero parts: \(\ell(\lambda) = \max\{i \mid l_i \neq 0\} \leq d\). Putting \(n = |\lambda| = \sum_i l_i\) and letting \(V^\lambda \in \text{Irr}\, S_n\) denote the representation associated to \(\lambda\)
(§4.3.2), we consider the following \( k \)-vector space:

\[
S^4V = \text{Hom}_{S_n}(V^4, V^\otimes n).
\]

The algebra \( A' = \text{End}_{S_n}(V^\otimes n) \) acts on \( S^4V \) by \((a.f)(w) = a(f(w)) \) for \( a \in A' \), \( w \in V^4 \) and \( f \in S^4V \) and this action makes \( S^4V \) an irreducible representation of \( A' \) (Theorem 4.29). It follows from (8.36) that \( S^4V \) is an irreducible representation of \( \text{GL}(V) \), as already discussed in §4.7.3, and also of \( \text{gl}(V) \) and \( \text{sl}(V) \). Thus:

**Proposition 8.37.** Let \( \lambda \) be a partition with \( \ell(\lambda) \leq d = \dim V \). Then \( S^4V \) is an irreducible representation of \( \text{gl}(V) \) and of \( \text{sl}(V) \).

For each, \( 0 \neq w \in V^4 \), the map \( f \mapsto f(w) \) gives an embedding \( S^4V \hookrightarrow V^\otimes n \) in \( \text{Rep} \; A' \). These maps yield the decomposition of \( V^\otimes n \) into homogeneous components for \( \text{GL}(V) \) and for \( \text{gl}(V) \) and \( \text{sl}(V) \); see Theorem 4.32 and (4.53):

\[
V^\otimes n \cong \bigoplus_{n: \ell(\lambda) \leq d} S^\lambda V \otimes V^\lambda.
\]

(8.37)

**Example 8.38** (Symmetric and exterior powers). As we have observed in Example 4.30, the above embedding identifies \( S^{(n)}V = \text{Hom}_{S_n}(1, V^\otimes n) \) with \( (V^\otimes n)^{S_n} \), which in turn maps isomorphically onto \( \text{Sym}^n V \) under the canonical map \( V^\otimes n \twoheadrightarrow \text{Sym}^n V \) (Lemma 3.36). Since the latter map is a morphism of \( \text{gl}(V) \)-representations, we obtain the following isomorphism in \( \text{Rep} \; \text{gl}(V) \):

\[
S^{(n)}V \cong \text{Sym}^n V.
\]

Similarly, for the partition \((1, \ldots, 1) \vdash n \), we obtain an identification of \( S^{(1,\ldots,1)}V = \text{Hom}_{S_n}(1, V^\otimes n) \) with \( (V^\otimes n)^{S_n} \cong \Lambda^n V \), which gives the following \( \text{gl}(V) \)-isomorphism:

\[
S^{(1,\ldots,1)}V \cong \Lambda^n V.
\]

Thus, all symmetric and exterior powers of \( V \) are (finite-dimensional) irreducible representations of \( \text{gl}(V) \) and \( \text{sl}(V) \). The exterior powers \( \Lambda^n V \), as representations of \( \text{sl}(V) \), have already been discussed in §8.4.2.

#### 8.8.3. \( S^4V \) as a Highest Weight Representation

Since \( S^4V \) is a finite-dimensional irreducible representation of the semisimple Lie algebra \( \text{sl}(V) \) (Proposition 8.37), it follows that \( S^4V \cong V(\lambda') \) for a unique weight \( \lambda' \in \Lambda^+_+ \) (8.18). Our goal in this subsection is to identify \( \lambda' \).

Let us start with a few reminders about the special linear Lie algebra. Since \( \text{sl}(V) \cong \text{sl}_d \) via a choice of basis for \( V \), we may use the setup of §8.4.2 with some small adjustments of notation. In particular, we will work with the Cartan subalgebra \( \mathfrak{h} = \text{sl}_d \cap \mathfrak{g}_d \). The set of roots for \( \mathfrak{h} \) is \( \Phi = \{ e_i - e_j \mid 1 \leq i \neq j \leq d \} \),
where \((e_j)^d\) denotes the dual basis of \(s^*_d\) for the standard basis \((e_i)^d\) of \(s_d\). The fundamental weights for the usual base of \(\Phi\) are given by \(\lambda_i = \mu_1 + \cdots + \mu_i\) \((1 \leq i \leq d - 1)\) with

\[
\lambda_i = e_i - \frac{1}{d} \sum_{j=1}^{d} e_j.
\]

As was already noted in (8.30), an arbitrary weight in \(\Lambda\) can be written in the following two ways, with unique \(m_j, l_i \in \mathbb{Z}\) that are related by \(l_i = \sum_{j \geq i} m_j\):

\[
m_1 \lambda_1 + \cdots + m_{d-1} \lambda_{d-1} = l_1 \mu_1 + \cdots + l_{d-1} \mu_{d-1}.
\]

The weight in question is dominant (i.e., all \(m_i \geq 0\)) if and only if \((l_1, \ldots, l_{d-1})\) is a partition (i.e., \(l_1 \geq l_2 \geq \cdots \geq 0\)). In this way, we obtain a bijection between \(\Lambda_+\) and the set of all 

\[
\lambda \in \mathcal{P} \mid \ell(\lambda) \leq d \quad \mapsto \quad \lambda' = \sum_{i=1}^{d} l_i \mu_i = \sum_{i=1}^{d-1} (l_i - l_d) \mu_i
\]

Proposition 8.39. Let \(\lambda\) be a partition with \(\ell(\lambda) \leq d = \dim V\). Then \(S^d V \cong V(\lambda')\) in \(\text{Rep} \; \mathfrak{sl}(V)\) with \(\lambda' \in \Lambda_+\) as in (8.38).

Before proving this, let us state a corollary, which answers a question that was left open in connection with (4.54), and give some examples. The corollary follows by applying formula (8.32) to the weight \(\lambda' = \sum_{i=1}^{d-1} (l_i - l_d) \mu_i\).

Corollary 8.40. If \(\lambda = (l_1, l_2, \ldots, l_d)\), then \(\dim S^d V = \prod_{1 \leq i < j \leq d} \frac{l_i - l_j + j - i}{j - i}\).

Example 8.41 (Symmetric and exterior powers, again). Applying the proposition to the partition \((1, 1, \ldots, 1)\) with \(n \leq d\), we obtain

\[
S^{(1,\ldots,1)} V \cong V(\mu_1 + \cdots + \mu_n) = V(\lambda_n),
\]

which we knew already: \(S^{(1,\ldots,1)} V \cong \Lambda^n V \cong V(\lambda_n)\) (Example 8.38 and §8.4.2). For the partition \((n)\), we obtain

\[
S^{(n)} V \cong \text{Sym}^n V \cong V(n \mu_1).
\]

Proof of Proposition 8.39. Fix a \(\mathbb{K}\)-basis \((x_i)^d\) of \(V\) and put \(n = |\lambda|\). The monomials \(x_{i_1} \otimes x_{i_2} \otimes \cdots \otimes x_{i_n}\) form a \(\mathbb{K}\)-basis of \(V^{\otimes n}\) that is permutated by the \(S_n\)-action. A transversal for the \(S_n\)-orbits on these monomials is given by the monomials

\[
x_m = x_{m_1} \otimes x_{m_2} \otimes \cdots \otimes x_{m_d}
\]
with $\mathbf{m} = (m_1, m_2, \ldots, m_d) \in \mathbb{Z}^d$ and $|\mathbf{m}| = \sum_i m_i = n$. The isotropy group of $x_\mathbf{m}$ is the Young subgroup $S_n$; so $\mathbb{k}S_n x_\mathbf{m} \cong 1 \uparrow_{S_n}^{S^m}$. This yields the familiar description (3.70) of $V^{\otimes n} \in \text{Rep} S_n$:

$$V^{\otimes n} = \bigoplus_{\mathbf{m} \in \mathbb{Z}^d : |\mathbf{m}| = n} (V^{\otimes n})_\mathbf{m} \quad \text{with} \quad (V^{\otimes n})_\mathbf{m} := \mathbb{k}S_n x_\mathbf{m} \cong 1 \uparrow_{S_n}^{S^m}.$$ 

Putting $H_\mathbf{m} := \text{Hom}_{S_n} (V^A, (V^{\otimes n})_\mathbf{m})$, we obtain the decomposition

$$\mathbb{S}^A V = \text{Hom}_{S_n} (V^A, V^{\otimes n}) \cong \bigoplus_{\mathbf{m} \in \mathbb{Z}^d : |\mathbf{m}| = n} H_\mathbf{m}$$

with

$$H_\mathbf{m} \cong \text{Hom}_{S_n} (V^A, 1 \uparrow_{S_n}^{S^m}) \cong \text{Hom}_{S_n} (V^A \downarrow_{S_n}, 1) \cong ((V^A)^{S^m})^*.$$ 

Here, the last isomorphism is given by restriction of homomorphisms to invariants.

Now let us turn to the action of $\mathfrak{gl}(V)$, which we identify with $\mathfrak{gl}_d$, by means of the basis $(x_i)^d$ of $V$ as above. The monomials $x_\mathbf{m}$ for $\mathbf{m} = (m_1, m_2, \ldots, m_d)$ are weight vectors for the diagonal subalgebra $\mathfrak{d}_d = \bigoplus_{i=1}^d \mathbb{k} e_i \subseteq \mathfrak{gl}_d$:

$$e_{i,j} x_\mathbf{m} = m_i x_\mathbf{m}.$$ 

Since the actions of $\mathfrak{gl}(V)$ and $S_n$ on $V^{\otimes n}$ commute, it follows that the summand $(V^{\otimes n})_\mathbf{m} = \mathbb{k}S_n x_\mathbf{m}$ is the $\mathfrak{d}_d$-weight space of $V^{\otimes n}$ for the weight $\sum_{i=1}^d m_i e_i \in \mathfrak{d}_d$. Identifying this weight with $\mathbf{m}$, note that $\mathbf{m}|_{\mathfrak{h}} = \sum_{i=1}^d m_i \mu_i \in \mathfrak{h}^*$. Similarly, one sees that $e_{i,j}(V^{\otimes n})_\mathbf{m} = 0$ if $m_j = 0$ and $e_{i,j}(V^{\otimes n})_\mathbf{m} \subseteq (V^{\otimes n})_{\mathbf{m}+\epsilon_i-\epsilon_j}$ if $m_j > 0$.

All this transfers directly to the $\mathfrak{gl}(V)$-action on $\mathbb{S}^A V$. The matrix $e_{i,j}$ acts as the scalar $m_i$ on the summand $H_\mathbf{m}$, which is therefore the $\mathfrak{d}_d$-weight space of $\mathbb{S}^A V$ with weight $\mathbf{m} = \sum_{i=1}^d m_i e_i$; elements of $H_\mathbf{m}$ have weight $\sum_{i=1}^d m_i \mu_i$ for $\mathfrak{h}$. Further, $e_{i,j} H_\mathbf{m} = 0$ if $m_j = 0$ and $e_{i,j} H_\mathbf{m} \subseteq H_{\mathbf{m}+\epsilon_i-\epsilon_j}$ if $m_j > 0$. Recall that $H_\mathbf{m} \cong ((V^A)^{S^m})^*$. We have seen earlier that $H_\mathbf{m} \cong ((V^A)^{S^m})^*$ is 1-dimensional for $\mathbf{m} = \lambda = (l_1, l_2, \ldots, l_d)$ while $H_\mathbf{m} = 0$ for $\mathbf{m} = \lambda + \epsilon_i - \epsilon_j$ with $i < j$ (Lemma 4.31). Thus, the unique (up to scalar multiples) $0 \neq f_\lambda \in H_\lambda$ has weight $\lambda' = \sum_{i=1}^d l_i \mu_i$ for $\mathfrak{h}$ and $e_{i,j} f_\lambda = 0$ for $i < j$; so $n_+ f_\lambda = 0$. Finally, $f_\lambda$ generates $\mathbb{S}^A V$ by irreducibility (Proposition 8.37). This completes the proof of the proposition. \hfill \Box

**Representations of $\mathfrak{gl}(V)$**. The difference between $\text{Irr} \mathfrak{sl}(V)$ and $\text{Irr} \mathfrak{gl}(V)$ is not very significant and it comes, naturally enough, from the trace. By Proposition 8.39 and (8.18), the restrictions of the representations $\mathbb{S}^A V$ yield all of $\text{Irr} \mathfrak{sl}(V)$, with some duplication due to the non-injectivity of the map (8.38). Observe that the element $1d_V \in \mathfrak{gl}(V)$ acts on $\mathbb{S}^A V$ as the scalar $n = |\lambda| \in \mathbb{Z}_+$, because it does so on $V^{\otimes n}$. Twisting the various $\mathbb{S}^A V \in \text{Irr} \mathfrak{gl}(V)$ with 1-dimensional representations
defined by scalar multiples of the trace, one obtains actions of $\text{Id}_V$ by arbitrary scalars and a bijection

$$\text{Irr}_{\text{fin}} gl(V) \xrightarrow{\sim} \mathbb{k} \times \Lambda_+.$$ 

For details, see Exercises 8.8.2 and 8.8.3.

### Exercises for Section 8.8

*In these exercises, $V$ denotes a nonzero $\mathbb{k}$-vector space with $d = \dim V < \infty$.

**8.8.1** (Isomorphism). Let $\lambda$ and $\mu$ be distinct partitions with $\ell(\lambda), \ell(\mu) \leq d$. Recall that $S^\lambda V \not\cong S^\mu V$ in $\text{Rep GL}(V)$ (Theorem 4.32). Prove:

(a) If $|\lambda| = |\mu|$, then $S^\lambda V \not\cong S^\mu V$ in $\text{Rep sl}(V)$, and hence also in $\text{Rep gl}(V)$.

(b) If $|\lambda| \neq |\mu|$, then $S^\lambda V \not\cong S^\mu V$ in $\text{Rep gl}(V)$, but not necessarily in $\text{Rep sl}(V)$.

**8.8.2** (Extending the definition of $S^\lambda V$). (a) For $\xi = (x_1, \ldots, x_d) \in \mathbb{k}^d$, put $|\xi| = \sum_i x_i$ and $\xi' = \sum_{i=1}^d x_i \mu_i = \sum_{i=1}^{d-1} (x_i - x_d) \mu_i$ as in (8.38). Let the additive group $(\mathbb{k}, +)$ act on $\mathbb{k}^d$ by putting $\xi + k = (x_1 + k, \ldots, x_d + k) \in \mathbb{k}^d$ for $k \in \mathbb{k}$ and let $\mathcal{X} = \{\lambda + k \mid \lambda = (l_1, \ldots, l_d) \in \mathcal{P}, k \in \mathbb{k}\}$ denote the union of the $\mathbb{k}$-orbits of all partitions with at most $d$ nonzero parts. Show that the map $\xi \mapsto (|\xi|, \xi')$ gives a bijection $\mathcal{X} \xrightarrow{\sim} \mathbb{k} \times \Lambda_+$.

(b) Let $\xi = (x_1, \ldots, x_d) \in \mathcal{X}$ and note that $\xi - x_d$ is a partition. Define

$$S^\xi V \overset{\text{def}}{=} S^{\xi - x_d} V \otimes T^{x_d} \in \text{Rep gl}(V),$$

where we have put $T^k = \mathbb{k}_{k\text{trace}} \in \text{Irr}_{\text{fin}} gl(V)$ for $k \in \mathbb{k}$ (Exercise 5.5.5). Show that $\text{Id}_V \in gl(V)$ acts on $S^\xi V$ as the scalar $|\xi|$ and $S^\xi V \downarrow_{\text{sl}(V)} = V(\xi')$. Conclude that $S^\xi V$ is irreducible and $S^\xi V \not\cong S^\eta V$ for distinct $\xi, \eta \in \mathcal{X}$.

**8.8.3** (Representations of $gl(V)$). Let $W \in \text{Irr}_{\text{fin}} gl(V)$. Show that $W \cong S^\xi V$ for a unique $\xi \in \mathcal{X}$ (Exercise 8.8.2) by completing the following steps:

(a) The element $\text{Id}_V \in gl(V)$ acts on $W$ as a scalar, say $t \in \mathbb{k}$.

(b) $W \downarrow_{\text{sl}(V)} \cong V(\lambda)$ for some $\lambda \in \Lambda_+$.

(c) If $\xi \in \mathcal{X}$ is such that $(|\xi|, \xi') = (t, \lambda)$ (Exercise 8.8.2), then $W \cong S^\xi V$.

Uniqueness of $\xi$ follows from Exercise 8.8.2.

**8.8.4** (Decomposing tensor products). Let $\lambda$ and $\mu$ be partitions with $\ell(\lambda), \ell(\mu) \leq d$. Show that $S^\lambda V \otimes S^\mu V$ is a completely reducible representation of $\text{GL}(V), gl(V)$ and $\text{sl}(V)$ and that all its irreducible constituents have the form $S^\nu V$ with partitions $\nu$ such that $\ell(\nu) \leq \dim V, |\nu| = |\lambda + \mu|$ and $\lambda' + \mu' - \nu' \in \bigoplus_{i=1}^{d-1} \mathbb{Z}_+ (e_i - e_{i+1})$. Furthermore, $S^{\lambda + \mu} V$ occurs exactly once. (Use Proposition 8.39 and Exercise 8.8.4.)
Part IV

Hopf Algebras
Chapter 9

Coalgebras, Bialgebras
and Hopf Algebras

Much of this chapter is of a rather more formal nature than the preceding parts of this
book; its main purpose is to set forth the basic generalities concerning coalgebras,
bialgebras and, most importantly, Hopf algebras. We have already encountered all
these structures, if not by name, in two special cases: group algebras (Section 3.3)
and enveloping algebras of Lie algebras (Section 5.4). Hopf algebras provide a
common framework for many of the general constructions that we have studied
separately for groups and for Lie algebras earlier. Moreover, as we shall see in
this chapter, coalgebras, bialgebras and Hopf algebras occur naturally in myriad
contexts other than groups and Lie algebras and are objects worthy of investigation
in their own right.

Throughout this chapter, we work over an arbitrary base field $\mathbb{k}$ unless otherwise
specified. As before, $\otimes$ stands for $\otimes_{\mathbb{k}}$ and, for any $V \in \text{Vect}_{\mathbb{k}}$, we let $V^\ast$ denote the
linear dual and $\langle \cdot, \cdot \rangle: V^\ast \times V \rightarrow \mathbb{k}$ is the evaluation pairing: $\langle f, v \rangle = f(v)$.

9.1. Coalgebras

In brief, the definitions of coalgebras and their homomorphisms are obtained by
“dualizing” the corresponding definitions for algebras: all arrows in the algebra
diagrams (1.1) and (1.2) have to be reversed for coalgebras. It turns out that, rather
than being a mere formal exercise, this process actually results in a highly useful
algebraic structure. This section furnishes some examples to illustrate this point
and it also discusses the passage between the category of coalgebras and the familiar
category of algebras.
9.1.1. The Category of \(k\)-Coalgebras

In detail, a \(k\)-coalgebra is a \(k\)-vector space \(C\) that is equipped with two \(k\)-linear maps, the comultiplication \(\Delta = \Delta_C : C \to C \otimes C\) and the counit \(\varepsilon = \varepsilon_C : C \to k\), such that the following diagrams commute:

\[
\begin{array}{ccc}
C & \xrightarrow{\Delta} & C \otimes C \\
\downarrow{\Delta} & & \downarrow{\Delta \otimes \text{Id}} \\
C \otimes C & \xrightarrow{\text{Id} \otimes \Delta} & C \otimes C \otimes C
\end{array}
\quad \text{and} \quad
\begin{array}{ccc}
C & \xrightarrow{\varepsilon \otimes \text{Id}} & C \otimes k \\
\downarrow{\varepsilon} & & \downarrow{\varepsilon \otimes \text{Id}} \\
k \otimes C & \xrightarrow{\text{Id} \otimes \varepsilon} & C \otimes C
\end{array}
\]

The diagram on the left expresses coassociativity of the comultiplication \(\Delta\); the second diagram states the counit laws. The coalgebra \(C\) is called cocommutative, if \(\Delta\) also satisfies the rule

\[
\Delta = \tau \circ \Delta,
\]

where \(\tau : C \otimes C \to C \otimes C\) is the switch map \(\tau(a \otimes b) = b \otimes a\).

Given \(k\)-coalgebras \(C\) and \(D\), a homomorphism (or coalgebra map) from \(C\) to \(D\) is a \(k\)-linear map \(f : C \to D\) such that the following diagrams commute:

\[
\begin{array}{ccc}
C & \xrightarrow{\Delta_C} & D \\
\downarrow{f} & & \downarrow{\Delta_D} \\
C \otimes C & \xrightarrow{f \otimes f} & D \otimes D
\end{array}
\quad \text{and} \quad
\begin{array}{ccc}
C & \xrightarrow{\varepsilon_C} & k \\
\downarrow{\varepsilon} & & \downarrow{\varepsilon_D} \\
C & \xrightarrow{f} & D
\end{array}
\]

We thus have introduced a new category,

\[\text{Coalg}_k.\]

Here, for the record, are a few more definitions, none of which is likely to come as a surprise to the reader. Let \(C\) be a \(k\)-coalgebra and let \(D \subseteq C\) be a \(k\)-subspace. If the restriction \(\Delta|_D\) has image in \(D \otimes D\), viewed as a subspace of \(C \otimes C\) in the usual fashion, then \(D\) is called a subcoalgebra of \(C\). The restrictions \(\Delta_D = \Delta|_D\) and \(\varepsilon_D = \varepsilon|_D\) serve as comultiplication and counit for \(D\) in this case; so \(D \in \text{Coalg}_k\) and the inclusion \(D \hookrightarrow C\) is a map in \(\text{Coalg}_k\). The subspace \(D\) is called a coideal of \(C\) if \(\langle \epsilon, D \rangle = 0\) and \(\Delta|_D\) has image in \(D \otimes C + C \otimes D\), the kernel of the canonical epimorphism \(C \otimes C \to (C/D) \otimes (C/D)\) in \(\text{Vect}_k\). Then \(\varepsilon\) and \(\Delta\) pass down to to the quotient space \(C/D\), making \(C/D\) a \(k\)-coalgebra and the canonical epimorphism \(C \to C/D\) a map of coalgebras. We leave it to the care of the reader to verify that coideals of \(C\) are exactly the kernels (in \(\text{Vect}_k\)) of coalgebra maps \(f : C \to D\) and to formulate the coalgebra versions of the standard isomorphism theorems; see Sweedler [194, Theorem 1.4.7] for example.
Finally, keeping the counit $\varepsilon$ but replacing the comultiplication $\Delta$ of $C$ by $\Delta^{\text{cop}} := \tau \circ \Delta$ with $\tau$ as in (9.2), we obtain the **coopposite coalgebra**, $C^{\text{cop}}$. If $C$ is cocommutative, then $C^{\text{cop}} \cong C$.

### 9.1.2. Initial Examples

We have seen special incarnations of the diagrams (9.1) as well as the cocommutativity rule (9.2) before: for group algebras in (3.26), (3.27) and for enveloping algebras of Lie algebras in (5.22), (5.23). In this subsection, we explain the coalgebra structure of enveloping algebras in more detail. We also present a class of coalgebras that is modeled on group algebras but may be lacking the algebra structure maps, multiplication and unit.

#### Group-like Coalgebras

Let $C \in \text{Coalg}_k$. An element $c \in C$ is called **group-like** if $\Delta c = c \otimes c$ and $\langle \varepsilon, c \rangle = 1$. This is exactly how $\Delta$ and $\varepsilon$ have been defined on the $k$-basis $G$ of the group algebra $kG$ (Section 3.3). Note also that if $\Delta c = c \otimes c$ and $c \neq 0$, then $\langle \varepsilon, c \rangle c = c$ by the counit laws (9.1), and hence $\langle \varepsilon, c \rangle = 1$. Denoting the set of all group-like elements of $C$ by $G_C$, we obtain a functor $G: \text{Coalg}_k \to \text{Sets}$. Thus,

$$G_C \overset{\text{def}}{=} \{ c \in C \mid \Delta c = c \otimes c \text{ and } \langle \varepsilon, c \rangle = 1 \}$$

**Lemma 9.1.** For any $C \in \text{Coalg}_k$, distinct elements of $G_C$ are $k$-linearly independent.

**Proof.** Suppose, for a contradiction, that there is a nontrivial $k$-linear relation among elements of $G_C$. Then there is a relation of the form

$$g = \lambda_1 g_1 + \cdots + \lambda_n g_n,$$

with distinct $g, g_i \in G_C$ and $0 \neq \lambda_i \in k$. Assume that $n$ is chosen minimal. Then $g_1, g_2, \ldots, g_n$ are $k$-linearly independent and $n \geq 2$, because $g = \lambda_1 g_1$ would result in $1 = \langle \varepsilon, g \rangle = \lambda_1 \langle \varepsilon, g_1 \rangle = \lambda_1$, contradicting our assumption $g \neq g_1$. Applying $\Delta$ to the displayed relation gives

$$\sum_{i,j} \lambda_i \lambda_j g_i \otimes g_j = g \otimes g = \Delta g = \sum_i \lambda_i \Delta g_i = \sum_i \lambda_i g_i \otimes g_i.$$

Since $(g_i \otimes g_j)_{i,j}$ is a $k$-linearly independent family in $C \otimes C$, we must have $\lambda_i \lambda_j = 0$ for $i \neq j$, contradicting the fact that all $\lambda_i$ are nonzero. \qed

**Example 9.2** (The group-like coalgebra of a set). Given a set $X$, consider the vector space $kX$ is of all formal $k$-linear combinations of the elements of $X$ (Example A.5). We can make $kX$ into a $k$-coalgebra by decreeing all $x \in X$ to be group-like and extending the maps $\Delta$ and $\varepsilon$ from $X$ to all of $kX$ by $k$-linearity. Lemma 9.1 implies that $G(kX) = X$. Any subset $Y \subseteq X$ gives rise to the subcoalgebra $kY$ of
In fact, it is not hard to see that all subcoalgebras of $\kappa X$ arise in this way and that the foregoing gives a functor $\kappa : \text{Sets} \to \text{Coalg}_k$ that is left adjoint to the functor $G : \text{Coalg}_k \to \text{Sets}$ (Exercise 9.1.2).

**Enveloping Algebras**

Next, let us turn to the enveloping algebra $U\mathfrak{g}$ of a Lie algebra $\mathfrak{g} \in \text{Lie}_k$. We have observed in (5.21) that the comultiplication $\Delta : U\mathfrak{g} \to U\mathfrak{g} \otimes U\mathfrak{g}$ is in fact an algebra map, but the formulahas been explicitly stated only for the elements of the subspace $\mathfrak{g} \subseteq U\mathfrak{g}$:

$$\Delta x = x \otimes 1 + 1 \otimes x \quad (x \in \mathfrak{g}).$$

The comultiplication takes a more complicated form for elements of $U\mathfrak{g}$ not belonging to $\mathfrak{g}$. The next example details this.

**Example 9.3** (Enveloping algebras as coalgebras). Let $U = U\mathfrak{g}$ be the enveloping algebra of a Lie $\kappa$-algebra $\mathfrak{g}$ and assume that $\text{char} \kappa = 0$. Fix a $\kappa$-basis $(x_i)_{i \in I}$ of $\mathfrak{g}$ and choose a total ordering $\leq$ of $I$. By the Poincaré-Birkhoff-Witt Theorem ($\S$5.4.3), we know that a $\kappa$-basis of $U$ is given by the standard monomials $x_{i_1}x_{i_2}\cdots x_{i_n}$ with $i_1 \leq i_2 \leq \cdots \leq i_n$ ($n \geq 0$). In order to give an explicit formula for the comultiplication $\Delta$ of $U$, it will be convenient to streamline notation. Let $\mathbb{Z}^+_{<}(I)$ denote the (additive) monoid consisting of all functions $n : I \to \mathbb{Z}^+$ such that $n(i) = 0$ for almost all $i \in I$ (Example 3.3). Each standard monomials may be written in the form $x_{i_1}^{n(i_1)}x_{i_2}^{n(i_2)}\cdots x_{i_n}^{n(i_n)}$ for a unique $n \in \mathbb{Z}^+_{<}(I)$, where the superscript $<$ indicates that the $i$ occur in strictly increasing order. Now consider the following renormalization of this standard monomial:

$$x_n := \prod_{i \in I} x_{i}^{n(i)} / n(i)! \in U \quad (n \in \mathbb{Z}^+_{<}(I)).$$

The elements $x_n$ with $n \in \mathbb{Z}^+_{<}(I)$ also form a $\kappa$-basis of $U$. For this basis, the comultiplication of $U$ is given by the rule

$$\Delta(x_n) = \sum_{r+s=n} x_r \otimes x_s.$$

The verification of this formula is not difficult using the fact that $\Delta$ is an algebra map (Exercise 9.1.3). For the counit, recall that $\varepsilon$ is also an algebra map and $\langle \varepsilon, x \rangle = 0$ for all $x \in X$ by (5.20). Thus, $\langle \varepsilon, x_n \rangle = 0$ for $n \neq 0$ and $\langle \varepsilon, x_0 \rangle = \langle \varepsilon, 1 \rangle = 1$.

**9.1.3. Sweedler and Graphical Notation**

While the usual juxtaposition notation for the multiplication of an algebra is both simple and unproblematic, the comultiplication of a coalgebra $C$ requires a more elaborate notational scheme. Indeed, the expression $\Delta c = \sum_i c_{i,1} \otimes c_{i,2} \in C \otimes C$ for $c \in C$ is cumbersome and the components $c_{i,1}, c_{i,2} \in C$ are not unique unless further restrictions are imposed. Therefore, it is customary to symbolically write
\( \Delta c \) in the abbreviated form \( \Delta c = \sum c_{(1)} \otimes c_{(2)} \). This is known as the **Sweedler notation**. A variant of this notation also dispenses with the summation symbol:

\[
\Delta c = c_{(1)} \otimes c_{(2)} .
\]

Using the sumless version of the Sweedler notation, which we shall generally employ, the counit laws in (9.1) can be stated as the identities

\[
\langle \varepsilon, c_{(1)} \rangle c_{(2)} = c = c_{(1)} \langle \varepsilon, c_{(2)} \rangle \quad (c \in C).
\]

At this point, the reader is invited to revisit Sections 3.3 and 5.4.4, where the special cases \( C = kG \) and \( C = Ug \) are discussed, and write out this identity for \( c = g \in G \) and for \( c = x \in g \).

Coassociativity of \( \Delta \) allows for a convenient iteration of the Sweedler notation. Indeed, the coassociativity identity

\[
((\Delta \otimes \text{Id}) \circ \Delta) c = ((\text{Id} \otimes \Delta) \circ \Delta) c
\]

for \( c \in C \) can be stated as follows:

\[
(9.4) \quad c_{(1)(1)} \otimes c_{(1)(2)} \otimes c_{(2)} = c_{(1)} \otimes c_{(2)(1)} \otimes c_{(2)(2)} =: c_{(1)} \otimes c_{(2)} \otimes c_{(3)} .
\]

More generally, defining \( \Delta_n : C \to C^{\otimes n} \) inductively by \( \Delta_2 = \Delta \) and \( \Delta_{n+1} = (\Delta_n \otimes \text{Id}) \circ \Delta \), the resulting map \( \Delta_n \) is identical to the composite of any \( n-1 \) maps \( C \to \cdots \to C^{\otimes m} \to C^{\otimes (m+1)} \to \cdots \to C^{\otimes n} \) that apply \( \Delta \) to one factor of \( C^{\otimes m} \) and the identity to all others. We will write this map as

\[
\Delta_n c = c_{(1)} \otimes \cdots \otimes c_{(n)} .
\]

For example, if \( g \in GC \), then \( \Delta_n g = g \otimes g \otimes \cdots \otimes g \). If \( C = Ug \) and \( x \in g \), then \( \Delta_n x = \sum_{i=1}^n x_i \) with \( x_i = 1 \otimes \cdots \otimes x \otimes \cdots \otimes 1 \).

**Cobordism Diagrams**

Economical as the sumless Sweedler notation may be, it does require some practice to be able to use it with confidence and formulas tend to become non-intuitive and difficult to parse. It is therefore common practice in category theory and elsewhere to use various graphical representations of linear maps to explain their relations to each other in a more transparent manner. One such method are the cobordism diagrams employed in topological quantum field theory. Here, a linear map \( C^{\otimes m} \to C^{\otimes n} \) is depicted by a cobordism between \( m \) incoming boundary circles and \( n \) outgoing boundary circles. The input circles are arranged vertically and thought of as labeled from top to bottom by 1, 2, \ldots, \( m \); likewise for the output circles. Thus, the identity map \( \text{Id}_C \), the comultiplication \( \Delta : C \to C \otimes C \) and the counit \( \varepsilon : C \to \mathbb{1} = C^{\otimes 0} \) are represented by a cylinder, a “pair of pants” and a cup, respectively:

\[
\text{Id} = \begin{array}{c}
\hline
\end{array}
\Delta = \begin{array}{c}
\hline
\end{array}
\varepsilon = \begin{array}{c}
\cdots
\end{array}
\]

---

\(^{1}\)The Sweedler notation is also referred to as the Sweedler-Heyneman notation. It is often further streamlined by omitting the parentheses in the subscripts, but we shall not do so in this book.
The graphical rendition of the coassociativity identity (9.4) then looks like this:

And here are the counit laws \((\text{Id} \otimes \varepsilon) \circ \Delta = \text{Id} = (\varepsilon \otimes \text{Id}) \circ \Delta\) as cobordism diagrams:

**Example 9.4** (Coproduct for Frobenius algebras). Let \(A\) be a Frobenius \(k\)-algebra with Frobenius form \(\lambda \in A^\ast\) (Section 2.2). We will show here that \(A\) can be equipped with a comultiplication \(\delta : A \to A \otimes A\) such that \((A, \delta, \lambda)\) is a \(k\)-coalgebra. Thus, \(\lambda\) will play the role of the counit, represented by the cup diagram as above. We will also represent the multiplication \(m : A \otimes A \to A\) by a cobordism diagram:

Then \(\beta = \lambda \circ m : A \otimes A \to k\) gives a (non-degenerate) bilinear form on \(A\):

**Example 9.5** (Coproduct for Frobenius algebras). Let \(A\) be a Frobenius \(k\)-algebra with Frobenius form \(\lambda \in A^\ast\) (Section 2.2). We will show here that \(A\) can be equipped with a comultiplication \(\delta : A \to A \otimes A\) such that \((A, \delta, \lambda)\) is a \(k\)-coalgebra. Thus, \(\lambda\) will play the role of the counit, represented by the cup diagram as above. We will also represent the multiplication \(m : A \otimes A \to A\) by a cobordism diagram:

\[
\beta = \lambda \circ m : A \otimes A \to k \tag{9.5}
\]

Associativity of \(m\) implies that the form \(\beta\) is associative: \(\beta(a \otimes bc) = \beta(ab \otimes c)\) for \(a, b, c \in A\); so we may define \(\alpha := \beta \circ (m \otimes \text{Id}) = \beta \circ (\text{Id} \otimes m) : A \otimes^3 A \to k\) which looks like this:

There is a unique coform \(\gamma : k \to A \otimes A\) so that the following diagram commutes:

\[
\begin{align*}
A \otimes k & \xleftarrow{\text{Id} \otimes \gamma} A \xrightarrow{\sim} k \otimes A \\
A \otimes A \otimes A & \xrightarrow{\beta \otimes \text{Id}} A \otimes k \xleftarrow{\beta} A \otimes A \otimes A \\
A \otimes k & \xrightarrow{\gamma \otimes \text{Id}} A \otimes A \otimes A \xrightarrow{\gamma} k \otimes A
\end{align*}
\]

In the notation of (2.27), the coform \(\gamma\) is given by \(\gamma(1) = y_i \otimes x_i = c_{\lambda\gamma}\). Commutativity of the left and right halves of (9.7) are equivalent to (2.20) and (2.22), respectively. With
Commutativity of the diagram (9.7) translates into the “snake relations”:

\[
\gamma = \begin{array}{c}
\text{commutativity of the diagram (9.7) translates into the “snake relations”:}
\end{array}
\]

\[
(9.8)
\]

Define the comultiplication $\delta$ of $A$ by

\[
\delta : A \xrightarrow{\gamma \otimes \text{Id} \otimes \text{Id}} A \otimes 5 \xrightarrow{\text{Id} \otimes \alpha \otimes \text{Id}} A \otimes A \xrightarrow{\gamma \otimes \text{Id} \otimes \text{Id}} A \otimes A
\]

or, rendered graphically,

\[
(9.9)
\]

Redrawing $\alpha$ and using the second snake relation, we obtain

\[
(9.10)
\]

The graphical verification of the counit law $(\text{Id} \otimes \lambda) \circ \delta = \text{Id}$ is now straightforward using the first snake relation:

\[
(9.11)
\]
To summarize, the linear maps that we have assembled here satisfy

- \((A, m, u)\) is a \(\mathbb{k}\)-algebra;
- \((A, \delta, \lambda)\) is a \(\mathbb{k}\)-coalgebra;
- \((\text{Id}_A \otimes m) \circ (\delta \otimes \text{Id}_A) = \delta \circ m = (m \otimes \text{Id}_A) \circ (\text{Id}_A \otimes \delta)\).

The pair of identities above follows from (9.10) and (9.11); the verification is left to the reader (Exercise 9.1.4).

9.1.4. Convolution

Despite the close formal parallels to algebras, coalgebras do in fact allow for a construction that cannot naively be dualized to yield an analogous construction for algebras. In detail, let \(C = (C, \Delta, \varepsilon) \in \text{Coalg}_\mathbb{k}\) and \(A = (A, m, u) \in \text{Alg}_\mathbb{k}\) be given. Then the \(\mathbb{k}\)-vector space \(\text{Hom}_\mathbb{k}(C, A)\) carries the **convolution** product \(*\), which is defined by \(f \ast g = m \circ (f \otimes g) \circ \Delta\) for \(f, g \in \text{Hom}_\mathbb{k}(C, A)\) or, in elementwise form,

\[
(f \ast g)c = f(c_{(1)})g(c_{(2)}) \quad (c \in C).
\]

Coassociativity of \(\Delta\) and associativity of \(m\) together imply associativity of the convolution product: \(((f \ast g) \ast h)c = f(c_{(1)})g(c_{(2)})h(c_{(3)}) = (f \ast (g \ast h))c\). Similarly, it is straightforward to verify that \(u \circ \varepsilon : C \to \mathbb{k} \to A\) serves as a two-sided identity element for \(*\). Thus, \(\text{Hom}_\mathbb{k}(C, A)\) is a \(\mathbb{k}\)-algebra.

**The Dual Algebra of a Coalgebra.** The foregoing in particular yields a \(\mathbb{k}\)-algebra structure on the linear dual, \(C^* = \text{Hom}_\mathbb{k}(C, \mathbb{k})\). The unit map of \(C^*\) is the transpose of the augmentation, \(\varepsilon^* : \mathbb{k} \equiv \mathbb{k} \to C^*\), and multiplication is given by pulling back \(\Delta^* : (C \otimes C)^* \to C^*\) along the canonical embedding \(C^* \otimes C^* \to (C \otimes C)^*\). We will frequently consider dual algebras \(C^*\) for \(C \in \text{Coalg}_\mathbb{k}\) and generally denote the convolution multiplication in \(C^*\) by simple juxtaposition: \(fg = f \ast g\) for \(f, g \in C^*\).

**Structure of Convolution Algebras.** Returning to the case of a general convolution algebra, \(\text{Hom}_\mathbb{k}(C, A)\), consider the standard \(\mathbb{k}\)-linear embedding

\[
A \otimes C^* \hookrightarrow \text{Hom}_\mathbb{k}(C, A)
\]

\[
w \downarrow \downarrow
\]

\[
a \otimes f \longmapsto (c \mapsto a(f, c))
\]

This is in fact an algebra map. Indeed, identifying \(A \otimes C^*\) with its image in \(\text{Hom}_\mathbb{k}(C, A)\), it is routine to check that

\[
(a \otimes f) \ast (b \otimes g) = ab \otimes fg \quad (a, b \in A, f, g \in C^*).
\]

Thus, we may view \(A \otimes C^*\) as a subalgebra of \(\text{Hom}_\mathbb{k}(C, A)\), consisting of the elements of \(\text{Hom}_\mathbb{k}(C, A)\) that have finite rank. In particular, if \(C\) or \(A\) are finite-dimensional, then \(\text{Hom}_\mathbb{k}(C, A) \equiv A \otimes C^*\) as \(\mathbb{k}\)-algebras. In general, assuming \(A \neq 0\) and \(C \neq 0\),
the standard embeddings $A \hookrightarrow A \otimes C^*$ and $C^* \hookrightarrow A \otimes C^*$ composed with the embedding $A \otimes C^* \hookrightarrow \text{Hom}_\mathbb{k}(C, A)$ give algebra embeddings

\[ \varepsilon^*: \quad A \xrightarrow{\sim} \text{Hom}_\mathbb{k}(\mathbb{k}, A) \xleftarrow{\varepsilon} \text{Hom}_\mathbb{k}(C, A) \]

(9.12)

\[ a \quad \xrightarrow{\sim} \quad (c \mapsto a(\varepsilon, c)) \]

and

\[ u^*: \quad C^* = \text{Hom}_\mathbb{k}(C, \mathbb{k}) \xleftarrow{u} \text{Hom}_\mathbb{k}(C, A) \]

(9.13)

\[ f \quad \xrightarrow{\sim} \quad (c \mapsto \langle f, c \rangle 1) \]

**Example 9.5** (Convolution algebras of group-like coalgebras). Let $C = \mathbb{k}X$ be the group-like coalgebra of the set $X$ over $\mathbb{k}$ (Example 9.2). Then

\[ \text{Hom}_\mathbb{k}(\mathbb{k}X, A) \cong \text{Hom}_{\text{Sets}}(X, A|_{\text{Sets}}) \cong \prod_{x \in X} A, \]

with $f \in \text{Hom}(\mathbb{k}X, A)$ corresponding to the $X$-tuple $(f(x))_{x \in X}$. The convolution product and addition in $\text{Hom}(\mathbb{k}X, A)$ translate into the usual pointwise algebra operations in $\prod_{x \in X} A$: $(f * g)(x) = f(x)g(x)$ and similarly for $+$. In terms of the above isomorphism, (9.12) becomes the embedding of $A$ as the diagonal of $\prod_{x \in X} A$, that is, the constant functions $X \to A$.

**Example 9.6** (Convolution algebras of enveloping algebras). Now take the enveloping algebra $U = U_\mathbb{g}$ of $\mathbb{g} \in \mathfrak{lie}_\mathbb{k}$ for $C$. Assuming $\text{char} \mathbb{k} = 0$ as in Example 9.3, fix an ordered $\mathbb{k}$-basis $(x_i)_{i \in I}$ of $\mathbb{g}$ and consider the $\mathbb{k}$-basis of $U$ that is given by the renormalized standard monomials $x_n$ with $n \in \mathbb{Z}_+^{(I)}$. Let $A[t_i | i \in I]$ be the algebra of formal power series in the commuting variables $t_i$ over $A$ (Example 3.3). Using the description of $\Delta$ and $\varepsilon$ in Example 9.3, it is straightforward to check that the following map is an isomorphism in $\text{Alg}_\mathbb{k}$:

\[ \phi: \quad \text{Hom}_\mathbb{k}(U, A) \xrightarrow{\sim} A[t_i | i \in I] \]

\[ f \quad \xrightarrow{\sim} \quad \sum_n f(x_n)t^n \]

Here, the sum runs over all $n \in \mathbb{Z}_+^{(I)}$ and we have put $t^n = \prod_{i \in I} t_i^{n(i)}$. Under the isomorphism $\phi$, (9.12) becomes the embedding of $A$ as the constant power series in $A[t_i | i \in I]$. The unit map of $U$ yields a splitting of this embedding:

\[ u^*: \quad \text{Hom}_\mathbb{k}(U, A) \to A, \quad f \mapsto f(1), \]

which translates into the map $A[t_i | i \in I] \to A$, $\sum_n a_n t^n \mapsto a_0$. 

9.1.5. From Coalgebras to Algebras and Back

The construction of the dual algebra of a coalgebra (§9.1.4) is functorial: for any map $f : C \to D$ in $\text{Coalg}_k$, the transpose map $f^* : D^* \to C^*$ in $\text{Vec}_k$ is in fact easily seen to be a map in $\text{Alg}_k$ (Exercise 9.1.6). Thus, the linear dual yields a contravariant functor,

$$\cdot^* : \text{Coalg}_k \to \text{Alg}_k.$$

On the other hand, starting with $A = (A, m, u) \in \text{Alg}_k$, things are not quite so simple, because the multiplication transpose $m^* : A^* \to (A \otimes A)^*$ need not have image in the subspace $A^* \otimes A^* \subseteq (A \otimes A)^*$. So $A^*$ does not generally become a coalgebra via the ordinary linear dual. Of course, if $A$ is finite dimensional, then $A^* \otimes A^* \cong (A \otimes A)^*$ and $A^*$ is indeed a coalgebra with comultiplication $\Delta = m^*$ and counit $\epsilon = u^*$: commutativity of the coassociativity and counit diagrams (9.1) for $A$ follows immediately by transposing the corresponding diagrams (1.1) for $A$ (Exercise 9.1.6). For a general algebra $A$, however, the linear dual $A^*$ needs to be replaced by the so-called \textit{finite dual},\(^2\) which was already alluded to in §1.5.2:

$$A^o := \{ a^* \in A^* \mid a^* \text{ vanishes on some cofinite ideal of } A \}.$$

Recall that an ideal $I$ of $A$ is said to be cofinite if $\dim_k A/I < \infty$. We can write $A^o$ as a union of $k$-subspaces of $A^*$ that are indexed by the set of cofinite ideals $I$ of $A$:

$$A^o = \bigcup_I A_I^* \quad \text{with} \quad A_I^* = \{ a^* \in A^* \mid \langle a^*, I \rangle = 0 \}.$$

This union is directed in the sense that any two subspaces $A_I^*$ and $A_J^*$ are contained in a common third one, namely $A_{I \cap J}^*$. The subspace $A_I^*$ is canonically isomorphic to $(A/I)^*$, which, as we have seen above, carries a coalgebra structure that is obtained by dualizing the algebra structure of $A/I$, because $A/I$ is finite dimensional. Hence, by identification with $(A/I)^*$, each $A_I^*$ becomes a (finite-dimensional) coalgebra. The following proposition shows that these coalgebra structures can be assembled to make $A^o$ a coalgebra. In fact, $A^o$ is the largest $k$-subspace of $A^*$ that becomes a coalgebra with comultiplication given by $m^*$ (Exercise 9.1.8).

**Proposition 9.7.** (a) Let $A = (A, m, u) \in \text{Alg}_k$. Then $A^o$ is a $k$-coalgebra with comultiplication $\Delta = m^* |_{A^o}$ and counit $\epsilon = u^* |_{A^o}$.

(b) If $f : A \to B$ is a map in $\text{Alg}_k$, then $f^o = f^* |_{B^o}$ is a map $f^o : B^o \to A^o$ in $\text{Coalg}_k$.

**Proof.** (a) The first issue is to show that $m^*(A^o) \subseteq A^o \otimes A^o$, regarding $A^o \otimes A^o \subseteq A^* \otimes A^* \subseteq (A \otimes A)^*$. Inasmuch as $A$ is the directed union of the various subspaces $A_I^*$ for cofinite ideals $I$ of $A$, it suffices to show that $m^*(a^*) \in A_I^* \otimes A_J^*$ for each $a^* \in A_I^*$.

---

\(^2\) $A^o$ is also known as the \textit{Sweedler dual} of $A$.\footnote{A^o is also known as the \textit{Sweedler dual} of $A$.}
9.1. Coalgebras

But \( m^*(a^*) \in (A \otimes A)^* \) vanishes on the (cofinite) ideal \( I^{(2)} := (I \otimes A) + (A \otimes I) \) of \( A \otimes A \): for any \( x \in I^{(2)} \), \( \langle m^*(a^*), x \rangle = \langle a^*, m(x) \rangle \in \langle a^*, I \rangle = 0 \). Thus, \( m^*(a^*) \in (A \otimes A)^* \subseteq (A \otimes A)^* \). Furthermore, as subspaces of \( (A \otimes A)^* \), the following coincide:

\[
\begin{align*}
(A \otimes A)^*_{I^{(2)}} &= ((A \otimes A)/I^{(2)})^* = ((A/I) \otimes (A/I))^* \\
&= (A/I)^* \otimes (A/I)^* = A_I^* \otimes A_I^*.
\end{align*}
\]

Therefore, \( m^*(A_I^*) \subseteq A_I^* \otimes A_I^* \) and we obtain a well-defined linear map \( \Delta := m^*|_{A_I^*} : A^*_I \to A^* \otimes A^* \). Trivially, we also have a map \( \varepsilon := u^*|_{A^*} : A^* \to k \). It remains to check commutativity of the diagrams (9.1) for \( A^* \), which amounts to chasing an arbitrary element \( a^* \in A^* \) through the diagram in question. If \( a^* \in A_I^* \), say, then commutativity reduces to commutativity of the corresponding diagram for \( A_I^* \), and this in turn is a consequence of commutativity of the corresponding diagram for \( (A/I)^* \), which we have already settled above. This proves (a).

(b) Let \( b^* \in B^* \), say \( b^* \in B_J^* \) for a cofinite ideal \( J \) of \( B \). The preimage \( I = f^{-1}(J) \) is a cofinite ideal of \( A \) and \( \langle f^*(b^*), I \rangle = \langle b^*, f(I) \rangle \subseteq \langle b^*, J \rangle = 0 \). Therefore, \( f^*(b^*) \in A_I^* \subseteq A^* \) and we have shown that \( f^* = f^*|_{B^*} \) maps \( B^* \) to \( A^* \). Commutativity of the diagrams (9.3) is checked, exactly as we did with (9.1) in the proof of (a), by reduction to the case of finite-dimensional algebras, which is easily handled by transposing the algebra diagrams (1.2). We leave the details to the reader (Exercise 9.1.6).

By virtue of Proposition 9.7, we now also have a contravariant functor

\[
\cdot^* : \text{Alg}_k \to \text{Coalg}_k.
\]

The functors \( \cdot^* \) and \( \cdot^\circ \) are mutually adjoint on the right:

**Theorem 9.8.** For any \( A \in \text{Alg}_k \) and \( C \in \text{Coalg}_k \), there is a natural bijection of sets \( \text{Hom}_{\text{Alg}_k}(A, C^* \otimes k) \cong \text{Hom}_{\text{Coalg}_k}(C, A^* \otimes k) \).

**Proof.** We first show that the image of the canonical embedding of \( k \)-vector spaces \( C \hookrightarrow C^{**} \) is in fact contained in \( C^{* \otimes} \). To see this, we view \( C \in \text{Rep} C^* \) via the left action \( \cdot \) of the convolution algebra \( C^* \) on \( C \subseteq C^{**} \) that was already considered in (2.19) in the setting of algebras: \( \langle d^* \cdot c^* \cdot c, c \rangle = \langle d^*, c^*(c)_1 \rangle \langle c^*, (c^* \cdot c) \rangle \) for \( c \in C \) and \( c^*, d^* \in C^* \). This can be written more compactly as

\[
(9.14) \quad c^* \cdot c = (c^* \cdot c)_1 (c^*, c^* \cdot c) = (c^* \cdot c)_1 (c^* \cdot c)_2.
\]

For each \( c \in C \), the subspace \( C^* \cdot c \subseteq C \) is contained in the \( k \)-linear span of the components \( c^*(c)_1 \), and hence \( C^* \cdot c \) is finite dimensional. In other words, \( C \in \text{Rep} C^* \) is **locally finite**. Hence, the kernel of each subrepresentation \( C^* \cdot c \) is a cofinite ideal of \( C^* \), say \( I_c \), satisfying \( \langle I_c, c \rangle = \langle I_c, c^* \cdot c \rangle = \langle 1, I_c \cdot c \rangle = \langle 1, 0 \rangle = 0 \). This shows that \( c \in C^{* \otimes} \), providing us with the desired embedding

\[
(9.15) \quad C \hookrightarrow C^{* \otimes}.
\]
It is straightforward to check that this embedding is in fact a coalgebra map (Exercise 9.1.9). Therefore, for a given \( f \in \text{Hom}_{\text{Alg}}(A, C^\circ) \), we may define the coalgebra map \( \Psi(f) : C \to A^\circ \) to be the composite of the embedding \( C \hookrightarrow C^{\ast\ast} \) followed by the coalgebra map \( f^\circ : C^{\ast\ast} \to A^\circ \) (Proposition 9.7).

For the opposite direction, we start with \( A \). The dual of the inclusion \( A^\circ \hookrightarrow A^* \) is an epimorphism \( A^{**} \to A^* \) in \( \text{Vect}_k \). Composition with the canonical embedding of \( A \hookrightarrow A^{**} \) gives a \( k \)-linear map
\[
\alpha : A \to A^{**}.
\]
Again, it is easily verified that \( \alpha \) is an algebra map (which need not be injective; Exercise 9.1.9). Thus, for any \( g \in \text{Hom}_{\text{ComAlg}}(C, A^\circ) \), we may define \( \Phi(g) = g^\circ \circ \alpha \in \text{Hom}_{\text{Alg}}(A, C^\circ) \). We leave to the reader the task of checking that the maps \( \Psi \) and \( \Phi \) are inverse to each other and satisfy the requisite naturality condition for adjoint functors (Exercise 9.1.9).

**Example 9.9** (\( \text{End}_k(V)^\circ \)). Let us describe the coalgebra \( \text{End}_k(V)^\circ \) for a given \( V \in \text{Vect}_k \). If \( V \) is infinite dimensional, then \( \text{End}_k(V)^\circ = 0 \) (Exercise 9.1.10); so we assume that \( \dim_k V < \infty \). Then \( \text{End}_k(V)^\circ = \text{End}_k(V)^* \) and the standard isomorphisms \( \text{End}_k(V) \cong V \otimes V^* \) and \( V \cong V^{**} \) give rise to \( \text{Vect}_k \)-isomorphisms \( \text{End}_k(V)^* \cong \text{End}_k(V) \cong V \otimes V^* \). Identifying \( \text{End}_k(V) \) and \( \text{End}_k(V)^* \) with \( V \otimes V^* \), the evaluation pairing \( \langle \cdot, \cdot \rangle : \text{End}_k(V)^* \times \text{End}_k(V) \to k \) takes the form
\[
\langle v \otimes f, w \otimes g \rangle = \langle f, w \rangle \langle g, v \rangle \quad (v, w \in V, f, g \in V^*).
\]
Furthermore, the counit \( \varepsilon : \text{End}_k(V)^* \to k, \phi \mapsto \phi(\text{Id}_V) \), becomes the ordinary trace map (B.23):
\[
\langle \varepsilon, v \otimes f \rangle = \langle f, v \rangle.
\]
Indeed, \( \text{Id}_V \in \text{End}_k(V) \) becomes \( \sum_i x_i \otimes x^i \in V \otimes V^* \) with dual bases \( (x_i, x^i) \). Thus,
\[
\langle \varepsilon, v \otimes f \rangle = \langle v \otimes f, \sum_i x_i \otimes x^i \rangle = \sum_i \langle f, x_i \rangle \langle x^i, v \rangle = \langle f, v \rangle \quad \text{as claimed. Next, we show that comultiplication } \Delta \text{ of } \text{End}_k(V)^* \text{ takes the following form in } V \otimes V^*:
\]
\[
\Delta(v \otimes f) = \sum_i x_i \otimes f \otimes v \otimes x^i.
\]
To see this, recall that \( \Delta \) is the transpose of the multiplication of the algebra \( \text{End}_k(V) \), which in \( V \otimes V^* \) becomes \( (v \otimes f) \circ (w \otimes g) = v(f, w) \otimes g \). Thus,
\[
\langle v \otimes f, (v' \otimes f') \circ (v'' \otimes f'') \rangle = \langle v \otimes f, v'(f', v'') \otimes f'' \rangle = \langle f, v' \rangle \langle f', v'' \rangle \langle f'', v \rangle = \sum_i \langle f, x_i \rangle \langle f', x^i \rangle \langle v'', v \rangle = \sum_i \langle x_i \otimes f, v' \otimes f' \rangle \langle v \otimes x^i, v'' \otimes f'' \rangle,
\]
proving the asserted formula for the comultiplication.
Exercises for Section 9.1

9.1.1 (Sums and intersections of subcoalgebras etc.). Let \( C \in \text{Coalg}_\mathbb{K} \). Show that the sum and the intersection of any collection of subcoalgebras of \( C \) is again a subcoalgebra. Similarly for left and right coideals of \( C \); see Example 9.10 for the latter two notions. Finally, show that the sum of any collection of coideals of \( C \) is a coideal, but this generally fails for intersections of coideals.

9.1.2 (Group-likes and sets). (a) Show that the functor \( \mathbb{K} \cdot : \text{Sets} \to \text{Coalg}_\mathbb{K} \) (Example 9.2) is left adjoint to the functor \( G : \text{Coalg}_\mathbb{K} \to \text{Sets} \). Conclude that, for any two \( X, Y \in \text{Sets} \), there is a natural bijection \( \text{Hom}_{\text{Sets}}(X, Y) \cong \text{Hom}_{\text{Coalg}_\mathbb{K}}(\mathbb{K}X, \mathbb{K}Y) \).

(b) For any \( X \in \text{Sets} \), show that the subcoalgebras of \( \mathbb{K}X \) are exactly the subspaces of the form \( \mathbb{K}Y \) for subsets \( Y \subseteq X \) and the coideals of \( \mathbb{K}X \) are the subspaces of the form \( \langle x - x' | x \sim x' \rangle_\mathbb{K} \), where \( \sim \) is an equivalence relation on \( X \).

9.1.3 (Comultiplication in enveloping algebras). Using the fact that \( \Delta \) is an algebra map (5.21), prove the identity \( \Delta(x_a) = \sum_{r+s=n} x_r \otimes x_s \) for \( n \in \mathbb{Z}_+^\infty(X) \) in Example 9.3.

9.1.4 (Frobenius algebras). Using (9.10) and (9.11), verify the identities \( (\text{Id}_A \otimes m) \circ (\Delta \otimes \text{Id}_A) = \Delta \circ m = (m \otimes \text{Id}_A) \circ (\text{Id}_A \otimes \delta) \) in Example 9.4.

9.1.5 (Primes in convolution algebras). The algebra map \( \varepsilon^* : A \hookrightarrow B = \text{Hom}_\mathbb{K}(C, A) \) in (9.12) generally does not induce a map \( \text{Spec} B \to \text{Spec} A \), \( P \mapsto (\varepsilon^*)^{-1}(P) \), because this fails for the extension \( A \hookrightarrow A[[x]] \) (Example 9.6 and Exercise 1.3.5). Consider the following condition, which evidently holds for commutative algebras as well as for right noetherian algebras:

\[(*)_r \quad \text{Every finitely generated ideal of } A \text{ is finitely generated as right ideal.}\]

(a) Assume that \( A \) satisfies \((*)_r \). Show that \( B\varepsilon^*(a) \subseteq \varepsilon^*(Aa)B \) for all \( a \in A \). Conclude the existence of a map \( \text{Spec} B \to \text{Spec} A \), \( P \mapsto (\varepsilon^*)^{-1}(P) \).

(b) PI-algebras \( A \) need not satisfy \((*)_r \) as the following example shows:\(^3\) Given a field extension \( K/\mathbb{K} \), form the matrix algebra \( A = \begin{pmatrix} K & K \\ 0 & K \end{pmatrix} \). Show that \( I = \begin{pmatrix} 0 & K \\ 0 & 0 \end{pmatrix} \) is an ideal of \( A \) that is generated by the matrix \( \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \), even as left ideal, but \( I \) is only finitely generated as right ideal of \( A \) if the extension \( K/\mathbb{K} \) is finite.

9.1.6 (Duals of algebras and coalgebras). (a) Check that if \( f : C \to D \) is a map in \( \text{Coalg}_\mathbb{K} \), then the transpose map \( f^\circ : D^* \to C^* \) is a map in \( \text{Alg}_\mathbb{K} \).

(b) If \( A \in \text{Alg}_\mathbb{K} \) is finite dimensional, then \( A^* \otimes A^* \cong (A \otimes A)^* \) and \( A^* \) is a \( \mathbb{K} \)-coalgebra with comultiplication \( \Delta = m^* \) and counit \( \varepsilon = u^* \).

(c) Let \( f : A \to B \) be a homomorphism of finite-dimensional \( \mathbb{K} \)-algebras. Then the transpose \( f^* : B^* \to A^* \) is a map of \( \mathbb{K} \)-coalgebras.

\(^3\)The example was pointed out to me by Louis Rowen.
9.1.7 (Opposites and cooppesites). For given $C \in \text{Coalg}_k$ and $A \in \text{Alg}_k$, prove the following isomorphisms.

(a) $\text{Hom}_k(C^\text{cop}, A^\text{op}) \cong \text{Hom}_k(C, A)^\text{op}$ in $\text{Alg}_k$; in particular, $(C^\text{cop})^* \cong (C^*)^\text{op}$

(b) $(A^\text{op})^* \cong (A^*)^\text{cop}$ in $\text{Coalg}_k$.

9.1.8 (On the finite dual). For $A = (A, m, u) \in \text{Alg}_k$, consider the $(A, A)$-bimodule structure on $A^*$ that is given by the actions $\rho$ and $\psi$ (2.19): $a \cdot f \cdot b = f \circ b \circ_A a$ for $f \in A^*$ and $a, b \in A$. Show that the following are equivalent for $f \in A^*$:

(i) $f \in A^\circ$; (ii) $m^*(f) \in A^\circ \otimes A^\circ$; (iii) $m^*(f) \in A^* \otimes A^*$; (iv) $A \cdot f$ is finite dimensional; (v) $f \cdot A$ is finite-dimensional; and (vi) $A \cdot f \cdot A$ is finite dimensional.

Here, we view $A^* \otimes A^* \subseteq (A \otimes A)^*$ as usual. In particular, $A^\circ = (m^*)^{-1}(A^* \otimes A^*)$, and hence $A^\circ$ largest subspace of $A^*$ that is a coalgebra via $m^*$.

9.1.9 (Details for the proof of Theorem 9.8). Fill in the details left open in the proof: the embedding $C \hookrightarrow C^{\text{cop}}$ is a coalgebra map; $\alpha$ is an algebra map, which need not be injective; the maps $\Psi$ and $\Phi$ are inverse to each other and satisfy the requisite naturality condition for adjoint functors.

9.1.10 (Endomorphism algebras of infinite-dimensional vector spaces). Let $V \in \text{Vec}_k$ be infinite dimensional. Show that $\text{End}_k(V)^\circ = 0$.

9.1.11 (Tensor products of coalgebras). (a) Let $C, D \in \text{Coalg}_k$. Define $\Delta_{C \otimes D} = (2 \ 3) \circ (\Delta_C \otimes \Delta_D) : C \otimes D \rightarrow (C \otimes D) \otimes (C \otimes D)$, where (2 3) switches the factors in positions 2 and 3, and let $\varepsilon_{C \otimes D} = \varepsilon_C \otimes \varepsilon_D : C \otimes D \rightarrow k \otimes k \xrightarrow{\sim} k$. Show that $C \otimes D$ is a $k$-coalgebra with comultiplication $\Delta_{C \otimes D}$ and counit $\varepsilon_{C \otimes D}$.

(b) For $A, B \in \text{Alg}_k$, show that $A^\circ \otimes B^\circ \cong (A \otimes B)^\circ$ in $\text{Coalg}_k$.

(c) For $X, Y \in \text{Sets}$, show that $kX \otimes kY \cong k[X \times Y]$ in $\text{Coalg}_k$.

9.2. Comodules

We now turn to the variant of the standard notion of a module over an algebra that is appropriate for coalgebras: comodules. Unsurprisingly, the definition is obtained by reversing arrows in the module diagrams (1.17). However, we will change sides here and spell things out for right comodules at the outset, since they will turn out to be most directly related to representations of algebras, that is, left modules.

9.2.1. Comodules

Let $C = (C, \Delta, \varepsilon) \in \text{Coalg}_k$. A right comodule over $C$ is a $k$-vector space, $M$, together with a map $\delta = \delta_M : M \rightarrow M \otimes C$ in $\text{Vec}_k$ such that the following two
diagrams commute:

\[
\begin{array}{ccc}
M & \xrightarrow{\delta} & M \otimes C \\
\downarrow{\delta} & & \downarrow{\delta \otimes \text{Id}_C} \\
M \otimes C & \xrightarrow{\text{Id}_M \otimes \Delta} & M \otimes C \otimes C \\
\end{array}
\quad \text{and} \quad
\begin{array}{ccc}
M & \xrightarrow{\delta} & M \otimes C \\
\downarrow{\delta} & & \downarrow{\text{Id}_M \otimes \varepsilon} \\
M \otimes \mathbb{K} & & \\
\end{array}
\]

In analogy with the “action” terminology for modules, we will say that the map \( \delta \) defines a \textit{coaction} of \( C \) on \( M \). The sumless Sweedler notation will also be employed for the coaction, indexing by 0 the factor(s) belonging to \( M \):

\[
\delta m = m_{(0)} \otimes m_{(1)} \quad (m \in M).
\]

Given right \( C \)-comodules \( M \) and \( N \), a \textit{homomorphism} (or \( C \)-comodule map) from \( M \) to \( N \) is a map \( f: M \to N \) in \( \text{Vect}_\mathbb{K} \) such that the following diagram commutes:

\[
\begin{array}{ccc}
M & \xrightarrow{f} & N \\
\downarrow{\delta_M} & & \downarrow{\delta_N} \\
M \otimes C & \xrightarrow{f \otimes \text{Id}_C} & N \otimes C \\
\end{array}
\]

This yields a new category, \( \text{Mod}^C \).

Let \( M \in \text{Mod}^C \). A \( \mathbb{K} \)-subspace \( M' \subseteq M \) is called a \textit{subcomodule} if \( \delta_M(M') \subseteq M' \otimes C \), viewed as a subspace of \( M \otimes C \) in the usual fashion. We may then take \( \delta_{M'} = \delta_M|_{M'} \) to equip \( M' \) with the structure of a right \( C \)-comodule. The embedding \( M' \hookrightarrow M \) and the canonical surjection \( M \twoheadrightarrow M/M' \) both become maps in \( \text{Mod}^C \).

Again, we leave to the reader the routine task of checking that subcomodules of \( M \) are the same as the kernels (in \( \text{Vect}_\mathbb{K} \)) of comodule maps \( f: M \to N \) and to formulate the comodule variants of the standard isomorphism theorems; see Sweedler [194, Proposition 2.0.1] for example.

There is of course an analogous category of left \( C \)-comodules, \( \text{CMod} \). The formulation of the left-sided version of the foregoing can safely be left to the imagination of the reader. Corresponding to the earlier equivalence of module categories, \( \text{Mod}_A \cong A^{\text{op}} \text{Mod} \) for \( A \in \text{Alg}_\mathbb{K} \), we now have an equivalence

\[
\text{CMod} \cong \text{Mod}^{C^{\text{cop}}},
\]

where \( C^{\text{cop}} \) is the coopposite coalgebra (§9.1.1). For \( M \in \text{CMod} \), the coaction \( \delta: M \to C \otimes M \) will be written as follows, preserving the convention that \( m_{(i)} \in C \) for \( i \neq 0 \):

\[
\delta m = m_{(-1)} \otimes m_{(0)} \quad (m \in M).
\]
Example 9.10 (The regular comodules). Any coalgebra $C$ can be viewed as both a right and a left comodule over itself using the comultiplication $\Delta$ as coaction in either case. In analogy with the case of algebras, these comodules will be referred to as the regular comodules. Subcomodules of the right regular comodule $C$ are called right coideals of $C$; similarly for left coideals. If $C$ is cocommutative, then right and left coideals of $C$ both are the same as subcoalgebras.

Example 9.11 ($kX$-comodules and $X$-graded $k$-vector spaces). Consider the group-like colalgebra $kX$ of a given $X \in \text{Sets}$ (Example 9.2). Since $kX \cong (kX)^{\text{op}}$, right and left comodules are “the same” for $kX$. Let $M$ be such a comodule, with coaction $\delta : M \to M \otimes kX = \bigoplus_{x \in X} M \otimes x$. If $m \in M$, then $\delta m = \sum_{x \in X} m^x \otimes x$ for unique $m^x \in M$ that are almost all 0. Defining $\delta^x \in \text{End}_k(M)$ by $\delta^x m = m^x$, the two commutative diagrams (9.19) are easily seen to be equivalent to the identities

$$\delta^x \delta^y = \delta_{x,y} \delta^x \quad (x, y \in X) \quad \text{and} \quad \text{Id}_M = \sum_{x \in X} \delta^x,$$

where $\delta_{x,y}$ is the Kronecker delta. It follows that $M = \bigoplus_{x \in X} \delta^x M$ is an $X$-graded $k$-vector space (Exercise 1.1.11). Conversely, any $X$-graded $k$-vector space $M = \bigoplus_{x \in X} M^x$ may be equipped with a $kX$-coaction by defining $\delta m = \sum_{x \in X} m^x \otimes x$, where $m^x \in M^x$ is the $x$-homogeneous component of $m$. Finally, $kX$-comodule maps $f : M \to N$ are just $k$-linear maps satisfying $f(M^x) \subseteq N^x$ for all $x \in X$. In sum, $kX$-comodules are the same as $X$-graded $k$-vector spaces:

$$\text{Mod}^{kX} \equiv kX\text{Mod} \equiv \text{Vect}^X_k.$$

9.2.2. Local Finiteness

Let $A \in \text{Alg}_k$ and $M \in A\text{Mod}$. Recall that $M$ is said to be locally finite, if $\dim_k A.m < \infty$ for all $m \in M$—we have encountered such modules in the context of enveloping algebras (e.g., §5.7.3) and, more recently, in the proof of Theorem 9.8. In general, we define the locally finite part of $M$ by

$$M_{\text{fin}} \overset{\text{def}}{=} \{ m \in M \mid \dim_k A.m < \infty \}$$

Thus $M_{\text{fin}}$ is the sum of all locally finite submodules of $M$. It is easy to see that $(\cdot)_{\text{fin}}$ gives a functor from $A\text{Mod}$ onto the full subcategory of $A\text{Mod}$ consisting of all locally finite $A$-modules.

By contrast, all comodules and coalgebras are automatically locally finite as the theorem below shows. Part (a) is often called the Fundamental Theorem on Coalgebras (e.g., Sweedler [194, Theorem 2.2.1]); we shall refer to parts (a) and (b) as the Finiteness Theorem for coalgebras and for comodules, respectively.

Finiteness Theorem for Coalgebras and Comodules. Let $C \in \text{Coalg}_k$ and let $M$ be a right or left $C$-comodule. Then:
(a) Every finite subset of $C$ is contained in a subcoalgebra of $C$ that is finite dimensional over $k$.

(b) Every finite subset of $M$ is contained in a finite-dimensional subcomodule of $M$.

Proof. (a) Since the sum of subcoalgebras of $C$ is again a subcoalgebra, it suffices to show that every $c \in C$ is contained in a finite-dimensional subcoalgebra of $C$. For this, we again consider the left module action $\mu$ of the convolution algebra $C^*$ on $C$ (9.14). We have already observed in the proof of Theorem 9.8 that $C^* \rightarrow c$ is finite dimensional and that $I = \ker(C^* \rightarrow c)$ is a cofinite ideal of $C^*$ satisfying $\langle I, c \rangle = 0$. Thus, $c$ belongs to the $k$-subspace $I^\perp := \{ x \in C \mid \langle I, x \rangle = 0 \}$ of $C$. It suffices to show that $I^\perp$ is in fact a finite-dimensional subcoalgebra of $C$. To this end, consider the canonical epimorphism $\pi: C^* \rightarrow C^*/I$ in $\text{Alg}_k$ and the corresponding map $\pi' = \pi^*: (C^*/I)^*= (C^*/I)^* \rightarrow C^{*\circ}$ in $\text{Coalg}_k$ (Proposition 9.7). The image of $\pi^*$ is $C^{*\circ}_{I^\perp} = \{ x \in C^{*\circ} \mid \langle I, x \rangle = 0 \}$, a finite-dimensional subcoalgebra of $C^{*\circ}$. Recalling that $C$ is a subcoalgebra of $C^{*\circ}$ as well by (9.15), we obtain that $C \cap C^{*\circ}_{I^\perp} = I^\perp$ is a finite-dimensional subcoalgebra of $C^{*\circ}$ (Exercise 9.1.1), and hence of $C$ as desired.

(b) will be proved in §9.2.3. □

Note that the Finiteness Theorem for comodules implies in particular that irreducible comodules must be finite dimensional. Here, as the reader will have guessed, a comodule $M$ is said to be irreducible if $M \neq 0$ and the only subcomodules of $M$ are 0 and $M$ itself. Similarly, it follows from the Finiteness Theorem for coalgebras that simple coalgebras must be finite dimensional, where, perhaps somewhat more surprisingly, a coalgebra $C$ is called simple if $C \neq 0$ and $C$ has no subcoalgebras other than 0 and itself.

9.2.3. The Passage between Comodules and Modules

We now consider the passage between modules and comodules, keeping the notation of §9.2.2. It will turn out that every right $C$-comodule is automatically a left module over the algebra $C^*$ in a natural way, whereas only locally finite left modules of an algebra $A$ are also right $A^{*\circ}$-comodules.

Recall that a left $A$-module structure on $M \in \text{Vect}_k$ is given by an “action” map $\mu = \mu_M: A \otimes M \rightarrow M$, $a \otimes m \mapsto a.m$, in $\text{Vect}_k$ that is subject to certain axioms. Under the $k$-linear isomorphism

$$\text{Hom}_k(A \otimes M, M) \cong \text{Hom}_k(M, \text{Hom}_k(A, M))$$

that is given by $\text{Hom}-\otimes$ adjunction (B.16), $\mu$ corresponds to the map

$$\mu' = \mu'_M: M \longrightarrow \text{Hom}_k(A, M)$$

(9.22) $\begin{array}{ccc}
\mu' & : & M \\
\downarrow & \cong & \downarrow \\
m & \mapsto & (a \mapsto a.m)
\end{array}$
The subspace of $\text{Hom}_k(A, M)$ consisting of all finite-rank homomorphisms coincides with the image of the canonical monomorphism $M \otimes A^* \hookrightarrow \text{Hom}_k(A, M)$, $m \otimes a^* \mapsto (a \mapsto m(a^*, a))$ (B.18); we will view this map as an embedding. Thus, $\mu'(m) \in M \otimes A^*$ if and only if $m \in M_{fin}$. In this case, $I = \{a \in A \mid aA.m = 0\}$ is a cofinite ideal of $A$ and $I \subseteq \text{Ker} \mu'(m)$. Therefore, we do in fact have $\mu'(m) \in M \otimes A^\circ$.

Applying the foregoing to $M_{fin}$ in place of $M$, we see that $\mu(M_{fin}) \subseteq M_{fin} \otimes A^\circ$. Thus, for any $M \in A_{\text{Mod}}$, we have a $k$-linear map

$$\mu_{\text{fin}} = \mu'|_{M_{\text{fin}}} : M_{\text{fin}} \rightarrow M_{\text{fin}} \otimes A^\circ. \tag{9.23}$$

It is easy to check that $\mu_{\text{fin}}$ makes $M_{\text{fin}}$ into an $A^\circ$-comodule; we will elaborate on this in the proof of the following proposition.

**Proposition 9.12.**

(a) Let $C \in \text{Coal}_{k\text{-Alg}}$. Then every right $C$-comodule is a locally finite left module over the algebra $C^*$ via $\text{Hom} \otimes$ adjunction, and every $C$-comodule map becomes a $C^*$-module map.

(b) Let $A \in \text{Alg}_{k\text{-Alg}}$. A left $A$-module $M$ is a right $A^\circ$-comodule via $\text{Hom} \otimes$ adjunction if and only if $M$ is locally finite. Module maps between locally finite left $A$-modules become right $A^\circ$-comodule maps in this way.

**Proof.** (a) Let $M \in \text{Mod}^C$. The standard embeddings $C \hookrightarrow C^{**}$ and $M \otimes C^{**} \hookrightarrow \text{Hom}_k(C^*, M)$ give rise to a map $M \otimes C \hookrightarrow M \otimes C^{**} \hookrightarrow \text{Hom}_k(C^*, M)$, which in turn yields a map

$$\text{Hom}_k(M, M \otimes C) \leftarrow \text{Hom}_k(M, \text{Hom}_k(C^*, M)) \xrightarrow{\sim} \text{Hom}_k(C^* \otimes M, M).$$

The image of the coaction $\delta_M$ under this map is as follows:

$$c^* \otimes m \mapsto c^*.m := m_{(0)}(c^*, m_{(1)}) \quad (c^* \in C^*, m \in M). \tag{9.24}$$

This defines a left $C^*$-module action on $M$: the module axioms (1.17) are readily verified using (9.24) and the comodule axioms (9.19). For example, for $c^*, d^* \in C^*$ and $m \in M$, we compute

$$c^*. (d^*.m) = c^*. m_{(0)}(d^*, m_{(1)}) \overset{\text{(9.19)}}{=} m_{(0)}(c^*. m_{(1)}) d^*. m_{(2)} = m_{(0)}(c^*. d^*. m_{(1)}) = (c^*. d^*. m).$$

It is similarly straightforward to check that $C$-comodule maps are module maps for the $C^*$-action (9.24). Finally, (9.24) also shows that $C^*.m$ is contained in the $k$-span of the (finite) set of components $m_{(0)} \in M$. Thus, $C^*.m$ is finite dimensional for all $m \in M$, proving that the $C^*$-module $M$ is locally finite.

(b) We have already seen that, under the isomorphism (9.21), the image $\mu'$ of the action map $\mu$ of $M$ belongs to $\text{Hom}_k(M, M \otimes A^\circ) \subseteq \text{Hom}_k(M, \text{Hom}_k(A, M))$ if and only if $M$ is locally finite. In this case, writing $\mu' : M \rightarrow M \otimes A^\circ, m \mapsto m_{(0)} \otimes m_{(1)}$, we have

$$a.m = m_{(0)}(m_{(1)}, a) \quad (a \in A, m \in M). \tag{9.25}$$
In particular, \( m = 1.m = m_{(0)}(m_{(1)}, 1) = (\text{Id}_M \otimes e_A) \circ \mu'(m) \) as required in (9.19).
Similarly, the “coassociativity” axiom in (9.19) follows from the corresponding “associativity” module axiom in (1.17) by retraceing the steps of the computation in the proof of (a). Thus, \( M \) becomes a right \( A^\circ \)-comodule. Finally, if \( f: M \to N \) is an \( A \)-module map between locally finite left \( A \)-modules, then

\[
f(m_{(0)})(m_{(1)}, a) = f(a.m) = a.f(m) = f(m_{(0)})(f(m_{(1)}), a) \quad (a \in A, m \in M).
\]

Therefore, \( f(m_{(0)}) \otimes m_{(1)} = f(m_{(0)}) \otimes f(m_{(1)}) \); so \( f \) is an \( A^\circ \)-comodule map. \( \square \)

There is of course a version of Proposition 9.12 where the sides of the actions and coactions are switched. We will freely use this version as well. Formally, the switch in sides can be effected by applying the proposition to \( C^{\text{cop}} \) and \( A^{\text{op}} \) and using the isomorphisms \((C^{\text{cop}})^{\circ} \equiv (C^\circ)^{\text{cop}} \) and \((A^{\text{op}})^{\circ} \equiv (A^\circ)^{\text{cop}} \) (Exercise 9.1.7).

**Proof of the Finiteness Theorem for Comodules.** We are now ready to prove part (b) of the Finiteness Theorem in §9.2.2: if \((M, \delta_M)\) is a right or left comodule over \( C \in \text{Coalg}_k \), then every finite subset \( S \subseteq M \) is contained in a finite-dimensional subcomodule of \( M \). By the foregoing, we may focus on the case where \( M \) is a right comodule. Viewing \( M \) as (locally finite) left module over the algebra \( C^\circ \) as in Proposition 9.12(a), the \( C^\circ \)-submodule \( M' := C^\circ.S \) of \( M \) is finite dimensional and \( S \subseteq M' \). Thus, it suffices to show that \( M' \) is also a \( C \)-submodule of \( M \). But Proposition 9.12(b) interpretes the \( C^\circ \)-module embedding \( M' \hookrightarrow M \) as a \( C^\circ \)-comodule map, with \( M \) being viewed as right \( C^{\text{cop}} \)-comodule via the composite of \( \delta_M: M \to M \otimes C \) with the embedding \( M \otimes C \hookrightarrow M \otimes C^{\text{cop}} \) coming from (9.15). Thus, \( \delta_M(M') \subseteq (M \otimes C) \cap (M' \otimes C^{\text{cop}}) = M' \otimes C \) as desired. \( \square \)

**Exercises for Section 9.2**

**9.2.1 (Convolution).** (a) Let \( \alpha: A \to A' \) be a map in \( \text{Alg}_k \) and \( \gamma: C' \to C \) a map in \( \text{Coalg}_k \). Show that \( \alpha \circ \gamma^*: \text{Hom}_k(C, A) \to \text{Hom}_k(C', A') \) is an algebra map for the convolution algebra structures.

(b) Let \( M \in C^\text{Mod} \) and \( V \in A^\text{Mod} \). Show that \( \text{Hom}_k(M, V) \) becomes a left module over the convolution algebra \( \text{Hom}_k(C, A) \), with the convolution action \((f \ast g)(m) = f(m_{(-1)}).g(m_{(0)}) \) for \( f \in \text{Hom}_k(C, A) \), \( g \in \text{Hom}_k(M, V) \) and \( m \in M \). Further, if \( \nu: V \to V' \) is a map in \( A^\text{Mod} \) and \( \mu: M' \to M \) a map in \( C^\text{Mod} \), then \( \nu \circ \mu^*: \text{Hom}_k(M, V) \to \text{Hom}_k(M', V') \) is a map of \( \text{Hom}_k(C, A) \)-modules.

**9.2.2 (Locally finite modules).** Let \( A \in \text{Alg}_k \) and let \( 0 \to M' \to M \to M'' \to 0 \) be a short exact sequence in \( A^\text{Mod} \). Show:

(a) If \( M \) is locally finite, then so are \( M' \) and \( M'' \).

(b) Assuming \( A \) to be left noetherian or affine, show that the converse of (a) holds: if \( M' \) and \( M'' \) are both locally finite, then so is \( M \). (Use Exercise 1.1.8.)
(c) Give an example, where \( M' \) and \( M'' \) are both locally finite, but \( M \) is not.

9.2.3 (The subcoalgebra \( C^M \subseteq C \)). Let \( C \in \text{Coalg}_k \) and let \( (M, \delta) \in \text{Mod}_C^C \). The purpose of this exercise is to construct a subcoalgebra \( C^M \subseteq C \) that serves as an analog of the algebra \( A_V \cong A/\text{Ker} V \) for a module \( V \in \text{A} \text{Mod}\). To construct \( C^M \), first assume that \( \dim_k M < \infty \). Viewing \( M \in C^* \text{Mod} \) as in Proposition 9.12(a), we obtain an algebra map \( C^* \rightarrow \text{End}_k(M) \). By Proposition 9.7, this map in turn yields a coalgebra map \( \delta' : \text{End}_k(M) \rightarrow C^* \) with image in \( C \), where \( C \hookrightarrow C^* \) as in (9.15). Define \( C^M := \text{Im} \delta' \); this is a subcoalgebra of \( C \). In general, define the subcoalgebra \( C^M \) to be the sum of all \( C^M' \) with \( M' \) ranging over the finite-dimensional subcomodules of \( M \). Verify the following properties of \( C^M \):

(a) \( C^M = \sum_{f \in M^*} ((f \otimes \text{Id}_C) \circ \delta)M \); it suffices to let \( f \) run over any \( X \subseteq M^* \) that is dense in the sense that \( X^\perp = \{ m \in M \mid \langle x, m \rangle = 0 \text{ for all } x \in X \} = \{0\} \). Furthermore, \( C^M \) is the unique smallest \( k \)-subspace \( U \subseteq C \) such that \( \delta M \subseteq M \otimes U \).

(b) \( \ker_{C^M} M = (C^M)^* = \{ c^* \in C^* \mid \langle c^*, C^M \rangle = 0 \} \).

(c) If \( M \) is finite dimensional, then so is \( C^M \).

(d) If \( M \) is an irreducible comodule, then \( C^M \) is a simple coalgebra.

9.3. Bialgebras and Hopf Algebras

Group algebras and enveloping algebras are algebras and that are also coalgebras, equipped with a comultiplication \( \Delta \) and a counit \( \varepsilon \) that respect the underlying algebra structure. Such coalgebras are called bialgebras. This section introduces bialgebras formally along with the subclass of Hopf algebras, around which most of the remainder if this book will revolve. Both group algebras and enveloping algebras are in fact Hopf algebras.

9.3.1. Bialgebras

A \( k \)-bialgebra is a \( k \)-vector space \( B \) that is simultaneously a \( k \)-algebra and a \( k \)-coalgebra and these two structures are compatible in the sense that the following two equivalent conditions are satisfied:

(i) The comultiplication \( \Delta \) and the counit \( \varepsilon \) of \( B \) are algebra maps;

(ii) the multiplication \( m \) and the unit \( u \) of \( B \) are coalgebra maps.

To establish the equivalence of (i) and (ii), note that \( m \) is a coalgebra map if and only if \( \Delta \circ m = (m \otimes m) \circ \Delta_{B \otimes B} \) and \( \varepsilon \circ m = \varepsilon_{B \otimes B} \), where the definitions of \( \Delta_{B \otimes B} \) and \( \varepsilon_{B \otimes B} \) from Exercise 9.1.11 are understood: for \( x, y \in B \),

\[
\Delta(xy) = x_{(1)}y_{(1)} \otimes x_{(2)}y_{(2)} = \Delta x \Delta y \quad \text{and} \quad \langle \varepsilon, xy \rangle = \langle \varepsilon, x \rangle \langle \varepsilon, y \rangle.
\]
Similarly, $u$ is a coalgebra map if and only if $\Delta \circ u = (u \otimes u) \circ \Delta_k$ and $\varepsilon \circ u = \varepsilon_k$
or, explicitly, with $1 = 1_B$:

\[ \Delta 1 = 1 \otimes 1 \quad \text{and} \quad \langle \varepsilon, 1 \rangle = 1_k. \]

The displayed equations above together state exactly that (i) is satisfied.

Naturally, a homomorphism of bialgebras is defined to be a $k$-linear map between $k$-bialgebras that is both an algebra and a coalgebra map. The resulting category of $k$-bialgebras will be denoted by $\text{BiAlg}_k$.

A bi-ideal of $B \in \text{BiAlg}_k$, by definition, is a $k$-subspace $I \subseteq B$ that is simultaneously an ideal and a coideal of $B$; in this case, the factor $B/I$ inherits the obvious bialgebra structure from $B$. Similarly, a subbialgebra of $B$ is a subalgebra $A \subseteq B$ that is also a subcoalgebra; so $\Delta A \subseteq A \otimes A$. More generally, any subalgebra $A \subseteq B$ satisfying $\Delta A \subseteq A \otimes B$—so $A$ is a right coideal of $B$—is called a right coideal subalgebra of $B$; likewise for left coideal subalgebras. Algebra terminology such as “commutative” and “semisimple” and coalgebra notions such as “cocommutative,” when applied to a bialgebra $B$, naturally refer to the appropriate underlying algebra or coalgebra structures of $B$. Simultaneously passing to the opposite algebra and coopposite coalgebra structures, one obtains a bialgebra, $B^{\text{bi-op}}$,

that is sometimes handy when switching sides is desirable or necessary.

Group-like Elements. Let $B \in \text{BiAlg}_k$. Recall that the group-like elements of $B$, viewed as a coalgebra, form the set $GB = \{ b \in B \mid \Delta b = b \otimes 1 + 1 \otimes b, \langle \varepsilon, b \rangle = 1_k \}$. Since $\Delta$ and $\varepsilon$ are now algebra maps, $GB$ is in fact a submonoid of the multiplicative monoid $(B, \cdot)$: $1 = 1_B \in GB$ and $gh \in GB$ for $g, h \in GB$. The functor $G: \text{Coalg}_k \rightarrow \text{Sets}$ restricts to a functor $G: \text{BiAlg}_k \rightarrow \text{Monoids}$, the category of monoids. Furthermore, since distinct elements of $GB$ are $k$-linearly independent in $B$ (Lemma 9.1), the $k$-linear span of $GB$ in $B$ is a subbialgebra of $B$ that is isomorphic to the monoid algebra $k[GB]$ (§3.1.2).

9.3.2. Primitive Elements

Let $B \in \text{BiAlg}_k$ with identity element $1 = 1_B$. The set of primitive elements of $B$ is defined by

\[ LB = \{ b \in B \mid \Delta b = b \otimes 1 + 1 \otimes b \}. \]

The counit axiom implies $b = \langle \varepsilon, b \rangle 1 + \langle \varepsilon, 1 \rangle b$ for $b \in LB$ and $\langle \varepsilon, 1 \rangle = 1$, since $\varepsilon$ is an algebra map. Thus, for $b \in LB$,

\[ \langle \varepsilon, b \rangle = 0. \]
Plainly, $LB$ is a $k$-subspace of $B$; in fact, $LB$ is a Lie subalgebra of $B_{\text{Lie}}$: for $b, c \in LB$, the equality $\Delta[b, c] = [\Delta b, \Delta c]$, which holds because $\Delta$ is an algebra map, becomes

$$\Delta[b, c] = (b \otimes 1 + 1 \otimes b)(c \otimes 1 + 1 \otimes c) - (c \otimes 1 + 1 \otimes c)(b \otimes 1 + 1 \otimes b) = [b, c] \otimes 1 + 1 \otimes [b, c].$$

Again, the construction of $LB$ is functorial, giving a functor $L: \BiAlg_k \to \Lie_k$.

If $\text{char } k = p > 0$, then the binomial theorem gives the following relation:

$$\Delta(b^p) = (b \otimes 1 + 1 \otimes b)^p = b^p \otimes 1 + 1 \otimes b^p \quad (b \in LB).$$

Thus, we obtain a map $(\cdot)^{[\cdot]}: LB \to LB, b \mapsto b^p$, which equips $LB$ with the structure of a Lie $p$-algebra. For the formal definition of Lie $p$-algebras, see Jacobson [110, Section V.7].

Outlook: Primitive Elements and Enveloping Algebras. For any $g \in \text{Lie}_k$, the counit (5.20) and the comultiplication (5.21) of the enveloping algebra $Ug$ are algebra maps, making $Ug$ a $k$-bialgebra. If $\text{char } k = 0$, then it is easy to see from the description of $\Delta$ in Example 9.3 that $L(Ug) = g$. For background, we mention the following facts without proof.

- If $\text{char } k = p > 0$, then $L(Ug)$ is the $k$-linear span of all $x^p k$ with $x \in g$ and $k \in \mathbb{Z}_{+}$; see [153, Proposition 5.5.3].

Next, let $B \in \BiAlg_k$ be arbitrary and put $g := LB$. By the universal property (5.15) of enveloping algebras, the embedding $g \subseteq B_{\text{Lie}}$ gives rise to a map $\beta: Ug \to B$ in $\Alg_k$, which is in fact a map in $\BiAlg_k$ by (5.20) and (5.21). Moreover:

- If $\text{char } k = 0$, then $\beta$ is injective. For a proof, see Bourbaki [27, chap. II §1 Théorème 1] or Montgomery [153, Theorem 5.3.1 and Proposition 5.5.3].
- If $\text{char } k = p > 0$, then $\beta$ factors through the following quotient of $Ug$, called the restricted enveloping algebra of the Lie $p$-algebra $g$:

$$U[[p]]_g := Ug/(b^p - b^{[p]} \mid b \in g).$$

The resulting map $U[[p]]_g \to B$ is in fact injective; see [110] or [153, loc. cit.].

9.3.3. Further Examples of Bialgebras

We have already mentioned (restricted) enveloping algebras of Lie algebras and group algebras as examples of bialgebras. More generally, for any monoid $M$, the monoid algebra $kM$, when equipped with its group-like coalgebra structure (Example 9.2), is a bialgebra with $M = G(kM)$. In this subsection, we explain how an associative unital pairing on a finite-dimensional vector space $V$ gives rise to a bialgebra structure on the algebra of polynomial functions, $O(V) = \text{Sym } V^*$ (Section C.3).
Algebras of Polynomial Functions

Let \( V \in \text{Vect}_k \) be finite dimensional and assume that \( V \) is equipped with an associative pairing, that is, a \( k \)-bilinear map

\[
\mu : V \times V \longrightarrow V
\]
satisfying the condition \( \mu(\mu(x, y), z) = \mu(x, \mu(y, z)) \) for all \( x, y, z \in V \). Regarding \( \mu \) as a linear map \( \mu : V \otimes V \to V \), the associativity condition can be written in the form \( \mu \circ (\text{Id} \otimes \mu) = \mu \circ (\mu \otimes \text{Id}) : V \otimes V \otimes V \to V \). The dual of \( \mu \) yields a linear map,

\[
V^* \overset{\mu^*}{\longrightarrow} (V \otimes V)^* \overset{\Delta}{\longrightarrow} V^* \otimes V^* \hookrightarrow O(V) \otimes O(V).
\]

Here, \( V^* \hookrightarrow O(V) \) is the standard embedding of \( V^* \) as the degree-1 component of \( O(V) \). We will allow ourselves to identify \( V^* \otimes V^* \) and \( (V \otimes V)^* \). By the universal property of symmetric algebras (1.8), the above linear map extends uniquely to an algebra map,

\[
\Delta = \Delta_\mu : O(V) \otimes O(V) \longrightarrow O(V) \otimes O(V).
\]

Thus, \( \Delta|_{V^*} = \mu^* \) and so, dualizing the associativity condition for \( \mu \), we obtain the identity \( ((\text{Id} \otimes \Delta) \circ \Delta)|_{V^*} = ((\Delta \otimes \text{Id}) \circ \Delta)|_{V^*} \). This implies that \( \Delta \) is coassociative, because \( V^* \) generates the algebra \( O(V) \). Now assume that the given pairing \( \mu \) is also unital, that is, there exists an element \( e \in V \) such that \( \mu(e, x) = x = \mu(x, e) \) for all \( x \in V \). Then the evaluation map \( V^* \to k, f \mapsto \langle f, e \rangle \), lifts to an algebra map,

\[
\varepsilon = \varepsilon_\mu : O(V) \longrightarrow k,
\]

which is readily checked to satisfy the counit laws. In summary, the pairing \( \mu \) equips the algebra \( O(V) \) with a bialgebra structure.

Evaluation gives a linear embedding \( V^* \hookrightarrow \k^V \), the algebra of of all \( k \)-valued functions on \( V \) with pointwise addition and multiplication. This embedding extends to an algebra map \( O(V) \to \k^V \). Similarly, since \( O(V) \otimes O(V) \cong O(V \otimes V) \), we also have an algebra map \( O(V) \otimes O(V) \to \k^{V \times V} \). For infinite \( k \), all these maps are monomorphisms (Exercise C.3.2) and we may regard each \( f \otimes g \in O(V) \otimes O(V) \) as a \( k \)-valued function on \( V \times V \) via \( (f \otimes g)(x, y) = f(x)g(y) \). The comultiplication \( \Delta : O(V) \to O(V) \otimes O(V) \) then becomes \( \Delta(f)(x, y) = f(\mu(x, y)) \), because this holds for the generators \( f \in V^* \). Thus, if \( 0 \neq f \in O(V) \) is multiplicative, in the sense that \( f(\mu(x, y)) = f(x)f(y) \) for all \( x, y \in V \), then \( \Delta f = f \otimes f \) and so \( f \) is group-like. Similarly, if \( f(\mu(x, y)) = f(x) + f(y) \) for all \( x, y \in V \), then \( \Delta f = f \otimes 1 + 1 \otimes f \) and hence \( f \) is primitive.

**Example 9.13** (Polynomial functions on \( n \times n \) matrices). We now describe the counit and the comultiplication of \( O(V) \) for the vector space \( V = \text{Mat}_n(k) \) of \( n \times n \) matrices over \( k \), using matrix multiplication as the pairing \( \mu \). Note that matrix multiplication is indeed associative and unital. We will often omit \( k \) from our notation, writing \( \text{Mat}_n = \text{Mat}_n(k) \). Choosing as our basis of the linear dual \( \text{Mat}_n^* \)
the forms $X_{ij}$ that associate to each matrix $x = (x_{ij}) \in \text{Mat}_n$ the $(i, j)$-entry $x_{ij} \in \mathbb{k}$, we may identify $O(\text{Mat}_n)$ with the polynomial algebra in $n^2$ variables:

$$O(\text{Mat}_n) = O(\text{Mat}_n(\mathbb{k})) = \mathbb{k}[X_{ij} \mid 1 \leq i, j \leq n].$$

The matrix multiplication formula $(xy)_{ij} = \sum_k x_{ik} y_{kj}$ leads to the following expression for the comultiplication of $O(\text{Mat}_n)$ on the generators $X_{ij}$; to evaluate $\Delta$ on arbitrary polynomials, one uses the fact that $\Delta$ is an algebra map:

$$\Delta X_{ij} = \sum_k X_{ik} \otimes X_{kj}.$$  

(9.26)

The counit of $O(\text{Mat}_n)$ comes from the identity matrix $1_{n \times n}$. Since $\langle X_{ij}, 1_{n \times n} \rangle = \delta_{ij}$, we obtain the following formula on the generators $X_{ij}$:

$$\langle \varepsilon, X_{ij} \rangle = \delta_{ij}.$$  

(9.27)

An important element of $O(\text{Mat}_n)$, homogeneous of degree $n$, is the determinant,

$$D = \sum_{s \in S_n} \text{sgn}(s) X_{1s(1)}X_{2s(2)} \cdots X_{ns(n)}.$$  

(9.28)

Since $\det(xy) = \det(x)\det(y)$ for all $x, y \in \text{Mat}_n$, the above discussion (after embedding $\mathbb{k}$ into some infinite field, if necessary) shows that $D$ is a group-like element of the bialgebra $O(\text{Mat}_n)$.

### 9.3.4. Hopf Algebras

For any $B \in \text{BiAlg}_k$, we may equip $\text{End}_k(B)$ with the convolution algebra structure ($\S 9.1.4$). If $\text{Id}_B$ is invertible in $\text{End}_k(B)$, then $B$ is called a **Hopf algebra** and the inverse of $\text{Id}_B$ in $\text{End}_k(B)^\times$ is called the **antipode** of $B$. We will generally denote the antipode by $S$ or $S_B$. Thus, $S$ is characterized by the condition

$$S(b_{(1)})b_{(2)} = \langle \varepsilon, b \rangle 1 = b_{(1)}S(b_{(2)}) \quad (b \in B),$$  

(9.29)

where $1 = 1_B$ is the identity element of $B$. The equalities (9.29) show in particular, that $S$ also serves as an antipode for $B^{\text{bi-op}}$; so $B^{\text{bi-op}}$ is a Hopf algebra as well.

The notions of a **Hopf subalgebra** and a **Hopf ideal** are obtained from those of a subbialgebra and bi-ideal by also requiring stability under the antipode $S$. Thus, for example, a Hopf ideal of a Hopf algebra $H$ is a $k$-subspace $I \subseteq H$ such that $SI \subseteq I$ and such that $I$ is an ideal for the underlying algebra structure of $H$ and also a coideal for the coalgebra structure of $H$, that is, $\langle \varepsilon, I \rangle = 0$ and $\Delta I \subseteq I \otimes H + H \otimes I$. In this case, $H/I$ becomes a Hopf algebra in the evident fashion.

It can be shown that any bialgebra homomorphism $f : H \to K$ between arbitrary Hopf algebras $H$ and $K$ automatically preserves the antipode: $f \circ S_H = S_K \circ f$ (Exercise 9.3.1). Thus, **homomorphisms of Hopf algebras** are just bialgebra homomorphisms between Hopf algebras. Thus, Hopf algebras over $\mathbb{k}$ form a full subcategory of $\text{BiAlg}_k$, which will be denoted by

$$\text{HopfAlg}_k.$$
It also turns out that, in a finite-dimensional Hopf algebra, all subbialgebras and all bi-ideals are automatically stable under the antipode, and hence they are Hopf subalgebras and Hopf ideals, respectively (Exercise 9.3.2). However, this fails for general Hopf algebras.

**Group-like and Primitive Elements.** For any \( H \in \text{HopfAlg}_k \), the monoid of group-like elements \( GH = \{ g \in H \mid \Delta g = g \otimes g, \langle \epsilon, g \rangle = 1_k \} \) is in fact a group. Indeed, for \( g \in GH \), condition (9.29) says that \( (Sg)g = 1 = g(Sg) \) or, equivalently,

\[
Sg = g^{-1} \quad (g \in GH).
\]

Thus, \( GH \) is a subgroup of \( H^\times \), the group of invertible elements of the underlying algebra of \( H \). In this way, we obtain a functor \( G : \text{HopfAlg}_k \rightarrow \text{Groups} \). The \( k \)-linear span of \( GH \) in \( H \) is a Hopf subalgebra of \( H \) that is isomorphic to the group algebra \( k[GH] \) (Lemma 9.1); so we have the following embedding in \( \text{HopfAlg}_k \):

\[
(9.31) \quad k[GH] \hookrightarrow H.
\]

Next, let us consider the Lie algebra \( LH = \{ x \in H \mid \Delta x = x \otimes 1 + 1 \otimes x \} \) of all primitive elements of \( H \). Since \( \langle \epsilon, x \rangle = 0 \) for \( x \in LH \) and \( S1 = 1 \) by (9.30), condition (9.29) gives \( 0 = \langle \epsilon, x \rangle 1 = (Sx)1 + (S1)x = Sx + x \). Thus,

\[
(9.32) \quad Sx = -x \quad (x \in LH).
\]

**Standard Example: Group Algebras.** For any group \( G \), the group algebra \( kG \) has an antipode \( S \) that is given by \( Sg = g^{-1} \) for \( g \in G \) (3.28). Thus, the group algebra \( kG \) is in fact a Hopf algebra, evidently cocommutative. Moreover, \( G \) is the collection of group-like elements: \( G(kG) = G \) (Example 9.2). The earlier group algebra functor \( \mathbb{k} \cdot : \text{Groups} \rightarrow \text{Alg}_k \) is the composite of the forgetful functor \( \text{HopfAlg}_k \rightarrow \text{Alg}_k \) dropping the coalgebra structure and the antipode with the functor sending a group to its group (Hopf) algebra,

\[
\mathbb{k} \cdot : \text{Groups} \rightarrow \text{HopfAlg}_k.
\]

This functor is left adjoint to the group-like functor \( G : \text{HopfAlg}_k \rightarrow \text{Groups} \). Namely, for any \( G \in \text{Groups} \) and \( H \in \text{HopfAlg}_k \), the earlier bijection (3.2) restricts to a bijection, functorial in both \( G \) and \( H \),

\[
(9.33) \quad \text{Hom}_{\text{HopfAlg}_k}(kG, H) \cong \text{Hom}_{\text{Groups}}(G, GH)
\]

**Standard Example: Enveloping Algebras.** The enveloping algebra \( Ug \) of any \( g \in \text{Lie}_k \) admits an antipode \( S \) that is given by formula (9.32) for \( x \in g \); see (5.24). Thus, \( Ug \) is a Hopf algebra, also clearly cocommutative. As for groups and group algebras, the earlier enveloping algebra functor \( U : \text{Lie}_k \rightarrow \text{Alg}_k \) arises from a functor

\[
U : \text{Lie}_k \rightarrow \text{HopfAlg}_k
\]
that is left adjoint to the primitive-element functor $L : \text{HopfAlg}_k \to \text{Lie}_k$: the bijection (5.15) restricts to a bijection, functorial in both $\mathfrak{g} \in \text{Lie}_k$ and $H \in \text{HopfAlg}_k$,

\begin{equation}
\text{Hom}_{\text{HopfAlg}_k}(U\mathfrak{g}, H) \cong \text{Hom}_{\text{Lie}_k}(\mathfrak{g}, LH)
\end{equation}

### 9.3.5. Properties of the Antipode

The next proposition lists some fundamental properties of the antipode, which will be used tacitly in numerous calculations below without explicit reference.

**Proposition 9.14.** Let $H \in \text{HopfAlg}_k$. Then the antipode $S$ gives a map $H \to H^{\text{biop}}$ in $\text{HopfAlg}_k$. Explicitly:

(a) $S1 = 1$ and $S(hk) = SkSh$ for $h, k \in H$.

(b) $\varepsilon \circ S = \varepsilon$ and $(Sh)_1 \otimes (Sh)_2 = S(h_2) \otimes S(h_1)$ for $h \in H$.

**Proof.** Parts (a) and (b) state respectively that $S$ gives a map $H \to H^{\text{op}}$ in $\text{Alg}_k$ and a map $H \to H^{\text{cop}}$ in $\text{Coalg}_k$. Thus, $S$ gives a bialgebra map $H \to H^{\text{biop}}$ and hence a map in $\text{HopfAlg}_k$. It remains to prove (a) and (b).

(a) We have already pointed out that $S1 = 1$. Next, letting $m$ denote the multiplication of $H$ and $\tau$ the endomorphism of $H^\otimes 2 = H \otimes H$ that switches tensor factors, we need to establish the following equality in $C := \text{Hom}_k(H^\otimes 2, H)$:

$$S \circ m = m \circ (S \otimes S) \circ \tau.$$

For this, we will use the convolution algebra structure of $C$, viewing $H = (H, m, u)$ as the underlying algebra of $H$ and $H^\otimes 2$ as being equipped with the tensor coalgebra structure (Exercise 9.1.11): $\Delta_{H^\otimes 2} = (23) \circ (\Delta \otimes \Delta)$. The identity element of $C$ is given by $1_C = u \circ \varepsilon^\otimes 2$, where $\varepsilon^\otimes 2 = \varepsilon \otimes \varepsilon$ is the counit of $H^\otimes 2$. Thus, putting $\rho = S \circ m$ and $\nu = m \circ (S \otimes S) \circ \tau$ for brevity, it suffices to show that

$$\rho \ast m = 1_C = m \ast \nu$$

holds in $C$, because this will imply the desired equality $\rho = \rho \ast m \ast \nu = \nu$. But since $m : H^\otimes 2 \to H$ is a coalgebra map, it is easy to see that $m^* : \text{Hom}_k(H, H) \to C$ is an algebra map for the convolution algebra structures (Exercise 9.2.1). Therefore, $m^*(S) = \rho$ is the inverse of $m^*(\text{Id}_H) = m$ in $C$, giving the first of the above two equalities. As for the second, we compute, for $h, k \in H$,

$$(m \ast \nu)(h \otimes k) = m(h_1 \otimes k_1)\nu(h_2 \otimes k_2) = h_1 k_1 S(k_2) S(h_2)$$

\begin{align*}
&= h_1 \langle \varepsilon, h \rangle S(h_2) + h_1 S(h_2) \langle \varepsilon, k \rangle \\
&= \langle \varepsilon, h \rangle \langle \varepsilon, k \rangle 1 \\
&= (u \circ \varepsilon^\otimes 2)(h \otimes k).
\end{align*}

This completes the proof of (a). The proof of (b) proceeds in a similar fashion, replacing $C$ by the convolution algebra $\text{Hom}_k(H, H^\otimes 2)$ that arises from the coalgebra
structure of $H$ and the algebra structure of $H^\otimes 2$. We leave the details to the care of the reader. It may be instructive to regard both parts of the proposition as special cases of Exercise 9.3.6 with $f = \text{Id}_H$. □

Hopf algebras $H$ whose antipode satisfies $S \circ S = \text{Id}_H$ are called **involutory**. This class of Hopf algebras includes all group algebras and enveloping algebras and, more generally, all commutative and all cocommutative Hopf algebras (Exercise 9.3.7).

### 9.3.6. Duals of Bialgebras and Hopf Algebras

A remarkable feature of the definition of bialgebras is its symmetry with regard to the roles played by the underlying algebra and coalgebra structures. As we shall see presently, this symmetry also manifests itself in the fact that the finite dual $B^\circ$ of any bialgebra $B$ is again a bialgebra and likewise for Hopf algebras. To wit:

**Proposition 9.15.** Let $(B, m, u, \Delta, \varepsilon) \in \text{BiAlg}_k$. Then:

(a) The finite dual $B^\circ$ becomes a bialgebra with structure maps given by the restrictions of the transposes $m^*, u^*, \Delta^*$ and $\varepsilon^*$.

(b) If $B$ is a Hopf algebra with antipode $S$, then $B^\circ$ is a Hopf algebra with antipode $S^\circ := S^*|_{B^\circ}$.

(c) If $f : A \to B$ is a bialgebra map, then so is $f^* : B^\circ \to A^\circ$.

**Proof.** (a) We already know that $B^\circ$ is a coalgebra with comultiplication $m^*|_{B^\circ}$ and counit $u^*|_{B^\circ}$ (Proposition 9.7), and we have also seen in our discussion of convolution algebras that the full linear dual $B^*$ is an algebra with unit map $\varepsilon^*$ and with multiplication given by pulling back $\Delta^* : (B \otimes B)^* \to B^*$ along the canonical embedding $B^* \otimes B^* \hookrightarrow (B \otimes B)^*$. We need to show that $B^\circ$ is a subalgebra of $B^*$; conditions (i) and (ii) in §9.3.1 are then readily seen to hold for $B^\circ$ by dualizing the commutative diagrams expressing these conditions for $B$.

First note that $\varepsilon = 1_{B^\circ} \in B^\circ$, because $\ker \varepsilon$ is an ideal of codimension 1 in $B$. In order to show that $B^\circ$ is closed under multiplication, consider linear forms $f, g \in B^*$ vanishing on cofinite ideals $I$ and $J$ of $B$, respectively. Then $I \otimes B + B \otimes J$ is a cofinite ideal of $B \otimes B$, being the kernel of the canonical map $B \otimes B \to (B/I) \otimes (B/J)$, and hence the preimage $\Delta^{-1}(I \otimes B + B \otimes J)$ is a cofinite ideal of $B$. Since $fg$ vanishes on this ideal, it follows that $fg \in B^\circ$. Therefore, $B^\circ$ is a subalgebra of $B^*$, and hence $B^\circ$ is a bialgebra.

(b) Assume that $S : B \to B$ is an antipode for $B$. We will show that the transpose $S^*$ maps $B^\circ$ to itself. But if $f \in B^\circ$ vanishes on the cofinite ideal $I$ of $B$, say, then $S^* f$ vanishes on the preimage $S^{-1} I = \{ b \in B \mid S b \in I \}$. It follows from Proposition 9.14 that $S^{-1} I$ is an ideal of $B$, which is evidently cofinite. Therefore,
S^o f \in B^o as desired. The fact that S^o is a convolution inverse for \text{Id}_{B^o} follows readily by dualizing condition (9.29) for B (Exercise 9.3.8).

(c) Let f : A \to B be a bialgebra map. Then f^o : B^o \to A^o is a map in \text{Coalg}_k (Proposition 9.7). Moreover, f^o is a map in \text{Alg}_k (Exercise 9.1.6), and hence this is also true of the restriction f^o. This completes the proof of the proposition. \hfill \Box

By the proposition, the contravariant functor \cdot^o : \text{Alg}_k \to \text{Coalg}_k restricts to contravariant functors \cdot^o : \text{BiAlg}_k \to \text{BiAlg}_k and \cdot^o : \text{HopfAlg}_k \to \text{HopfAlg}_k. These functors commute with the bi-opposite functor \cdot^{\biop} (Exercise 9.1.7) and they are self-adjoint:

**Proposition 9.16.** For any B, C \in \text{BiAlg}_k, there is a natural bijection of sets Hom_{\text{BiAlg}_k}(B, C^o) \isom Hom_{\text{BiAlg}_k}(C, B^o); likewise for Hopf algebras.

**Proof.** The bijection arises by restricting the bijection in Theorem 9.8. In detail, let f : B \to C^o be a bialgebra homomorphism and consider the algebra map B \to C that is the composite of f with the inclusion C^o \hookrightarrow C^o. By Theorem 9.8, this map corresponds to a unique coalgebra map f' : C \to B^o, which is given by f'(c) = (b \mapsto (f(b), c)). Since f is a coalgebra map, f' is also an algebra map and hence it is a bialgebra map. Thus, we have an injective map Hom_{\text{BiAlg}_k}(B, C^o) \to Hom_{\text{BiAlg}_k}(C, B^o). The inverse map is obtained by reversing the roles of B and C. The statement about Hopf algebras is clear, because Hopf homomorphisms are the same as homomorphisms for the underlying bialgebra structures. \hfill \Box

**Example 9.17** (Representative functions on a group). Consider the Hopf algebra \kk G, the group algebra of the group G. The finite dual \((\kk G)^\circ\) is called the Hopf algebra of **representative functions** on G. In view of Exercise 9.1.8,

\[(\kk G)^\circ = \{f \in (\kk G)^* \mid \dim\kk < \infty\}.\]

The \kk-algebra \((\kk G)^\circ\) is isomorphic to the algebra of all function G \to \kk with pointwise algebra operations (Example 9.5). This algebra is commutative and reduced: it has no nonzero nilpotent elements. Hence, the same is true for the subalgebra \((\kk G)^o\). For general G, the Hopf algebra \((\kk G)^o\) may be isomorphic to \kk even if G is infinite (Exercise 9.3.9). Let us now focus on the case where G is finite. Then \((\kk G)^o = (\kk G)^\circ\). Let \((\delta_x)_{x \in G}\) denote the dual basis for the standard basis G of \kk G; so \(\langle \delta_x, y \rangle = \delta(x,y) 1_k\) for x, y \in G. This basis satisfies \(\sum_{x \in G} \delta_x = \varepsilon = 1_{(\kk G)^\circ}\) and it consists of orthogonal idempotents of \((\kk G)^\circ\):

\[\delta_x \delta_y = \delta_{x,y} \delta_x \quad (x, y \in G).\]

Thus, as a \kk-algebra, \((\kk G)^\circ \equiv \kk^{\times |G|}, the direct product of \(|G|\) copies of \kk; see also Example 9.5. The coalgebra structure maps and the antipode of \((\kk G)^\circ\) are given by

\[\Delta \delta_x = \sum_{y \in G} \delta_y \otimes \delta_{y^{-1}x}, \quad \langle \varepsilon, \delta_x \rangle = \delta_{x,1} 1_k \quad \text{and} \quad S \delta_x = \delta_{x^{-1}}.\]
Example 9.18 (The finite dual of a polynomial algebra). Now let \( \mathbb{k}[x] \) be the polynomial algebra, with \( \Delta x = x \otimes 1 + 1 \otimes x \) and \( \langle \varepsilon, x \rangle = 0 \). We may view \( \mathbb{k}[x] \) as the enveloping algebra of the 1-dimensional Lie algebra with its usual Hopf algebra structure. Thus, assuming \( \text{char} \mathbb{k} = 0 \) and putting \( x_n = x^n/n! \), we have an isomorphism of \( \mathbb{k} \)-algebras \((\text{Example } 9.6)\),

\[
\phi: \mathbb{k}[x]^* \longrightarrow \mathbb{k}[x]
\]

\[
\psi \quad \quad \psi
f \longmapsto \sum_{n=0}^{\infty} f(x_n) t^n
\]

In order to describe the part of \( \mathbb{k}[t] \) that corresponds to the finite dual \( \mathbb{k}[x]^\circ \), consider a power series \( s = \sum_{n \geq 0} \sigma_n t^n/n! \in \mathbb{k}[t] \) and put \( f = \phi^{-1}(s) \); so \( \sigma_n = f(x^n) \). By definition of the finite dual, \( f \) belongs to \( \mathbb{k}[x]^\circ \) if and only if \( f \) vanishes on some nonzero ideal of \( \mathbb{k}[x] \) or, equivalently, \( f(x^n a) = 0 \) for some monic polynomial \( a \in \mathbb{k}[x] \) and all \( n \in \mathbb{Z}_+ \). Writing \( a = \sum_{i=0}^{d} \alpha_i x^i \) with \( \alpha_i \in \mathbb{k} \) and \( \alpha_d = 1 \), we obtain the condition \( \sum_{i=0}^{d} \alpha_i f(x^{n+i}) = 0 \), which in turn can be expressed as the linear recursion \( \sigma_{n+d} = -\sum_{i=0}^{d-1} \alpha_i \sigma_{n+i} = 0 (n \geq 0) \). Thus,

\[
\mathbb{k}[x]^\circ = \left\{ \sum_{n \geq 0} \sigma_n t^n/n! \in \mathbb{k}[t] \mid \text{the sequence } \sigma_n \text{ satisfies a linear recursion} \right\}
\]

For more on linearly recursive power series, see \([29, \text{chap. IV } 4, \text{Exercise } 1]\).

### 9.3.7. Group-like and Primitive Elements of the Finite Dual

Let \( B \in \text{BiAlg}_\mathbb{k} \). The monoid of group-like elements of \( B^\circ \) has the following description:

\[
(9.35) \quad \mathbb{G}(B^\circ) = \text{Hom}_{\text{Alg}_\mathbb{k}}(B, \mathbb{k})
\]

This follows from the fact that any algebra map \( f: B \rightarrow \mathbb{k} \) is a \( \mathbb{k} \)-linear form that vanishes on a codimension-1 ideal of \( B \), and hence \( f \in B^\circ \). Moreover, \( f \) preserves the multiplication and the unit of \( B \), which says exactly that \( f \in \mathbb{G}(B^\circ) \). Since the coalgebra structure of \( B \) does not enter here, (9.35) is in fact valid for any \( B \in \text{Alg}_\mathbb{k} \) but \( \mathbb{G}(B^\circ) \) is then merely a set.

In order to describe the Lie algebra \( L(B^\circ) \), let \( B^+ = \ker \varepsilon \) denote the augmentation ideal of \( B \) and \( B^{+2} \) its square. Put \( \mathbb{I} = \mathbb{k} \equiv B/B^+ \), viewed as \((B, B)\)-bimodule with \( B \) acting from both sides via \( \varepsilon \), and let \( \text{Der}(B, \mathbb{I}) \subseteq B^+ \) denote the subspace consisting of all linear forms \( f \in B^+ \) that satisfy the identity

\[
(9.36) \quad \langle f, xy \rangle = \langle f, x \rangle \varepsilon(y) + \varepsilon(x) \langle f, y \rangle \quad (x, y \in B).
\]

**Proposition 9.19.** Let \( B \in \text{BiAlg}_\mathbb{k} \). Then \( L(B^\circ) = \text{Der}(B, \mathbb{I}) \). Furthermore, \( L(B^\circ) \cong (B^+/B^{+2})^* \) in \( \text{Vect}_\mathbb{k} \). If the ideal \( B^+ \) is finitely generated, then \( \dim_{\mathbb{k}} L(B^\circ) < \infty \).
Let us write \((y - f, x) = (f, xy)\) for \(f \in B^*\) and \(x, y \in B\) and recall that \(f \in B^*\) if and only if \(B - f\) is finite dimensional (Exercise 9.1.8). If \(f \in \text{Der}(B, \mathbb{I})\), then \(y - f = \langle e, y \rangle f + \langle f, y \rangle e\) by (9.36) and hence \(B - f \subseteq \mathbb{k}f + \mathbb{k}e\). Therefore, \(\text{Der}(B, \mathbb{I}) \subseteq B^\times\) and, in fact, \(L(B^\times) = \text{Der}(B, \mathbb{I})\), because the condition \(\Delta f = f \otimes 1 + 1 \otimes f\) spells out exactly to (9.36).

To prove the isomorphism \(\text{Der}(B, \mathbb{I}) \cong (B^+/B^{+2})^*\), observe that the map \(x \mapsto x - \langle e, x \rangle 1_B\) is a \(\mathbb{k}\)-linear projection \(B \rightarrow B^+\). Composing this map with the canonical map \(B^+ \rightarrow B^+/B^{+2}\), we obtain an epimorphism \(\pi \colon B \rightarrow B^+/B^{+2}\) in \(\text{Vect}_\mathbb{k}\) and a monomorphism \(\pi^+ \colon (B^+/B^{+2})^* \hookrightarrow B^\times\). Every \(f \in \text{Der}(B, \mathbb{I})\) vanishes on \(1\) and on \(B^{+2}\). Therefore, restriction of linear forms to \(B^+\) gives an embedding \(\text{Der}(B, \mathbb{I}) \hookrightarrow (B^+/B^{+2})^*\). This embedding is in fact an isomorphism. For, a straightforward verification shows that \(\pi(xy) = \pi(x)\langle e, y \rangle + \langle e, x \rangle \pi(y)\) for \(x, y \in B\). Thus, for any \(l \in (B^+/B^{+2})^*\), the form \(\pi^+(l) = l \circ \pi\) belongs to \(\text{Der}(B, \mathbb{I})\) and it restricts to \(l\).

Finally, the congruence \(xmy \equiv \langle e, x \rangle m \langle e, y \rangle \mod B^{+2}\) for \(x, y \in B\) and \(m \in B^+\) implies that \(\dim_k B^+/B^{+2}\) is bounded above by the minimal number of ideal generators of \(B^+\). Therefore, \(\text{Der}(B, \mathbb{I}) \cong (B^+/B^{+2})^*\) is finite dimensional if \(B^+\) is finitely generated.

\[\square\]

### 9.3.8. Examples of Hopf Algebras

So far, we have only met the standard examples of Hopf algebras, group algebras and enveloping algebras, along with their finite duals. The algebra \(O(V)\) of polynomial functions on a \(\mathbb{k}\)-vector space \(V\), with the bialgebra structure provided by a pairing \(V \times V \rightarrow V\) (§9.3.3), generally does not admit an antipode. For example, the bialgebra \(O(\text{Mat}_n) = \mathbb{k}[X_{ij} | 1 \leq i, j \leq n]\) (Example 9.13) cannot have an antipode, because the determinant (9.28) is a group-like element \(D \in G(O(\text{Mat}_n))\) that is not invertible. The first two examples remedy this deficiency in two ways.

**Example 9.20** (Polynomial functions on \(\text{GL}_n(\mathbb{k})\)). First, let us consider an arbitrary \(B \in \text{BAlg}_\mathbb{k}\) having a group-like element \(g \in GB\) that is not invertible in \(B\). If \(g \in \mathcal{Z}(B)\) or, more generally, if \(g\) is a normal element of \(B\) in the sense that \(gB = Bg\), then we may consider the classical localization \(B[g^{-1}]\); see [87] for noncommutative rings of fractions. The counit and comultiplication of \(B\) extend uniquely to algebra maps \(\varepsilon \colon B[g^{-1}] \rightarrow \mathbb{k}\) and \(\Delta \colon B[g^{-1}] \rightarrow (B \otimes B)((g \otimes g)^{-1}) \rightarrow B[g^{-1}] \otimes B[g^{-1}]\) making \(B[g^{-1}]\) a \(\mathbb{k}\)-bialgebra. Applying these remarks with \(B = O(\text{Mat}_n)\) and \(g = D\), the determinant, we obtain the bialgebra

\[
O(\text{GL}_n) = O(\text{GL}_n(\mathbb{k})) \overset{\text{def}}{=} O(\text{Mat}_n)[D^{-1}]
\]

This bialgebra does in fact have an antipode, and hence it is a Hopf algebra. To see this, let us consider the “generic matrix” \(X := (X_{ij})_{i,j} \in \text{Mat}_n(B)\); so \(D = \text{det} X\).
With $C = (C_{ij}) \in \text{Mat}_n(B)$ denoting the matrix of cofactors of $X$, the equation $XC^T = D1_{n\times n}$ holds in $\text{Mat}_n(B)$. Therefore, $X^{-1} = D^{-1}C^T$ in $\text{Mat}_n(B[D^{-1}])$. We define the antipode by

$$SX = X^{-1}.$$ 

In more detail, define an algebra map $S: B = \mathbb{k}[X_{ij} \mid i, j] \to B[D^{-1}]$ by $S_{ij} = D^{-1}C_{ji}$ and note that $SD = D^{-1}$; this follows from the equation

$$(SX)X = X(SX) = 1_{n\times n}$$

in $\text{Mat}_n(B[D^{-1}])$, which implies $1 = \det(SX)\det X = (SD)D$. Thus, $S$ extends uniquely to an algebra endomorphism of $B[D^{-1}]$. We leave it to the reader to check that, with comultiplication (9.26) and counit (9.27), equation (9.37) is equivalent to the defining property (9.29) of the antipode.

**Example 9.21** (Polynomial functions on $SL_n(\mathbb{k})$). Again, start with an arbitrary bialgebra $B$ and a group-like element $g \in GB$. Then $(e, g-1) = 0$ and $\Delta(g-1) = (g-1) \otimes g + 1 \otimes (g-1)$. Thus, the ideal $(g-1)$ of $B$ is a bi-ideal. Furthermore, if $B$ has an antipode, $S$, then $S(g-1) = (1-g)g^{-1}$. So $(g-1)$ is a Hopf ideal and the quotient $B/(g-1)$ inherits a Hopf algebra structure. Specializing to $B = O(\text{GL}_n)$ and $g = D$, we obtain the Hopf algebra

$$O(\text{SL}_n) = O(\text{SL}_n(\mathbb{k})) \overset{\text{def}}{=} O(\text{GL}_n)/(D-1) \cong O(\text{Mat}_n)/(D-1)$$

The underlying algebras of the above Hopf algebras $O(\text{GL}_n)$ and $O(\text{SL}_n)$ are both commutative. Noncommutative examples of bialgebras and Hopf algebras can be constructed by “quantizing” bialgebras whose underlying algebra structure is commutative. This is illustrated in our next example.

**Example 9.22** (Quantum matrices). We will alter the multiplication of the bialgebra $O(\text{Mat}_n) = \mathbb{k}[X_{ij} \mid 1 \leq i, j \leq n]$ of polynomial functions on $n \times n$-matrices (Example 9.13) in such a way that the resulting algebra is noncommutative while still retaining the structure of a bialgebra. In order to do so in a reasonable manner, we take into account that $\text{Mat}_n$ acts on affine $n$-space $\mathbb{k}^n$, viewed as column or row vectors, by left and right matrix multiplication. The left action yields a $\mathbb{k}$-linear map $\text{Mat}_n \otimes \mathbb{k}^n \to \mathbb{k}^n$ along with its dual, $(\mathbb{k}^n)^* \to (\text{Mat}_n \otimes \mathbb{k}^n)^* \equiv \text{Mat}_n^* \otimes (\mathbb{k}^n)^*$. As in §9.3.3, the latter map in turn lifts uniquely to a map $O(\mathbb{k}^n) \to O(\text{Mat}_n) \otimes O(\mathbb{k}^n)$ in $\text{Alg}_{\mathbb{k}}$, where $O(\mathbb{k}^n) = \text{Sym}(\mathbb{k}^n)^* = \mathbb{k}[x_1, \ldots, x_n]$. Thus, at the level of algebras of polynomial functions, the left multiplication operation of $\text{Mat}_n$ on $\mathbb{k}^n$ corresponds to an algebra map that is explicitly given by

$$O(\mathbb{k}^n) \to O(\text{Mat}_n) \otimes O(\mathbb{k}^n)$$

$$x_i \mapsto \sum_j X_{ij} \otimes x_j$$

(9.38)
Similarly, right matrix multiplication corresponds to the algebra map

$$O(\mathbb{k}^n) \rightarrow O(\mathbb{k}^n) \otimes O(\text{Mat}_n)$$

\(x_i \rightarrow \sum_j x_j \otimes X_{ji}\)

Adopting a viewpoint promulgated by Manin \[143\], \[144\], we quantize \(O(\text{Mat}_n)\) by first quantizing affine \(n\)-space \(\mathbb{k}^n\) in the following manner. Fixing a parameter \(q \in \mathbb{k}^\times\), we define quantum affine \(n\)-space over \(k\) to be the following algebra:

$$O_q(\mathbb{k}^n) \overset{\text{def}}{=} \mathbb{k} \langle x_1, \ldots, x_n \rangle / (x_i x_j - q x_j x_i \mid i < j)$$

For \(n = 2\), we obtain the familiar quantum plane (e.g., Exercise 1.1.15 and Example 1.24). In order to arrive at the desired quantization of \(O(\text{Mat}_n)\), which will be denoted by \(O_q(\text{Mat}_n)\), we use algebra generators \(X_{ij}\) as before and stipulate exactly the relations needed so that (9.38) and (9.39) define maps \(O_q(\mathbb{k}^n) \rightarrow O_q(\text{Mat}_n) \otimes O_q(\mathbb{k}^n)\) and \(O_q(\mathbb{k}^n) \rightarrow O_q(\mathbb{k}^n) \otimes O_q(\text{Mat}_n)\) in \(\text{Alg}_k\). We leave to the reader the burden of verifying that the relations in question can be spelled out as follows. For each \(2 \times 2\)-submatrix \(\begin{pmatrix} a & b \\ c & d \end{pmatrix}\) of the generic matrix

$$X = \begin{bmatrix} X_{11} & \ldots & X_{1n} \\ \vdots & \ddots & \vdots \\ X_{n1} & \ldots & X_{nn} \end{bmatrix}$$

we have the following relations (Exercise 9.3.13):

$$ab = q ba \quad ac = q ca \quad bc = cb \quad bd = q db \quad cd = q dc \quad ad - da = (q - q^{-1})bc$$

It is a simple matter to check that the “old” counit \(\langle \epsilon, X_{ij} \rangle = \delta_{ij}\) and comultiplication \(\Delta X_{ij} = \sum_k X_{ik} \otimes X_{kj}\) of \(O(\text{Mat}_n)\) also respect the above relations of \(O_q(\text{Mat}_n)\) and that \(O_q(\text{Mat}_n)\) becomes a bialgebra in this way (Exercise 9.3.13). However, \(O_q(\text{Mat}_n)\) is not a Hopf algebra, because there is a non-invertible group-like element: the so-called quantum determinant

$$D_q = \sum_{s \in S_n} (-q)^{\ell(s)} X_{1s(1)} X_{2s(2)} \cdots X_{ns(n)}$$

Here, \(\ell(s)\) is the length of the permutation \(s\) (Example 7.10). The reader is referred to the monograph \[34\] by Brown and Goodearl for the fact that \(D_q\) is a central group-like element of \(O_q(\text{Mat}_n)\). Accepting this, we may proceed as in Examples 9.20
9.3. Bialgebras and Hopf Algebras

and 9.21 and define bialgebras, called quantum $GL_n$ and quantum $SL_n$, by

$$O_q(GL_n) = O_q(GL_n([k])) \equiv O_q(Mat_n)[D^{-1}_q]$$

and

$$O_q(SL_n) = O_q(SL_n([k])) \equiv O_q(Mat_n)/(D_q - 1)$$

It can be shown $O_q(GL_n)$ and $O_q(SL_n)$ are in fact Hopf algebras; the definition of the antipode uses so-called “quantum minors” in place of the cofactors employed in (9.37). Details and a wealth of further information, including multi-parameter generalizations of the foregoing, can be found in [34].

Finally, let us turn to finite-dimensional Hopf algebras. The Hopf algebras $H_{n,q}$ ($n = 2, 3, \ldots$) in our next example were originally constructed by Taft [195]. The salient feature of $H_{n,q}$ is that its antipode has order $2n$, viewed as an element of $GL(H_{n,q})$. All Hopf algebras that we have considered thus far, in particular group algebras and enveloping algebras, are involutory: their antipodes have order $\leq 2$.

**Example 9.23** (Taft algebras). Consider the quantum plane $O_q([k]^2) = [k][g, x]$ with defining relation $gx = q xg$, where we choose $q \in [k]^*$ to be a root of unity of order $n \geq 2$. This algebra can be equipped with a bialgebra structure as follows:

$$\Delta g = g \otimes g \quad \langle \epsilon, g \rangle = 1$$
$$\Delta x = x \otimes 1 + g \otimes x \quad \langle \epsilon, x \rangle = 0$$

The relations $\Delta g \Delta x = q g \Delta x g$ and $\langle \epsilon, g \rangle \langle \epsilon, x \rangle = q \langle \epsilon, x \rangle \langle \epsilon, g \rangle$ are readily checked—so $\Delta$ and $\epsilon$ do indeed define algebra maps—and the coassociativity and counit identities $(\Delta \otimes \Id) \circ \Delta = (\Id \otimes \Delta) \circ \Delta$ and $(\epsilon \otimes \Id) \circ \Delta = \Id = (\Id \otimes \epsilon) \circ \Delta$ are also straightforward to verify by evaluation on the generators $g$ and $x$ of $O_q([k]^2)$. Thus, $O_q([k]^2)$ is a $[k]$-bialgebra. It requires more care to ascertain that the ideal $(g^n - 1, x^n)$ of $O_q([k]^2)$ is a bi-ideal, but this is indeed the case (Exercise 9.3.14). Thus, we may define the bialgebra

$$H_{n,q} \equiv O_q([k]^2)/(g^n - 1, x^n) = [k][g, x]/(gx - q xg, g^n - 1, x^n)$$

To determine the dimension of $H_{n,q}$, recall that the standard monomials $g^i x^j$ form a $[k]$-basis of $O_q([k]^2)$ (Exercise 1.1.15); so $O_q([k]^2) \equiv [k][g] \otimes [k][x]$ in $\text{Vect}_k$. Since $g^n - 1$ and $x^n$ are both central elements of $O_q([k]^2)$, we have $(g^n - 1, x^n) \equiv (g^n - 1) \otimes [k][x] + [k][g] \otimes (x^n)$ under this isomorphism. Hence, we obtain an isomorphism $H_{n,q} \equiv [k][g]/(g^n - 1) \otimes [k][x]/(x^n)$ in $\text{Vect}_k$, which shows that

$$\dim_k H_{n,q} = n^2.$$
Allowing ourselves to write $g, x \in H_{n,q}$ for the images of the generators of $O_q(k^2)$, a $k$-basis of $H_{n,q}$ is given by the standard monomials $g^i x^j$ $(0 \leq i, j < n)$. It remains to show that $H_{n,q}$ has an antipode, $S$, and that its order is $2n$. Since $g$ is group-like, $S$ must satisfy $Sg = g^{-1} = g'^{-1}$ by (9.30). Furthermore, the formula for $\Delta$ together with (9.29) forces $0 = \langle e, x \rangle 1 = Sx + (Sg)x$. Therefore, we must have $Sx = -g^{-1}x$.

Defining $Sg$ and $Sx$ by these equations, one readily checks the relations $(Sg)^n = 1$, $(Sx)^n = 0$ and $Sx Sg = qSg Sx$ in $H_{n,q}$. So $g' := (Sg)^{op}$, $x' := (Sx)^{op} \in H_{n,q}^{op}$ satisfy $g'' = 1$, $x'' = 0$ and $g'x' = q x'g'$, and hence $S$ gives a well-defined algebra map $H_{n,q} \rightarrow H_{n,q}^{op}$. The requisite convolution identity $S \ast \text{Id} = u \circ e = \text{Id} * S$ is easily verified by evaluation on the generators $g$ and $x$, which implies that it holds for all elements of $H_{n,q}$. Hence, $S$ is the desired antipode for $H_{n,q}$. Finally, $S^2 x = q^{i}x$, proving that $S$ has order $2n$ as claimed.

The Hopf algebra $H_{2,1}$, for char $k \neq 2$, was constructed earlier by Sweedler and is known as the **Sweedler algebra**: it is the unique non-commutative and non-cocommutative Hopf algebra of dimension 4. This algebra was already considered in Exercise 2.1.13.

**Exercises for Section 9.3**

**9.3.1** (Preservation of antipode and counit). For $H, K \in \text{HopfAlg}_k$, show:

(a) If $f : H \rightarrow K$ is a map in $\text{Alg}_k$ such that $(f \otimes f) \circ \Delta_H = \Delta_K \circ f$, then $\varepsilon_K \circ f = \varepsilon_H$ and $S_K \circ f = f \circ S_H$; so $f$ is a map in $\text{HopfAlg}_k$. (Show that $S_K \circ f$ and $f \circ S_H$ are both inverses for $f$ in the convolution algebra $\text{Hom}_k(H, K)$.)

(b) For any commutative $k$-algebra $R$, the subset $\text{Hom}_\text{Alg}_k(H, R) \subseteq \text{Hom}_k(H, R)$ is a subgroup of the group of units $\text{Hom}_k(H, R)^\times$.

**9.3.2** (Finite-dimensional Hopf algebras). Let $H \in \text{HopfAlg}_k$ be finite dimensional. Use the fact that the convolution algebra $\text{End}_k(H)$ is finite dimensional (and hence $\text{Id}_H \in \text{End}_k(H)$ is algebraic) to prove the following results of Nichols [156]:

(a) Every subbialgebra of $H$ is stable under the antipode $S$.

(b) Every proper ideal $I$ of $H$ such that $\Delta I \subseteq H \otimes I + I \otimes H$ is stable under $S$ and satisfies $\langle e, I \rangle = 0$; so $I$ is a Hopf ideal.

**9.3.3** (Group algebras of finite abelian groups). Let $kG$ be the group algebra of a finite abelian group $G$ over a field $k$ containing a root of unity of order equal to the exponent of $G$. Show that the Hopf algebra $kG$ is self-dual: $kG \cong (kG)^\ast$ as Hopf algebras.

**9.3.4** (Hopf subalgebras and Hopf ideals of group algebras). For an arbitrary group algebra $kG$, show:
(a) The Hopf subalgebras of \( \mathbb{k}G \) are exactly the subgroup algebras \( \mathbb{k}H \) for subgroups \( H \leq G \).

(b) The Hopf ideals of \( \mathbb{k}G \) are exactly the relative augmentation ideals \( \mathbb{k}G(\mathbb{k}N)^\circ \), where \( N \) is a normal subgroup of \( G \) (Exercise 3.3.6).

9.3.5 (Hopf ideals of enveloping algebras). Assume that \( \text{char} \mathbb{k} = 0 \). Using the fact that, for any \( H \in \text{HopfAlg}_\mathbb{k} \), the subalgebra that is generated by the primitive elements of \( H \) is isomorphic to the enveloping algebra of the Lie algebra \( LH \) (§9.3.2), show that the Hopf ideals of \( UG \) are exactly the relative augmentation ideals \( (Ug)a = a(Ug) \), where \( a \) is an ideal of \( g \) (Exercise 5.4.5).

9.3.6 (Inverses in convolution algebras). Let \( A \in \text{Alg}_\mathbb{k}, B \in \text{BiAlg}_\mathbb{k}, \) and \( C \in \text{Coalg}_\mathbb{k} \).

(a) Assume that \( f \in \text{Hom}_\mathbb{k}(B, A) \) has a convolution inverse, say \( g \). Show that \( f \) is an algebra map if and only if \( g \) is an anti-algebra map, that is, \( g : B \rightarrow A^{\text{op}} \) is an algebra map.

(b) Assume that \( f \in \text{Hom}_\mathbb{k}(C, B) \) has a convolution inverse, \( g \). Show that \( f \) is a coalgebra map if and only if \( g \) is an anti-coalgebra map, that is, \( g : C \rightarrow B^{\text{cop}} \) is a coalgebra map.

9.3.7 (Involutory Hopf algebras). Let \( H \in \text{HopfAlg}_\mathbb{k} \) with antipode \( S \). Show that the following are equivalent:

(i) \( H \) is involutory, that is, \( S \circ S = \text{Id}_H \);
(ii) \( S(h_{(2)})h_{(1)} = \langle e, h \rangle 1_H \) for all \( h \in H \);
(iii) \( h_{(2)}S(h_{(1)}) = \langle e, h \rangle 1_H \) for all \( h \in H \).

In particular, all commutative and all cocommutative Hopf algebras are involutory.

9.3.8 (Detail for Proposition 9.15). Let \( H \in \text{HopfAlg}_\mathbb{k} \) with antipode \( S \). Show that \( S^\circ := S \big|_{H^\circ} \) is a convolution inverse for \( \text{Id}_H^\circ \).

9.3.9 (Representative functions of groups). Assume that \( \mathbb{k} \) is finite.

(a) Show that, for each \( f \in (\mathbb{k}G)^\circ \), there is some normal subgroup \( N \leq G \) that has finite index in \( G \) such that \( f \big|_N \) is constant.

(b) Conclude that if \( G \) is infinite and simple, then \( (\mathbb{k}G)^\circ = \mathbb{k}e \).

9.3.10 (Dedekind’s Lemma). Let \( A \in \text{Alg}_\mathbb{k} \). By (9.35) and Lemma 9.1, \( \text{Hom}_{\text{Alg}}(A, \mathbb{k}) \) consists of \( \mathbb{k} \)-linearly independent elements of \( A^\circ \). Deduce Dedekind’s Lemma from field theory: distinct field homomorphisms \( K \rightarrow \mathbb{k} \) are \( \mathbb{k} \)-linearly independent.

\(^{4}\) It is also true that the Hopf subalgebras of \( Ug \) are exactly subalgebras \( Uh \) for Lie subalgebras \( h \subseteq g \); see [153, 5.5.3 and 5.6.5], for example.
9.3.11 (The trigonometric Hopf algebra). Assume that $\text{char } \kappa \neq 2$ and put $H = \kappa \langle s, c \rangle / (s^2 + c^2 - 1) \in \text{Alg}_k$. Define

$$
\Delta c = c \otimes c - s \otimes s \\
\Delta s = c \otimes s + s \otimes c
$$

$$
\langle e, c \rangle = 1 \\
\langle e, s \rangle = 0
$$

$Sc = c \\
Ss = -s$

Show:

(a) $H$ is a Hopf algebra.

(b) $H \cong \mathcal{O}(\text{GL}_1(\kappa)) = \kappa[t, t^{-1}]$ if and only if there exists $i \in \kappa$ with $i^2 = -1$.

9.3.12 (Polynomial functions on $\text{GL}_n$). Using the notation of Example 9.20, check that the equation $X(SX)(SX) = (SX)X = 1_{n \times n}$ is equivalent to (9.29) for $\mathcal{O}(\text{GL}_n)$.

9.3.13 (Quantum affine space and quantum matrices). Using the notation of Example 9.22, show:

(a) The standard monomials $x_1^{i_1}x_2^{i_2} \ldots x_n^{i_n}$ form a $\kappa$-basis of $\mathcal{O}_q(\kappa^n)$.

(b) The relations (9.40) express exactly that formulas (9.38) and (9.39) define algebra maps $\mathcal{O}_q(\kappa^n) \rightarrow \mathcal{O}_q(\text{Mat}_n) \otimes \mathcal{O}_q(\kappa^n)$ and $\mathcal{O}_q(\kappa^n) \rightarrow \mathcal{O}_q(\kappa^n) \otimes \mathcal{O}_q(\text{Mat}_n)$.

(c) Defining $\Delta X_{ij} = \sum_k X_{ik} \otimes X_{kj}$ and $\langle e, X_{ij} \rangle = \delta_{ij}$ yields well-defined algebra maps $\Delta: \mathcal{O}_q(\text{Mat}_n) \rightarrow \mathcal{O}_q(\text{Mat}_n) \otimes \mathcal{O}_q(\text{Mat}_n)$ and $\varepsilon: \mathcal{O}_q(\text{Mat}_n) \rightarrow \kappa$. Moreover, $\mathcal{O}_q(\text{Mat}_n)$ becomes a bialgebra in this way.

9.3.14 (Taft algebras). (a) For $n \in \mathbb{N}$ and a parameter $q$, the $q$-binomial coefficients are defined by

$$
\binom{n}{i}_q = \frac{(1-q^n)(1-q^{n-1}) \cdots (1-q^{n-i+1})}{(1-q)(1-q^2) \cdots (1-q^i)} \\
(0 \leq i \leq n).
$$

Show that $\binom{n}{i}_q = q^i \binom{n-1}{i-1}_q + \binom{n-1}{i-1}_q$. In particular, $\binom{n}{n}_q \in \kappa[q]$.

(b) Let $A \in \text{Alg}_\kappa$ and let $a, b \in A$ be such that $ab = q ba$ with $q \in \kappa^\times$. Use (a) to prove the quantum binomial formula: $(a + b)^n = \sum_{i=0}^n \binom{n}{i}_q b^i a^{n-i}$. For $q$ a root of unity of order $n$, conclude that $(a + b)^n = a^n + b^n$.

(c) Use (b) to check that $(g^n - 1, x^n)$ is a bi-ideal of $\mathcal{O}_q(\kappa^2)$ (notation of Example 9.23).
Chapter 10

Representations and Actions

As in the special cases of group algebras and enveloping algebras, which featured in earlier chapters, a representation of a Hopf algebra $H$ is understood to be a representation of the underlying algebra of $H$. The additional Hopf data imbue the representation category $\text{Rep} H$ with extra structure: the counit $\varepsilon$ provides us with a distinguished “trivial” representation and a notion of invariants, the comultiplication $\Delta$ allows us to form tensor products of representations, and the antipode $S$ leads to duals of representations. The Grothendieck group $\mathcal{R}(H)$ of all finite-dimensional representations of $H$ thus carries a ring structure, which need not be commutative. The bialgebra structure of $H$ also makes it possible to consider left $H$-modules that are simultaneously left $H$-comodules in such a way that the action and coaction of $H$ acknowledge each other; this leads to the powerful concept of a left $H$-Hopf module. Finally, defining $H$-module algebras to be algebras in the tensor category $\text{Rep} H$, we obtain a meaningful notion of $H$-action on algebras, which permits us to venture into the territory of noncommutative invariant theory.

While much of this chapter is devoted to laying the foundations of representations of Hopf algebras and their (co)actions on algebras, some nontrivial applications may be found in Section 10.2. Throughout this chapter, we illustrate the abstract material developed here with a continued discussion of the special case of finite group algebras.

*In this chapter, unless otherwise specified, $H = (H, m, u, \Delta, \varepsilon, S)$ denotes an arbitrary Hopf algebra over the field $\mathbb{K}$, which can also be arbitrary.*
10.1. Representations of Hopf Algebras

Hopf algebras provide the proper context for material that was already covered twice in the more specialized settings of groups (§3.3.3) and Lie algebras (Section 5.5). In addition, there are now also comodule variants to consider. As in earlier chapters, we will write the algebra map \( H \to \text{End}_k(V) \) defining a representation \( V \in \text{Rep} H \) as \( h \mapsto h_V \) and the corresponding left \( H \)-module action \( H \otimes V \to V \) as \( h \otimes v \mapsto h_v \).

10.1.1. General Constructions

**Trivial (Co)actions and (Co)invariants.** The counit \( \varepsilon : H \to k \) gives rise to the so-called the **trivial representation** of \( H \),

\[ \mathbb{1} = k_{\varepsilon} \]

More generally, the \( H \)-action on a representation \( V \in \text{Rep} H \) is said to be trivial if the action map is given by \( \varepsilon \otimes \text{Id}_V : H \otimes V \to k \otimes V \to V \), that is, \( h.v = \langle \varepsilon, h \rangle v \) for all \( h \in H \), \( v \in V \). The \( k \)-subspace of \( H \)-invariants of an arbitrary \( V \in \text{Rep} H \) is defined to be the \( 1 \)-homogeneous component \( V(1) \):

\[ V^H \overset{\text{def}}{=} \{ v \in V \mid h.v = \langle \varepsilon, h \rangle v \text{ for all } h \in H \} \]

Note that this definition only refers to the counit and therefore makes sense for any algebra \( A \) with a distinguished augmentation \( \varepsilon : A \to k \).

Dually, the coaction of a right comodule \( M \in \text{Mod}^H \) is called **trivial** if it has the form \( \delta = \text{Id}_M \otimes u : M \to M \otimes k \to M \otimes H \); so \( \delta m = m \otimes 1 \) for all \( m \in M \). For an arbitrary \( M \in \text{Mod}^H \), one puts

\[ M^{\text{co}H} \overset{\text{def}}{=} \{ m \in M \mid \delta m = m \otimes 1 \} \]

This is a \( k \)-subspace of \( M \), called the space of \( H \)-coinvariants in \( M \).\(^1\) When \( M \) is viewed as left module over the algebra \( H^* \) as in Proposition 9.12 and (9.24), then the \( H^* \)-invariants with respect to \( u^* : H^* \to k \) coincide with the \( H \)-coinvariants:

\[ (10.1) \quad M^{H^*} = \{ m \in M \mid f.m = \langle f, 1 \rangle m \text{ for all } f \in H^* \} = M^{\text{co}H}. \]

Of course, invariants and coinvariants can also be defined, with the obvious notational adjustments, for right \( H \)-modules and left \( H \)-comodules, respectively. The above right-handed notations for invariants and coinvariants are being used for either side of the (co)action.

\(^1\)Note, however, that the term "coinvariants" also has a different meaning; see Exercise 10.1.1.
Tensor Products. Let \( V, W \in \text{Rep} H \). Then the tensor product \( V \otimes W \) becomes a representation of \( H \) via the algebra map
\[
H \xrightarrow{\Delta} H \otimes H \xrightarrow{\text{End}_k(V \otimes W)}.
\]
(1.51)

More generally, recall that a right coideal subalgebra of \( H \) is a subalgebra of \( K \subseteq H \) satisfying \( \Delta K \subseteq K \otimes H \). In this case, for any \( V \in \text{Rep} K \) and \( W \in \text{Rep} H \), the tensor product \( V \otimes W \) becomes a representation of \( K \) in the above manner. Explicitly, the \( K \)-action on \( V \otimes W \) is given by
\[
h_{V \otimes W} = (h_{(1)})_V \otimes (h_{(2)})_W \quad (h \in K)
\]
(10.2)

The tensor product in \( \text{Rep} H \) is clearly functorial: if \( f : V \to V' \) and \( g : W \to W' \) are maps in \( \text{Rep} H \), then \( f \otimes g : V \otimes W \to V' \otimes W' \) is also a map in \( \text{Rep} H \) for the action (10.2). Thus, we obtain a bifunctor, exact in both variables,
\[
\cdot \otimes \cdot : \text{Rep} H \times \text{Rep} H \to \text{Rep} H.
\]

By virtue of the counit laws, we have the following natural isomorphisms, for any \( V \in \text{Rep} H \),
\[
\mathbb{1} \otimes V \cong V \cong V \otimes \mathbb{1}.
\]
(10.3)

Coassociativity of \( \Delta \) implies associativity of the tensor product: the associativity isomorphism (B.8) in \( \text{Vect}_k \) restricts to a natural isomorphism in \( \text{Rep} H \), for any \( U, V, W \in \text{Rep} H \),
\[
U \otimes (V \otimes W) \cong (U \otimes V) \otimes W.
\]
(10.4)

As in §B.1.3, this isomorphism is usually implicitly understood when considering iterated tensor products such as \( U \otimes V \otimes W \) in \( \text{Rep} H \). In this way, \( \text{Rep} H \equiv H \text{Mod} \) becomes a tensor category. If \( H \) is cocommutative, then the switch map \( \tau : V \otimes W \rightleftharpoons W \otimes V \), \( v \otimes w \mapsto w \otimes v \), is an isomorphism in \( \text{Rep} H \). All this has been observed earlier for representations of groups and Lie algebras, where coassociativity and cocommutativity of \( \Delta \) are obvious. However, other than for groups and Lie algebras, \( V \otimes W \) and \( W \otimes V \) need not be isomorphic in \( \text{Rep} H \) for noncommutative Hopf algebras; see Exercises 10.1.2 and 10.1.6(a).

Tensor products can also be constructed for comodules: if \( M, N \in \text{Mod}^H \) have coactions \( \delta_M \) and \( \delta_N \), resp., then we may define a coaction \( \delta \) of \( H \) on \( M \otimes N \) by
\[
M \otimes N \xrightarrow{\delta_M \otimes \delta_N} M \otimes H \otimes N \otimes H \xrightarrow{\text{Id} \otimes \text{Id} \otimes m} M \otimes N \otimes H,
\]
where (2.3) switches the factors in positions 2 and 3. In terms of elements,
\[
\delta(m \otimes n) = m_{(0)} \otimes n_{(0)} \otimes m_{(1)}n_{(1)} \quad (m \in M, n \in N)
\]
(10.5)
The analogs of (10.3) and (10.4) hold for the tensor product in \( \text{Mod}^H \) as well, replacing 1 by \( 1' = \mathbb{k} \) with the trivial right \( H \)-comodule structure in (10.3). Note also that the material on trivial (co)actions and tensor products discussed so far works in exactly the same way for any bialgebra rather than a Hopf algebra.

**Changing Sides.** The antipode \( S \) of \( H \) allows us to switch sides. For instance, pullback along the algebra map \( S: H \to H^{op} \) (Proposition 9.14) as in §1.2.2 gives a functor

\[
S^*: \text{Mod}_H \equiv H^{op}\text{Mod} \longrightarrow H\text{Mod}.
\]

Similarly, pushing forward a given coaction \( \delta_M: M \to M \otimes H^{cop} \) for \( M \in \text{Mod}_H \) along \( S \), we obtain the coaction \( (\text{Id} \otimes S) \circ \delta_M: M \to M \otimes H^{cop} \). This results in a functor

\[
S_*: \text{Mod}^H \longrightarrow \text{Mod}^{H^{cop}} \equiv H\text{Mod}.
\]

Of course, left \( H \)-(co)modules can be turned into right (co)modules in this way as well. In fact, \( S \) is often bijective, in which case the above functors give equivalences of categories; this is certainly the case for all involutory Hopf algebras, such as group algebras and enveloping algebras, and we shall soon see that it also holds for all finite-dimensional Hopf algebras (Theorem 10.9). With modules, we shall continue to work mostly on the left, because \( H\text{Mod} \equiv \text{Rep} H \), but right-sided and two-sided features will also play a role.

**Adjoint Actions.** Any bimodule \( M \in H\text{Mod}_H \) can be viewed as a left module over the algebra \( H \otimes H^{op} \) in the usual way. Precomposing the resulting map \( H \otimes H^{op} \to \text{End}_k(M) \) with the map \( (\text{Id} \otimes S) \circ \Delta: H \to H \otimes H \to H \otimes H^{op} \) yields a map \( H \to \text{End}_k(M) \) in \( \text{Alg}_k \) and hence a representation of \( H \). This representation will be called the adjoint representation associated to \( M \) and denoted by \( M_{ad} \).

Explicitly, using a dot-less notation for the left and right \( H \)-actions on \( M \) coming from the original bimodule structure, the adjoint \( H \)-action on \( M = M_{ad} \) is given by

\[
(10.6) \quad h.m = h(1)mS(h(2)) \quad (h \in H, m \in M).
\]

The assignment of \( M_{ad} \) to \( M \) clearly yields an exact functor \( H\text{Mod}_H \to \text{Rep} H \). The \( H \)-invariants of \( M_{ad} \) turn out to be identical to the center of the bimodule \( M \), which is defined by \( \mathcal{Z}M = \{m \in M \mid hm = mh \text{ for all } h \in H\} \):

**Lemma 10.1.** \( M_{ad}^H = \mathcal{Z}M \) for any \( M \in H\text{Mod}_H \).

**Proof.** The computation \( hm = h(1)mS(h(2)) = h(1)S(h(2))m = \langle e, h \rangle m \) for \( m \in \mathcal{Z}M \) and \( h \in H \) shows that \( \mathcal{Z}M \subseteq M_{ad}^H \). To prove the reverse inclusion, first observe that, for any \( m \in M \) and \( h \in H \),

\[
(10.7) \quad hm = (h(1),m)h(2).
\]
For, the right-hand side evaluates to \( h_{(1)}mS(h_{(2)})h_{(3)} = h_{(1)}m(\epsilon, h_{(2)}) = hm \). If \( m \in M_{ad}^H \), then (10.7) gives \( hm = (\epsilon, h_{(1)})mh_{(2)} = mh \), proving the lemma. □

**Homomorphisms.** For given \( V, W \in \text{Rep } H \), the space \( \text{Hom}_k(V, W) \) carries a natural \((H, H)\)-bimodule structure given by

\[
hf k = hW \circ f \circ kV \quad (h \in H, f \in \text{Hom}_k(V, W))
\]

Representations of this form have already seen extensive use in the context of group algebras (3.29) and enveloping algebras (5.25). As in these special cases, we will dispense with the subscript \( \cdot_{ad} \) for \( \text{Hom}_k(V, W) \), because the new \( H \)-action (10.8) will be more important than the original left \( H \)-action coming from the \((H, H)\)-bimodule structure. Generalizing our earlier formulae (3.30) and (5.26), Lemma 10.1 yields the following description of the \( H \)-invariants of \( \text{Hom}_k(V, W) \):

\[
\text{Hom}_k(V, W)^H = \text{Hom}_H(V, W).
\]

It is easy to see that the bifunctor \( \text{Hom}_k \) for \( k \)-vector spaces (§B.3.2) restricts to a bifunctor, exact in both variables and contravariant in the first variable,

\( \text{Hom}_k(\cdot, \cdot) : (\text{Rep } H)^{op} \times \text{Rep } H \to \text{Rep } H \).

**Dual Representations.** Taking \( W = \mathbb{1} \) in the foregoing, the dual vector space \( V^* = \text{Hom}_k(V, k) \) becomes a representation of \( H \). Equation (10.8) can now be written in the following form, as in the proof of Lemma 3.20 and in (5.28):

\[
h_{V^*} = (S(h)V)^* \quad (h \in H)
\]

If \( V^* \cong V \) in \( \text{Rep } H \), then the representation \( V \) is called **self-dual**; this forces \( V \) to be finite dimensional by the Erdős-Kaplansky Theorem. For example, the trivial representation \( \mathbb{1} \) is self-dual by virtue of the identity \( \epsilon = \epsilon \circ S \). By our remarks about \( \text{Hom}_k(\cdot, \cdot) \) above, duality gives an exact contravariant functor,

\( \cdot^* : \text{Rep } H \to \text{Rep } H \).

**Basic Isomorphisms.** As for groups and Lie algebras, it is straightforward to verify that the canonical embedding (B.18) in \( \text{Vect}_k \) is actually a homomorphism in \( \text{Rep } H \):

\[
W \otimes V^* \xrightarrow{\pi} \text{Hom}_k(V, W)
\]

\[
w \otimes f \xrightarrow{\psi} (v \mapsto \langle f, v \rangle w)
\]
Recall that the image of this embedding consists of the finite-rank homomorphisms. Hence, (10.11) an isomorphism in $\text{Rep} H$ if at least one of $V, W$ is finite dimensional. In this case, the $\text{Vect}_k$-isomorphism (B.21) also is an isomorphism in $\text{Rep} H$:

\[
W^* \otimes V^* \xrightarrow{\sim} (V \otimes W)^*
\]

(10.12)

The switch in the order of $V$ and $W$ is necessitated by the fact that $S$ is an anti-homomorphism for the coalgebra structure of $H$ (Proposition 9.14): denoting the image of $g \otimes f$ in $(V \otimes W)^*$ by $(g \otimes f)'$, we have

\[
\langle h.(g \otimes f)', v \otimes w \rangle = \langle (g \otimes f)', S(h)_1.v \otimes S(h)_2.w \rangle
\]

(10.10)

\[
= \langle (g \otimes f)', S(h)_2.v \otimes S(h)_1.w \rangle
\]

(10.12)

\[
= \langle f, S(h)_2.v \otimes S(h)_1.w \rangle
\]

(10.8)

\[
= \langle (h.(g \otimes f))', v \otimes w \rangle.
\]

Lemma 10.2. Assume that the antipode $S$ of $H$ is bijective and let $\sigma$ denote the inverse of $S^2 = S \circ S \in \text{Aut}_{\text{Alg}}(H)$ (Proposition 9.14). Then:

(a) For any $V \in \text{Rep}_{\text{fin}} H$, there is a $\text{Rep} H$-isomorphism $V^{**} \cong \sigma V$, the $\sigma$-twist of $V$ (1.24). If $S^2$ is an inner automorphism of $H$, then $V^{**} \cong V$.

(b) Duality gives a bijection $\cdot^* : \text{Irr}_{\text{fin}} H \to \text{Irr}_{\text{fin}} H$.

Proof. (a) For any $V \in \text{Rep} H$, we have the canonical $k$-linear embedding (B.22),

\[
\mu : V \hookrightarrow V^{**}, \quad \mu(v) = (f \mapsto \langle f, v \rangle).
\]

For any $h \in H$ and $v \in V$, we calculate

\[
\langle \mu(S^2(h), v), f \rangle = \langle f, S^2(h).v \rangle = \langle S(h).f, v \rangle
\]

(10.10)

\[
= \langle \mu(v), S(h).f \rangle = \langle h.\mu(v), f \rangle.
\]

(10.10)

Thus, $\mu(S^2(h), v) = h.\mu(v)$. If $V$ is finite dimensional, then $\mu$ is bijective and the map $v \mapsto \sigma v$ is an isomorphism $V^{**} \xrightarrow{\sigma} V$ in $\text{Rep} H$:

\[
h.\mu(v) = \mu(S^2(h), v) \mapsto \sigma(S^2(h), v) = h.\sigma v.
\]

(1.24)

If $\sigma$ is inner, then $\sigma V \cong V$ in $\text{Rep} H$ (Exercise 1.2.3), and so $V^{**} \cong V$.

(b) For $S \in \text{Irr}_{\text{fin}} H$, we need to show that $S^*$ is irreducible. So let $f : U \hookrightarrow S^*$ be a nonzero monomorphism in $\text{Rep} H$. Then $f^* : S^{**} \to U^*$ is a nonzero epimorphism in $\text{Rep} H$. Since $S^{**} \cong \sigma S$ by (a) and $\sigma S$ is irreducible (Exercise 1.2.3), it follows that $f^*$ is an isomorphism. Hence $f$ is an isomorphism as well, proving
that $S^*$ is irreducible. Finally, twisting gives a bijection $\alpha : \text{Irr}_{\text{fin}} H \overset{\sim}{\rightarrow} \text{Irr}_{\text{fin}} H$ (Exercise 1.2.3), and hence duality also gives a bijection. \hfill \square

10.1.2. Hopf Modules

The Definition. A left $H$-Hopf module, by definition, is a $k$-vector space $M$ that is simultaneously a left $H$-module and a left $H$-comodule in such a way that the following two equivalent conditions are satisfied:

(i) the module action $\mu : H \otimes M \rightarrow M$ is an $H$-comodule map;

(ii) the comodule coaction $\delta : M \rightarrow H \otimes M$ is an $H$-module map.

Here, $H \otimes M$ carries the left $H$-comodule structure coming from the left regular comodule structure of $H$ (Example 9.10) and the left analog of the tensor coaction (10.5). Similarly, the $H$-module structure of $H \otimes M$ comes from the regular module structure $H = H_{\text{reg}}$ and the tensor action (10.2). Explicitly, writing the $H$-action on $M$ as $h \otimes m \mapsto h.m$ and the coaction as $m \mapsto m_{(-1)} \otimes m_{(0)}$ as usual, the $H$-coaction on $H \otimes M$ is given by $h \otimes m \mapsto h_{(1)} m_{(-1)} \otimes h_{(2)} \otimes m_{(0)}$ and the $H$-action by $h.(k \otimes m) = h_{(1)} k \otimes h_{(2)} m$ for $h, k \in H$ and $m \in M$. For the equivalence of (i) and (ii), note that both conditions amount to the following identity:

\[(10.13) \quad (h.m)_{(-1)} \otimes (h.m)_{(0)} = h_{(1)} m_{(-1)} \otimes h_{(2)} m_{(0)} \quad (h \in H, m \in M).
\]

Defining homomorphism of $H$-Hopf modules to be $k$-linear maps that are simultaneously $H$-module and $H$-comodule maps, we obtain a category.

\[H^H_{\text{Mod}}.\]

Variants. There is an obvious right-sided variant of left $H$-Hopf modules: they are “the same” as left $H^\text{biop}$-Hopf modules (Exercise 10.1.10). Moreover, Hopf modules may certainly be defined in the same way for arbitrary bialgebras. However, the Structure Theorem for Hopf Modules below, which is of fundamental importance, does depend on the existence of an antipode.

There is also a more general relative version of the notion of a Hopf module, which will play a role in §12.4.5. Specifically, let $K$ be any left coideal subalgebra of $H$; recall that this means that $K$ is a $k$-subalgebra of $H$ satisfying $\Delta K \subseteq H \otimes K$. Then a left $(H, K)$-Hopf module is a $k$-vector space $M$ that is a left $K$-module and a left $H$-comodule such that (10.13) holds for all $h \in K$ and $m \in M$. In this way, we obtain a category, $K^H_{\text{Mod}}$. As usual, we obtain the right-sided variant by passing to $^\text{biop}$. (Exercise 10.1.10).

---

2Hopf modules are called “bimodules” in Abe [1], which correlates well with the term “bialgebra”; however, bimodules already have a different meaning for us.
Some Examples. In short, the Structure Theorem for Hopf Modules states that every $M \in \text{HMod}$ is free as left $H$-module (Corollary 10.5). Since direct sums of left $H$-Hopf modules are again left $H$-Hopf modules in the obvious way, our first example below shows that, conversely, every free (left or right) $H$-module can be equipped with the structure of a (left or right) $H$-Hopf module.

Example 10.3 (Regular Hopf modules). The Hopf algebra $H$ becomes an object of $\text{HMod}$ by giving $H$ the left regular $H$-module structure, coming from the multiplication of $H$, along with the left regular comodule structure that comes from the comultiplication $\Delta$ of $H$ (Example 9.10). The identity (10.13) is clear, because $\Delta$ is an algebra map. Using the right regular module and comodule structures instead, $H$ becomes a right $H$-Hopf module.

Example 10.4 (Tensor products). For $M \in \text{HMod}$ and $V \in \text{HMod}$, the tensor product $M \otimes V$ becomes a left $H$-Hopf module by using the tensor module structure (10.2) and the coaction $\delta = \delta_M \otimes \text{Id}_V$:

$$h.(m \otimes v) = h_1(m) \otimes h_2(v)$$

and

$$\delta(m \otimes v) = m_{(-1)} \otimes m_{(0)} \otimes v$$

for $h \in H$, $m \in M$ and $v \in V$. Similarly, for $W \in \text{HMod}$, we may equip $M \otimes W$ with a left $H$-Hopf module structure with the tensor comodule structure and with module action $\mu_M \otimes \text{Id}_W$:

$$h.(m \otimes w) = h.m \otimes w$$

and

$$\delta(m \otimes w) = m_{(-1)}w_{(-1)} \otimes m_{(0)} \otimes w_{(0)}$$

In case $W$ has the trivial $H$-coaction, the above coaction becomes $\delta = \delta_M \otimes \text{Id}_W$. It is perhaps tempting to expect that, when $M$ and $N$ are both in $\text{HMod}$, the tensor module and comodule structures on $M \otimes N$ together should result in a left $H$-Hopf module structure, but that is not generally the case (Exercise 10.1.9).

The Structure Theorem and Consequences. For $M \in \text{HMod}$, we may view $M^{\text{co}H} = \{m \in M \mid \delta m = 1 \otimes m\}$ as a left $H$-comodule with trivial coaction. Forming the tensor product with the regular Hopf module $H \in \text{HMod}$ as in Example 10.4, we obtain $H \otimes M^{\text{co}H} \in \text{HMod}$; the $H$-action and $H$-coaction are given by $m \otimes \text{Id}$ and $\Delta \otimes \text{Id}$, respectively.

Structure Theorem for Hopf Modules. For any $M \in \text{HMod}$, the following map is an isomorphism in $\text{HMod}$:

$$H \otimes M^{\text{co}H} \xrightarrow{\sim} M$$

$$\begin{array}{ccc}
h \otimes m & \mapsto & h.m \\
\omega & \mapsto & \omega \\
\end{array}$$

Proof. Let $\Phi : H \otimes M^{\text{co}H} \rightarrow M$, $h \otimes m \mapsto h.m$, denote the map in the theorem; it is clearly a map in $\text{HMod}$. Furthermore, letting $\delta$ denote the coaction of $M$ and
\( \delta' = \Delta \otimes \text{Id} \) the coaction of \( H \otimes M^{coH} \), we compute for \( h \in H \) and \( m \in M^{coH} \),

\[
(\delta \circ \Phi)(h \otimes m) = (h.m)_{(-1)} \otimes (h.m)_{(0)} = h_{(1)}m_{(-1)} \otimes h_{(2)}.m_{(0)}
\]

(10.14)

\[
= h_{(1)} \otimes h_{(2)}.m = ((\text{Id} \otimes \Phi) \circ \delta')(h \otimes m).
\]

Thus, \( \Phi \) is also an \( H \)-comodule map, and hence it is a map in \( \text{Proj}_H \).

Next, define \( \pi \in \text{End}_H(M) \) by \( \pi m = S(m_{(-1)}).m_{(0)} \). The following calculation shows that the image of \( \pi \) consists of \( H \)-coinvariants in \( M \):

\[
\delta(\pi m) = (S(m_{(-1)}).m_{(0)})_{(-1)} \otimes (S(m_{(-1)}).m_{(0)})_{(0)}
\]

(10.13)

\[
= S(m_{(-1,1)}).m_{(0,1)} \otimes S(m_{(-1,2)}).m_{(0,2)}
\]

\[
= S(m_{(-2,1)}).m_{(-1)} \otimes S(m_{(-2,2)}).m_{(0)}
\]

\[
= S(m_{(-2)}).m_{(-1)} \otimes S(m_{(-3)}).m_{(0)}
\]

\[
= (\epsilon, m_{(-1)} \otimes S(m_{(-2)}).m_{(0)}
\]

\[
= 1 \otimes \pi m.
\]

Hence, we may consider the map

\[
\Psi : \quad M \longrightarrow H \otimes M^{coH}
\]

\[
\begin{array}{ccc}
\psi & \quad & \psi \\
\uparrow & \quad & \uparrow \\
m & \longmapsto & m_{(-1)} \otimes \pi(m_{(0)}
\end{array}
\]

To finish the proof, it suffices to show that \( \Psi \) and \( \Phi \) are inverse to each other. To this end, we compute for \( h \in H \) and \( m \in M^{coH} \),

\[
(\Phi \circ \Psi)(m) = m_{(-1)}(\pi(m_{(0)})) = m_{(-2)}S(m_{(-1)}).m_{(0)}
\]

\[
= (\epsilon, m_{(-1)}).m_{(0)} = m
\]

and, using (10.14) twice,

\[
(\Psi \circ \Phi)(h \otimes m) = (h.m)_{(-1)} \otimes \pi((h.m)_{(0)}) = h_{(1)} \otimes \pi(h_{(2)}.m)
\]

\[
= h_{(1)} \otimes S((h_{(2)}).m)_{(-1)})(h_{(2)}.m)_{(0)}
\]

\[
= h_{(1)} \otimes S(h_{(2)}).h_{(3)}.m = h \otimes m.
\]

This completes the proof.

\[\square\]

**Corollary 10.5.**

(a) If \( M \in \text{Proj}_H \), then \( M \) is free as left \( H \)-module.

(b) Let \( V, F \in \text{Rep} H \) with \( F \) free. Then, with the tensor action (10.2), \( F \otimes V \in \text{Rep} H \) is free. If \( P \in \text{Proj}_H \), then \( P \otimes V \in \text{Proj}_H \).

**Proof.** (a) The Structure Theorem for Hopf Modules gives an isomorphism \( M \cong H \otimes M^{coH} \cong H_{\text{reg}}^{\otimes d} \) in \( \text{Rep} H \), where \( d = \dim_k M^{coH} \). Thus, \( M \) is free as left \( H \)-module.
(b) We have seen that $F \in H^H \text{Mod}$ and $F \otimes V \in H^H \text{Mod}$ (Examples 10.3 and 10.4). Therefore, $F \otimes V$ is free by (a). If $P$ is projective, say $P$ is a direct summand of the free left $H$-module $F$, then $P \otimes V$ is a direct summand of $F \otimes V$, which is free. Hence $P \otimes V$ is projective. \hfill \Box

We remark that the map $\Psi$ in the proof of the Structure Theorem for Hopf Modules is the composite $\Psi = \varphi \circ \delta$, where $\varphi$ the following isomorphism in $\text{Rep } H$, a special case of the tensor product formula (Exercise 10.1.7):

$$
\psi: \quad H \otimes M = (\text{Ind}_k^H \mathbb{k}) \otimes M \xrightarrow{\sim} \text{Ind}_k^H (\mathbb{k} \otimes M) = H \otimes (\text{Res}_k^H M)
$$

Exercises for Section 10.1

10.1.1 ("Coinvariants"). Let $V \in \text{Rep } H$. Extending the definition for group algebras (Exercise 3.3.3), we may define the space of $H$-"coinvariants" of $V$ by $V_H = V/H^+ \cdot V$, where $H^+ = \text{Ker } \varepsilon$.

(a) Assuming that $\mathcal{S}$ is surjective, show that there is a natural isomorphism $(V_H)^* \xrightarrow{\sim} (V^*)^H$ in $\text{Veck}_k$. (This generalizes the isomorphism in Exercise 3.3.9.)

(b) Let $M \in H \text{Mod}_H$ and assume that $H$ is involutory. Show that $(M_{\text{ad}})_H = M/[H, M]$, where $[H, M]$ is the subspace of $M$ that is generated by the elements $[h, m] := hm - mh$ for $h \in H, m \in M$. (This generalizes Exercise 5.5.4.)

10.1.2 (Almost cocommutative Hopf algebras). Following Drinfeld [61], a Hopf algebra $H$ is called almost cocommutative if there exists $R \in (H \otimes H)^\times$ such that $(\tau \circ \Delta) h = R(\Delta h)R^{-1}$ for all $h \in H$, where $\tau$ is the switch map; so $h_{(2)} \otimes h_{(1)} = R(h_{(1)} \otimes h_{(2)})R^{-1}$. For cocommutative $H$, we may take $R = 1 \otimes 1$.

(a) Assuming $H$ to be almost cocommutative, show that the map $V \otimes W \to W \otimes V, v \otimes w \mapsto R^{-1}(w \otimes v)$, is an isomorphism in $\text{Rep } H$ for any $V, W \in \text{Rep } H$.

(b) Consider the Sweedler algebra $H_{2-1} = \mathbb{k} 1 \oplus \mathbb{k} g \oplus \mathbb{k} x \oplus \mathbb{k} g x$ with $g^2 = 1, x^2 = 0, xg = -gx$ and char $\mathbb{k} \neq 2$ (Example 9.23). The coalgebra structure and antipode are given by $\Delta g = g \otimes g$, $\langle \varepsilon, g \rangle = 1$, $Sg = g$ and $\Delta x = x \otimes 1 + g \otimes x$, $\langle \varepsilon, x \rangle = 0$, $Sx = -gx$. Show that $H_{2-1}$ is almost cocommutative (but not cocommutative), with $R = \frac{1}{2}(1 \otimes 1 + 1 \otimes g + g \otimes 1 - g \otimes g)$.

10.1.3 (Duality, induction and coinduction). Let $K \to H$ be a map in $\text{HopfAlg}_k$ and assume that the antipode of $H$ is invertible. Show that, for any $V \in \text{Rep } K$, there is a natural isomorphism $\text{Coind}_K^H V^* \cong (\text{Ind}_K^H V)^*$ in $\text{Rep } H$. (This generalizes an earlier isomorphism for group algebras in Exercise 3.3.10.)

10.1.4 (Dual representations). Let $U, V, W \in \text{Rep } H$. 

(a) Show that the isomorphism $\text{Hom}_k(U \otimes V, W) \xrightarrow{\cong} \text{Hom}_k(U, \text{Hom}_k(V, W))$ in (B.15) is in fact an isomorphism in $\text{Rep} \, H$. In particular, if $V$ or $W$ is finite dimensional, then $\text{Hom}_k(U \otimes V, W) \cong \text{Hom}_k(U, W \otimes V^*)$ in $\text{Rep} \, H$.

(b) Assuming $H$ to be cocommutative, show that the map $\text{Hom}_k(V, W) \to \text{Hom}_k(W^*, V^*)$ $f \mapsto f^*$, is a monomorphism in $\text{Rep} \, H$; it is an isomorphism if $V$ and $W$ are finite dimensional.

10.1.5 (The trace map). For $V \in \text{Rep}_{\text{fin}} \, H$, show that trace: $\text{End}_k(V) \to k = 1$ is a map in $\text{Rep} \, H$ if and only if $S(h_{(2)})_V(h_{(1)})_V = \langle e, h \rangle \text{Id}_V$ for all $h \in H$. In particular, this always holds if $H$ is involutory. Conversely, if all traces are $H$-equivariant and $\cap_{V \in \text{Rep}_{\text{fin}}} \, H \text{Ker} \, V = 0$, then $H$ must be involutory.

10.1.6 (1-dimensional representations, winding automorphisms, and twists). Let $k_\alpha$ denote the 1-dimensional representation of $H$ that is given by $\alpha \in \text{Hom}_{\text{Alg}}(H, k) = G(H^*)$ (9.35). Generalizing earlier facts for groups (Exercise 3.3.11), show:

(a) $k_\alpha \otimes k_\beta \cong k_{\alpha \beta}$ and $(k_\alpha)^* \cong k_{\alpha^{-1}}$. (Recall that $\alpha^{-1} = S^*(\alpha)$ in $G(H^*)$.)

(b) The map $\tilde{\alpha} : H \to H$, $h \mapsto \alpha(h_{(1)})h_{(2)}$, is an algebra automorphism of $H$; it is called a winding automorphism of $H$.

(c) The $\tilde{\alpha}$-twist (1.24) of $V$ is isomorphic to $k_{\alpha^{-1}} \otimes V$. Conclude that twisting gives an action of the group $G(H^*)$ on the set $\text{Irr} \, H$. (See Exercise 1.2.3.)

10.1.7 (Tensor product formula for Hopf algebras). Let $K$ be a right coideal subalgebra of $H$ and let $V \in \text{Rep} \, H$, $W \in \text{Rep} \, K$. Generalizing an earlier formula for groups (Exercise 3.3.13), prove the isomorphism $(\text{Ind}_K^H \, W) \otimes V \cong \text{Ind}_K^H \, (W \otimes V)$ in $\text{Rep} \, H$, with the tensor action (10.2) of $K$ on $W \otimes V$ and of $H$ on $(\text{Ind}_K^H \, W) \otimes V$.

10.1.8 (Hopf module identities). Let $M \in \text{Mod}_H^H$. Viewing the left $H$-coaction on $M$ as a right $H^*$-module action as in the right-sided version of Proposition 9.12(a), show that (10.13) implies the following identities, for $h \in H$, $m \in M$ and $f \in H^*$:

$$h \cdot (m, f) = (h_{(2)}, m) \cdot (f \circ S(h_{(1)})_H) \quad \text{and} \quad (h, m) \cdot f = h_{(2)} \cdot (m \cdot (f \circ (h_{(1)})_H)).$$

10.1.9 (Tensor products of Hopf modules). Let $H$ be a Hopf algebra with an element $1 \neq g \in GH$. Viewing $H \in \text{Mod}_H^H$ (Example 10.3), show that $H \otimes H$, when equipped with the tensor left $H$-module structure (10.2) and the left-handed version of the tensor $H$-comodule structure (10.5), is not a left $H$-Hopf module.

10.1.10 (Right Hopf modules). For a Hopf algebra $H$ and a right coideal subalgebra $K \subseteq H$, define a right $(H, K)$-Hopf module to be a $k$-vector space $M$ that is a right $K$-module and a right $H$-comodule such that $(m, k)_0 \otimes (m, k)_1 = m_{(0)} \otimes m_{(1)} k_{(2)}$ holds for all $k \in K$ and $m \in M$. Show that the resulting category $\text{Mod}_K^H$ of right $(H, K)$-Hopf modules is equivalent to $\text{Mod}_{K_{\text{fin}}}^{H_{\text{fin}}}$. 

10.1.11 (Hopf submodules of $H$). Show that 0 and $H$ are the only Hopf submodules of $H \in \text{Mod}_H^H$ (Example 10.3): they are the only $k$-subspaces of $H$ that are simultaneously a left ideal and a left coideal of $H$. 

10.2. First Applications

Lest the reader grow tired of the preponderance of abstract material in this chapter and the previous one, we postpone the development of further generalities in order to offer some applications of the foregoing, focusing for the most part on Hopf algebras that are finite dimensional. Chapter 12 will dig deeper into the structure of finite-dimensional Hopf algebras.

10.2.1. Finiteness of Hopf Algebras

In this subsection, we use the Structure Theorem for Hopf Modules to prove that, under certain circumstances, the Hopf algebra $H$ is “finite,” that is, $\dim_k H < \infty$.

**Proposition 10.6.** If a Hopf algebra $H$ has a nonzero finite-dimensional left or right ideal, then $H$ is itself finite dimensional.

**Proof.** First, let us assume that there exists a finite-dimensional left ideal $L$ of $H$. Giving $H$ the regular left $H$-Hopf module structure (Example 10.3), let $L'$ denote the $H$-subcomodule of $H$ that is generated by $L$. The Finiteness Theorem for comodules (§9.2.2) implies that $L'$ is finite dimensional. Therefore, it suffices to show that $L'$ is again a left ideal of $H$; for, then $L' \subseteq H^\text{HMod}$ and so $L'$ is free over $H$ by the Structure Theorem for Hopf Modules, forcing $H$ to be finite dimensional.

We will show more generally that, for any $M \in H^\text{HMod}$ and any $H$-submodule $V \subseteq M$, the $H$-subcomodule $V' \subseteq M$ that is generated by $V$ is also a left $H$-submodule of $M$. To see this, recall that the left $H$-coaction on $M$ can be expressed as a right $H^*$-module action and $V' = V.H^*$ (Proposition 9.12). The identity (10.13) then gives the following commutation rule between the actions of $H$ and $H^*$, whose verification we leave to the reader (Exercise 10.1.8):

$$h.(m.f) = (h_{(2)}.m).(f \circ S(h_{(1)})) \quad (h \in H, m \in M, f \in H^*).$$

If $m \in V$, then the right side belongs to $V' = V.H^*$, whence $H.V' \subseteq V'$ as desired.

This proves the proposition for left ideals. The result for right ideals follows by symmetry or by applying the foregoing to $H^{biop}$.

Proposition 10.6 implies in particular that any semisimple Hopf algebra $H$ must be finite dimensional, because the trivial representation $1$ gives rise to a 1-dimensional left ideal of $H$. However, there is an even stronger result. Recall that an algebra $A$ is said to be **left artinian** if $A$ satisfies the descending chain condition for left ideals; likewise for the right-sided version (Exercise 1.2.10). Finite-dimensional algebras are of course left and right artinian, but the converse is far from true due to the existence of infinite-dimensional division algebras. In light of this, the following theorem of Liu and Zhang [134] is remarkable. The theorem gives in particular that if a group algebra $\mathbb{k}G$ is left or right artinian, then $G$ must
be finite. This is an older result from the 1960s due to Connell [47]; see also [163, Theorem 10.1.1].

**Theorem 10.7.** Any left or right artinian Hopf algebra is in fact finite dimensional.

To prove the theorem, first note that, by symmetry, it suffices to consider the left artinian case. In view of Proposition 10.6 and the existence of the trivial representation, Theorem 10.7 is a consequence of the following purely ring-theoretic lemma.

**Lemma 10.8.** Let \( A \in \text{Alg}_k \) be left artinian. If \( A \) has a nonzero finite-dimensional representation, then \( A \) has a nonzero finite-dimensional right ideal.

**Proof.** We may assume that the given finite-dimensional representation of \( A \) is irreducible. Taking the kernel of this representation, we obtain a cofinite prime ideal of \( A \), which we shall denote by \( P \). Now we invoke the following standard facts about left artinian rings; see [125, (4.15) and Exercise 4 in §10]:

(a) \( A_{\text{reg}} \in \text{Rep} A \) has finite length (Hopkins-Levitzki Theorem).

(b) All prime ideals of \( A \) are maximal.

It follows from (a) that some finite product \( P_1P_2\ldots P_n \) of (not necessarily distinct) primes \( P_i \) of \( A \) must be zero: just consider the kernels of the factors in a composition series of \( A_{\text{reg}} \). Choosing \( n \) minimal, we may assume that omitting any \( P_i \) would result in a nonzero product. Our initially chosen prime \( P \) must occur among the factors \( P_i \), because \( 0 = P_1P_2\ldots P_n \subseteq P \) implies \( P_i \subseteq P \) for some \( i \) and so \( P = P_i \) by (b). Thus, letting \( X \) and \( Y \) denote the (possibly empty) subproducts of \( P_1P_2\ldots P_n \) before and after \( P \), respectively, we have \( XPY = 0 \) but \( XY \neq 0 \). Fact (a) also tells us that \( A \) is left noetherian (Exercise 1.2.10); so \( Y \) is finitely generated as left ideal of \( A \). Since \( P \) is cofinite, it follows that \( \dim_k Y/PY < \infty \) and so \( Y = V + PY \) for some finite-dimensional \( k \)-subspace \( V \). If \( x \in X \), then \( xV = xV + xPY = XY \) is a finite-dimensional right ideal of \( A \). Since \( XY \neq 0 \) for some \( x \), the lemma follows and hence Theorem 10.7 is proved as well. \( \square \)

### 10.2.2. Some Properties of Finite-Dimensional Hopf Algebras

The following theorem lists some fundamental properties of a finite-dimensional Hopf algebra \( H \). Part (a) implies that all free left \( H \)-modules are self-dual in \( \text{Rep} H \); this part certainly requires \( H \) to be finite dimensional. Part (c), on the other hand, evidently also holds for all involutory Hopf algebras; in particular all group algebras and all enveloping algebras of Lie algebras have a bijective antipode. However, there are Hopf algebras whose antipode fails to be bijective [196].

**Theorem 10.9.** Let \( H \in \text{HopfAlg}_k \) be finite dimensional. Then:

(a) The regular representation \( H_{\text{reg}} \) is self-dual: \( (H_{\text{reg}})^* \equiv H_{\text{reg}} \) in \( \text{Rep} H \).
(b) The space of invariants of $H_{\text{reg}}$ is 1-dimensional: $H_{\text{reg}}^H \cong \mathbb{1}$.

(c) The antipode $S$ of $H$ is bijective.

**Proof.** Put $M = (H_{\text{reg}})^*$, equipped with the standard left $H$-action (10.10).

(a) We will show that $M \in H^H_{\text{Mod}}$. It will then follow from the Structure Theorem for Hopf Modules that $M \cong H \otimes M^{\text{co}H}$, which forces $\dim_k M^{\text{co}H} = 1$ by a dimension count, and so $M \cong H_{\text{reg}}$ in $\text{Rep} H$ as desired.

The requisite coaction $M \rightarrow H \otimes M, \ f \mapsto f_{(-1)} \otimes f_{(0)}$, comes from the right regular module of the algebra $H^*$. Indeed, by the right-sided version of Proposition 9.12(b) and (9.25), we may write

\[
(10.15) \quad fg = \langle g, f_{(-1)} \rangle f_{(0)} \quad (f, g \in H^*).
\]

We need to verify (10.13): $(h.f)_{(-1)} \otimes (h.f)_{(0)} = h_{(1)} f_{(-1)} \otimes h_{(2)} f_{(0)}$ for all $f \in H^* = M$ and $h \in H$. By (10.15), this condition can be written as

\[
(h.f)g = \langle g, h_{(1)} f_{(-1)} \rangle h_{(2)} f_{(0)}.
\]

To check this equality, we write $h.f = f \rightleftharpoons h^{\text{sh}}$ using the standard right $H$-action — on $H^*$ (2.19). For $f, g \in H^*$ and $h, k \in H$, we have $\langle (fg) - h, k \rangle = \langle f, h_{(1)} k_{(1)} \rangle \langle g, h_{(2)} k_{(2)} \rangle$, which translates into the following identity:

\[
(10.16) \quad (fg) - h = (f - h_{(1)})(g - h_{(2)}).
\]

With this, the desired equality follows from the computation

\[
(h.f)g = (f - h^{\text{sh}})g
= (f - S(h_{(3)}))(g - h_{(1)} S(h_{(2)}))
= \langle f, h_{(1)} \rangle - S(h_{(2)})
= (10.16) \langle g - h_{(1)}, f_{(-1)} \rangle f_{(0)} - S(h_{(2)})
= (10.15) \langle g, h_{(1)}, f_{(-1)} \rangle h_{(2)} f_{(0)}.
\]

(b) By the right-handed version of (10.1), $M^{\text{co}H}$ is the space of invariants of the right regular module of the Hopf algebra $H^*$:

\[
M^{\text{co}H} = \{ f \in M \mid f_{(-1)} \otimes f_{(0)} = 1 \otimes f \}
= \{ f \in H^* \mid fg = \langle g, 1 \rangle f \text{ for all } g \in H^* \}.
\]

But $\dim_k M^{\text{co}H} = 1$, as we have remarked in the proof of (a), and we may replace the Hopf algebra $H$ in the foregoing by $H^*$ to conclude that the invariants of the right regular module of $H^{\ast \ast} \cong H$ are 1-dimensional. Further replacing $H$ by $H^{\text{bi} \text{op}}$, we obtain the same conclusion for the left regular module, $H_{\text{reg}}$. 
(c) The identity \( h.f = f - S h \) for \( f \in M \) shows that \( \text{Ker } S \subseteq \text{Ker } M \). But \( M \cong H_{\text{reg}} \) by (a) and so \( \text{Ker } M = \text{Ker } H_{\text{reg}} = 0 \), proving that \( S \) is bijective. \(\square\)

**Corollary 10.10.** If \( H \in \text{HopfAlg}_k \) is finite dimensional, then every \( S \in \text{Irr } H \) embeds into \( H_{\text{reg}} \).

**Proof.** Choosing an epimorphism \( H_{\text{reg}} \rightarrow S \) in \( \text{Rep } H \), we obtain a monomorphism \( S^* \hookrightarrow (H_{\text{reg}})^* \cong H_{\text{reg}} \) by Theorem 10.9. Since every irreducible representation of \( H \) has the form \( S^* \) for some \( S \in \text{Irr } H \) (Lemma 10.2), the corollary follows. \(\square\)

**10.2.3. Inner Faithful Representations**

To set the stage for the concept of inner faithfulness, let us briefly revisit representations of a Lie algebra \( g \). Recall that \( V \in \text{Rep } g \) is said to be \( g \)-faithful if the Lie algebra map \( g \rightarrow \text{gl}(V) \) is injective; this generally does not imply that the unique extension \( U_g \rightarrow \text{End}_k(V) \) in \( \text{Alg}_k \) is injective (Proposition 5.28). Similarly, a representation \( V \) of a group \( G \) over a field \( k \) is traditionally called faithful if the group homomorphism \( G \rightarrow \text{GL}(V) \) is injective. For clarity, we will say that \( V \) is \( G \)-faithful in this case, because we may also view \( V \) as a representation of the group algebra \( kG \), giving another notion of faithfulness (Section 1.2) that is generally stronger than \( G \)-faithfulness. For example, the standard permutation representation \( M_n \) of the symmetric group \( S_n \) (§3.2.4) is always \( S_n \)-faithful, but \( M_n \) is not faithful for \( kS_n \) (\( n \geq 4 \)) for dimension reasons.

To pinpoint the connection between the conflicting notions of faithfulness, we consider, for an ideal \( I \) of an arbitrary Hopf algebra \( H \), the unique largest Hopf ideal of \( H \) that is contained in \( I \):

\[
\mathcal{H} I \overset{\text{def}}{=} \text{the sum of all Hopf ideals of } H \text{ that are contained in } I
\]

If \( I = \text{Ker } V \) is the kernel of a representation \( V \in \text{Rep } H \), then we will refer to \( \mathcal{H} \text{Ker } V \) as the **Hopf kernel** of \( V \). The representation \( V \) is called **inner faithful** if \( \mathcal{H} \text{Ker } V = 0 \).

**Example 10.11** (Inner faithfulness for groups and Lie algebras). If \( H = kG \) is a group algebra, then Hopf ideals of \( kG \) are exactly the ideals of the form \( \langle g - 1 \mid g \in N \rangle \), where \( N \) is a normal subgroup of \( G \) (Exercise 9.3.4). It follows that \( \mathcal{H} I \) is the ideal of \( kG \) that is generated by the set \( I \cap \{ g - 1 \mid g \in G \} \). Thus, a representation \( V \in \text{Rep } kG \) is \( G \)-faithful if and only if \( \mathcal{H} \text{Ker } V = 0 \), whereas faithfulness in the general sense of Section 1.2 means that \( \text{Ker } V = 0 \). Thus, inner faithfulness for \( kG \) is the same as \( G \)-faithfulness. Similarly, for an ideal \( I \) of the enveloping algebra \( H = U_g \), the intersection \( I \cap g \) generates \( \mathcal{H} I \) (Exercise 9.3.5). Hence, \( g \)-faithfulness of \( V \in \text{Rep } U_g \) is equivalent to inner faithfulness for \( U_g \).
Tensor Powers. We now offer a version of Proposition 5.28 for finite-dimensional Hopf algebras. For a given $V \in \text{Rep} H$, we may consider the representation $TV = \bigoplus_{n \geq 0} V^\otimes n \in \text{Rep} H$. Here, $V^\otimes 0 = 1$ is the trivial representation, $V^\otimes 1 = V$, and the tensor power $V^\otimes n$ for $n \geq 2$ is regarded as a representation of $H$ by iterating (10.2); so $H$ acts via the map $\Delta_n : H \to H^\otimes n$ ($\S$9.1.3).

Lemma 10.12. Let $H \in \text{HopfAlg}_k$ be finite dimensional. Then $\mathcal{H} \ker V = \ker TV$ for any $V \in \text{Rep} H$.

Proof. For brevity, let us put $K_V = \ker TV = \bigcap_{n \geq 0} \ker V^\otimes n$. If $J$ is any Hopf ideal of $H$, then $\Delta_n J \subseteq \sum_{i=0}^{n-1} H^\otimes i \otimes J \otimes H^\otimes (n-i)$. Thus, if $J \subseteq \ker V$, then it follows that $J \subseteq \ker V^\otimes n$ for all $n$ and so $J \subseteq K_V$. It remains to show that $K_V$ is indeed a Hopf ideal of $H$. First, the inclusion $K_V \subseteq \ker V^\otimes 0 = \ker \varepsilon$ gives $\langle \varepsilon, K_V \rangle = 0$. Next, observe that $TV \otimes TV \in \text{Rep} H$ is the direct sum of the various tensor powers $V^\otimes n$ for $n \geq 0$ like $TV$ (but with $V^\otimes n$ occurring $n + 1$ times). Thus, $\ker(TV \otimes TV) = K_V$. On the other hand, the general equality $\ker_{A \otimes A'}(W \otimes W') = (\ker W) \otimes A' + A \otimes (\ker W')$ for any two representations $W, W'$ of arbitrary $A, A' \in \text{Alg}_k$ gives $\ker_{H \otimes H}(TV \otimes TV) = K_V \otimes H + H \otimes K_V$. Therefore, $\Delta K_V \subseteq K_V \otimes H + H \otimes K_V$. Finally, since $H$ is assumed finite dimensional, the inclusion $S K_V \subseteq K_V$ is automatically satisfied by Exercise 9.3.2. \(\square\)

Rieffel’s Theorem. A classical result of Burnside [36, Theorem IV on p. 299], which we have already referred to in $\S$3.5.2, states that if $G$ is a finite group and $V$ is a $G$-faithful complex representation of $G$, then every irreducible complex representation of $G$ occurs as a constituent of some tensor power $V^\otimes n$. The following theorem of Rieffel [175] generalizes Burnside’s Theorem while also giving a more transparent proof—the original proof of relied on character theory.

Theorem 10.13. Let $H \in \text{HopfAlg}_k$ be finite dimensional and let $V \in \text{Rep} H$ be inner faithful. Then the regular representation $H_{\text{reg}}$ embeds into some finite direct sum of tensor powers $V^\otimes n$ and every irreducible representation of $H$ embeds into some tensor power $V^\otimes n$.

Proof. Lemma 10.12 gives that $\ker TV = 0$. Finite dimensionality of $H$ also guarantees that $\ker TV = \bigcap_i \{ h \in H \mid h x_i = 0 \text{ for all } i \}$ for finitely many $x_i \in TV$ that may clearly be chosen homogeneous, say $x_i \in V^\otimes n_i$. Hence, we obtain a monomorphism $H_{\text{reg}} \hookrightarrow \bigoplus_i V^\otimes n_i$, $h \mapsto (h x_i)$, in $\text{Rep} H$. This proves the first assertion of the theorem. The second assertion is now clear in view of Corollary 10.10. \(\square\)
10.2.4. Non-Divisibility Results

In this subsection, we consider the condition that \( \text{char} \, k \) does not divide \( \dim_k V \) for certain \( 0 \neq V \in \text{Rep}_\text{fin} H \). Our main tool will be the trace map (B.23):

\[
\text{trace} : \quad \text{End}_k(V) \xrightarrow{(10.11)} V \otimes V^* \xrightarrow{\sim} 1
\]

The focus will be on involutory Hopf algebras. The reason for this is the following lemma; see Exercise 10.1.5 for a converse.

**Lemma 10.14.** If \( H \in \text{HopfAlg}_k \) is involutory, then \( \text{trace} : \text{End}_k(V) \rightarrow 1 \) is a morphism in \( \text{Rep} H \) for every \( V \in \text{Rep}_\text{fin} H \).

**Proof.** We already know from (10.11) that \( \text{End}_k(V) \cong V \otimes V^* \) in \( \text{Rep} H \). Applying \( S \) to the counit identity \( S(h(1))h(2) = (\varepsilon, h)1_H \) gives \( S(h(2))h(1) = (\varepsilon, h)1_H \), since \( S^2 = 1d_H \) by hypothesis. Therefore,

\[
\text{trace} (h.(v \otimes f)) = \text{trace}(h(1).v \otimes h(2).f) = \langle h(2).f, h(1).v \rangle = \langle f, S(h(2))h(1) \rangle = \langle f, h \rangle \text{trace} (v \otimes f).
\]

This shows that the trace map is \( H \)-equivariant, proving the lemma. \( \square \)

In addition to the trace map, we also have the following map in the other direction, which is clearly always a morphism in \( \text{Rep} H \):

\[
1_V : \quad 1 \longrightarrow \text{End}_k(V)
\]

The composite \( \text{trace} \circ 1_V \) is multiplication by \( \dim_k V \) on \( 1 = k \); so

(10.18) \( \text{trace} \circ 1_V \neq 0 \iff \text{char} \, k \nmid \dim_k V \).

**Complete Reducibility of Tensor Products.** Part (b) of the following proposition is due to Serre [182, Theorem 2.4], originally for the case of finite group algebras, but the proof given works in general.

**Proposition 10.15.** Let \( H \in \text{HopfAlg}_k \) be involutory and let \( V \in \text{Rep}_\text{fin} H \). Then:

(a) If \( \text{End}_H(V) = k \) and \( V \otimes V^* \) is completely reducible, then \( \text{char} \, k \nmid \dim_k V \).

(b) Assume that \( \text{char} \, k \nmid \dim_k V \). If \( W \otimes V \) or \( V \otimes W \) is completely reducible for \( W \in \text{Rep} H \), then \( W \) is completely reducible.
Proof. (a) Our hypothesis that $\text{End}_k(V) \cong V \otimes V^*$ is completely reducible forces trace: $\text{End}_k(V) \rightarrow \mathbb{1}$ to split in $\text{Rep} \, H$. Therefore, $\text{End}_k(V)^H = \text{End}_H(V) = k \cdot \text{Id}_V$ maps onto $\mathbb{1}$ and (10.18) gives that $\text{char} \, k \nmid \dim_k V$.

(b) Let $U$ be a subrepresentation of $W$ and let $\mu: U \hookrightarrow W$ denote the inclusion map. Consider the following commutative diagram:

\[
\begin{array}{ccc}
U & \overset{(10.3)}{\sim} & U \otimes \mathbb{1} \\
\mu \downarrow & & \downarrow \text{Id}_U \otimes \text{Id}_V \\
W & \overset{(10.3)}{\sim} & W \otimes \mathbb{1} \\
\end{array}
\]

The morphism $1_V: \mathbb{1} \rightarrow \text{End}_k(V)$ is split by the map $(\dim_k V)^{-1}$ trace in $\text{Rep} \, H$. Hence, we also have a splitting of $\text{Id}_U \otimes 1_V$ in $\text{Rep} \, H$; call it $\sigma$. If $W \otimes V$ is completely reducible, then the embedding $\mu \otimes \text{Id}_V: U \otimes V \hookrightarrow W \otimes V$ is split in $\text{Rep} \, H$, and hence $\mu \otimes \text{Id}_V \otimes \text{Id}_{V^*}$ is split as well, say by $\tau$. Altogether, viewing the isomorphisms (10.3) as identifications, $\sigma \circ \tau \circ (\text{Id}_W \otimes 1_V)$ is a splitting of $\mu$ in $\text{Rep} \, H$, proving that $W$ is completely reducible. If $V \otimes W$ is completely reducible, then identify $V$ with $V^{**}$ in $\text{Rep} \, H$ (Lemma 10.2) and argue similarly using the maps $1_{V^*} \otimes \text{Id} \dashv \mathbb{1} \otimes \cdot \rightarrow V^* \otimes V \otimes \cdot$, to construct a splitting of $\mu$. □

Semisimplicity. We close this subsection with a semisimplicity criterion for involutory Hopf algebras from [135, Theorem 2.3].

Theorem 10.16. Let $H \in \text{HopfAlg}_k$ be involutory. Then $H$ semisimple if and only if there exists $P \in \text{Proj}_{\text{fin}} H$ such that $\text{char} \, k \nmid \dim_k P$.

Proof. Recall that $H$ is semisimple if and only if all representations of $H$ are projective (Exercise 2.1.2). In particular, $\mathbb{1}$ is projective in this case and, of course, $\text{char} \, k \nmid \dim_k \mathbb{1} = 1$.

Conversely, let $P \in \text{Proj}_{\text{fin}} H$ be such that $\text{char} \, k \nmid \dim_k P$. Then $\text{End}_k(P) \cong P \otimes P^*$ is projective as well (Corollary 10.5). Since trace: $\text{End}_k(P) \rightarrow \mathbb{1}$ splits in $\text{Rep} \, H$ by (10.18), $\mathbb{1}$ is a direct summand of $\text{End}_k(P)$ and so $\mathbb{1}$ is also projective. Consequently, every $V \in \text{Rep} \, H$ is projective by Corollary 10.5, because $V \cong \mathbb{1} \otimes V$ by (10.3). This shows that $H$ is semisimple. □

Corollary 10.17. Let $H \in \text{HopfAlg}_k$ be involutory.

(a) If $H$ is finite dimensional with $\text{char} \, k \nmid \dim_k H$, then $H$ is semisimple.

(b) If $H$ is semisimple, then $\text{char} \, k \nmid \dim_k V$ for every absolutely irreducible $V \in \text{Rep} \, H$.

Proof. For (a), just take $P = H_{\text{reg}}$ in Theorem 10.16. To prove (b), assume that $H$ is semisimple (and so in particular finite dimensional) and let $V \in \text{Rep} \, H$ be
absolutely irreducible. Then \( \text{End}_H(V) = \mathbb{k} \) (Proposition 1.36) and \( V \otimes V^* \in \text{Rep } H \) is completely reducible. Therefore, \( \text{char } \mathbb{k} \nmid \dim_\mathbb{k} V \) by Proposition 10.15(a).

Part (a) of the corollary applies in particular to group algebras \( \mathbb{k}G \) of finite groups \( G \) with \( \text{char } \mathbb{k} \nmid |G| \), giving the more substantial direction of Maschke’s Theorem (§3.4.1). We will later prove a stronger result; see Corollary 12.13. Part (b) is due to Larson [130, Theorem 2.8]. In Exercise 10.2.2 and §10.4.4, we will say more about the condition that \( \text{char } \mathbb{k} \nmid \dim_\mathbb{k} V \) for absolutely irreducible \( V \in \text{Rep } H \).

### 10.2.5. The Chevalley Property

A classical result due to Chevalley [43, Proposition 2 in Chap. IV §5] states that if \( V \) and \( W \) are any two finite-dimensional completely reducible representations of an arbitrary group \( G \) over a field \( \mathbb{k} \) with \( \text{char } \mathbb{k} = 0 \), then the tensor product \( V \otimes W \in \text{Rep } \mathbb{k}G \) is again completely reducible. The corresponding statement holds for Lie algebras as well. The proof of these results will be given in §11.6.5. In reference to Chevalley’s Theorem, and following [5], we will say that an arbitrary Hopf algebra \( H \) has the **Chevalley property** if the tensor product of any two completely reducible representations in \( \text{Rep}_\text{fin} H \) is again completely reducible. Thus, all group algebras and all enveloping algebras of Lie algebras have the Chevalley property if \( \text{char } \mathbb{k} = 0 \). We will also say that a given \( V \in \text{Rep } H \) has the **Chevalley property** if the representation \( TV \in \text{Rep } H \) is completely reducible or, equivalently, all tensor powers \( V \otimes^n \) are completely reducible. In particular, \( V \) itself must certainly be completely reducible in this case, but this is generally not sufficient for \( V \) to have the Chevalley property. Evidently, the Chevalley property for \( H \) is equivalent to the Chevalley property for all completely reducible \( V \in \text{Rep}_\text{fin} H \).

**Proposition 10.18.** Let \( H \in \text{HopfAlg}_\mathbb{k} \) be finite dimensional and let \( V \in \text{Rep } H \). Then:

- (a) \( V \) has the Chevalley property if and only if the Hopf algebra \( H/\mathcal{H} \text{Ker } V \) is semisimple.

- (b) \( H \) has the Chevalley property if and only if the Jacobson radical \( \text{rad } H \) is a Hopf ideal.

**Proof.** (a) By Lemma 10.12, \( TV \) is a faithful representation of \( H/\mathcal{H} \text{Ker } V \). Thus, if \( TV \) is completely reducible, then \( H/\mathcal{H} \text{Ker } V \) is semisimple (Exercise 1.4.4). Conversely, if \( H/\mathcal{H} \text{Ker } V \) is semisimple, then all its representations are completely reducible. Consequently, \( TV \in \text{Rep } H \) is completely reducible; so \( V \) has the Chevalley property.

(b) Clearly, the Chevalley property for \( H \) is equivalent to the Chevalley property for \( V = \bigoplus_{S \in \text{Irr } H} S \). By (a), this in turn is equivalent to \( H/\mathcal{H} \text{rad } H \) being semisimple, because \( \text{rad } H = \text{Ker } V \). Finally, the algebra \( H/\mathcal{H} \text{rad } H \) is semisimple.
if and only if \( \text{rad} \, H = \mathcal{H} \text{rad} \, H \) (Theorem 1.39), which means that \( \text{rad} \, H \) is a Hopf ideal.

**Example 10.19** (The Chevalley property for finite group algebras). Let \( H = \mathbb{k}G \) be the group algebra of a finite group \( G \). Put \( p = \text{char} \, \mathbb{k} \geq 0 \) and let \( O_p(G) \) denote the \( p \)-core of \( G \) (Exercise 3.4.6), understood to be \( \{1\} \) if \( p = 0 \). It follows from Exercises 3.4.6 and 9.3.4 that

\[
\mathbb{k}G/\mathcal{H} \text{rad} \, \mathbb{k}G \cong \mathbb{k}[G/O_p(G)].
\]

In particular, by Proposition 10.18 and Maschke’s Theorem (§3.4.1), \( \mathbb{k}G \) has the Chevalley property if and only if \( p \) does not divide the order of \( G/O_p(G) \) or, equivalently, \( G \) has a normal Sylow \( p \)-subgroup (understood to be \( \{1\} \) if \( p \nmid |G| \)).

**Exercises for Section 10.2**

10.2.1 (Cocommutative Hopf algebras: exterior and symmetric powers). Let \( H \in \text{HopfAlg} \) be cocommutative and let \( V, W, V_1, \ldots, V_n \in \text{Rep} \, H \). Prove:

(a) The \( H \)-action on \( V^\otimes n \) via \( \Delta_n \) commutes with the place permutation action \( S_n \subset V^\otimes n \) (3.64) and, for any \( s \in S_n \), the map \( V_1 \otimes \cdots \otimes V_n \mapsto V_{s^{-1}1} \otimes \cdots \otimes V_{s^{-1}n} \), \( v_1 \otimes \cdots \otimes v_n \mapsto v_{s^{-1}1} \otimes \cdots \otimes v_{s^{-1}n} \), is an isomorphism in \( \text{Rep} \, H \). Similarly, \( \Lambda V, \Lambda^n V \in \text{Rep} \, H \).

(b) The \( H \)-action on \( TV \) (§10.2.3) passes down to actions on \( \text{Sym} \, V \) and all homogeneous components \( \text{Sym}^n V \). Similarly, \( \Lambda V, \Lambda^n V \in \text{Rep} \, H \).

(c) If \( \text{char} \, \mathbb{k} \nmid n! \), then the symmetrizer \( \mathcal{S} \) (Lemma 3.36) yields an isomorphism \( \text{Sym}^n V \cong (V^\otimes n)(\mathcal{S}) \) in \( \text{Rep} \, H \). Similarly, \( \Lambda^n V \cong (V^\otimes n)(\text{sgn}) \) via the antisymmetrizer \( \mathcal{A} \).

(d) The standard isomorphism \( \text{Sym}^n(V \oplus W) \cong \bigoplus_{i+j=n} \text{Sym}^i V \otimes \text{Sym}^j W \) (Exercise 1.1.12) is an isomorphism in \( \text{Rep} \, H \); likewise for \( \Lambda^n \).

(e) Show that \( (\text{Sym}^n(V^*)^* \cong \text{Sym}^n(V^*) \) in \( \text{Rep} \, H \) provided \( \text{char} \, \mathbb{k} \nmid n! \) and always \( (\Lambda^n V)^* \cong \Lambda^n(V^*) \) in \( \text{Rep} \, H \).

(f) If \( \dim \mathbb{k} \, V = n \) then \( \Lambda^n V \cong \bigotimes_{i \leq i \leq n} \Lambda^i V \) for some \( \Lambda^i V \in \mathcal{G}(H^i) \). For example, if \( H = \mathbb{k}G \), then \( \Lambda^i V = \text{det} \, V \) by (3.32). For \( H = U_3 \) and \( x \in g \), we have \( \delta^x_V(x) = \chi^x_V(x) \).

10.2.2 (Dimension of irreducibles). Give an example of a (non-semisimple) finite-dimensional involutory \( H \in \text{HopfAlg} \) and an absolutely irreducible representation \( V \in \text{Rep} \, H \) such that \( \dim \mathbb{k} \, V \) is divisible by \( \text{char} \, \mathbb{k} \). (You can find some in earlier exercises about group algebras or Lie algebras.)

10.2.3 (Chevalley property). Let \( H \in \text{HopfAlg} \) be finite dimensional and let \( P \), denote projective covers (Theorem 2.7 and Exercise 2.1.7). Show:

(a) \( H \) has the Chevalley property if and only if \( P(V \otimes W) \cong PV \otimes W \) for all \( V, W \in \text{Irr} \, H \).
(b) If $H$ has the Chevalley property, then $PV \equiv P \mathbb{1} \otimes \text{head} V$ for any $V \in \text{Rep} H$.$^3$

**10.2.4 (Taft algebras).** Let $H = H_{n,q}$ denote the Taft algebra, where $q \in \mathbb{K}^\times$ is a root of unity of order $n \geq 2$ (Example 9.23). Generalizing Exercise 2.1.13, show:

(a) $H$ has the Chevalley property.

(b) $P \mathbb{1} \equiv H e$, where $e = \frac{1}{n} \sum_{i=0}^{n-1} g^i \in H$ (notation of Example 9.23).

(c) The Cartan matrix of $H$ is the $n \times n$ matrix with all entries equal to 1.

**10.3. The Representation Ring of a Hopf Algebra**

Picking up the thread from Section 10.1, we now return to the general development of the representation theory of an arbitrary Hopf algebra $H$. The spotlight in this section is on the Grothendieck group $\mathcal{R}(H)$ of all finite-dimensional representations of $H$, as introduced in §1.5.5 for arbitrary algebras. Owing to the Hopf structure of $H$, the group $\mathcal{R}(H)$ is in fact a ring that is equipped with an augmentation and an anti-endomorphism; these maps are closely related to the counit and the antipode, respectively, of the finite dual $H^\circ$. In the special case where $H = U_\mathfrak{g}$ is the enveloping algebra of a Lie algebra $\mathfrak{g}$, the ring $\mathcal{R}(H)$ was denoted by $\mathcal{R}(\mathfrak{g})$ earlier in this book; see in particular Section 8.5, where the case of a semisimple Lie algebra $\mathfrak{g}$ was studied in detail.

**10.3.1. Ring Structure**

As for arbitrary algebras, the group $\mathcal{R}(H)$ is free abelian: a $\mathbb{Z}$-basis is given by the elements $[S] \in \mathcal{R}(H)$ coming from finite-dimensional irreducible representations $S \in \text{Irr}_{\text{fin}} H$ (Proposition 1.46). The bifunctor $\cdot \otimes \cdot : \text{Rep} H \times \text{Rep} H \to \text{Rep} H$ now also endows $\mathcal{R}(H)$ with a multiplication:

$$[V][W] \overset{\text{def}}{=} [V \otimes W] \quad (V, W \in \text{Rep}_{\text{fin}} H).$$

This multiplication is well-defined and it makes $\mathcal{R}(H)$ into a ring with identity element $1_{\mathcal{R}(H)} = [\mathbb{1}]$. A detailed justification for the case of enveloping algebras of Lie algebras was given in §5.5.8. In view of (10.3) and (10.4), the reasoning for $U_\mathfrak{g}$ applies verbatim to arbitrary Hopf algebras; so we shall refrain from repeating it here. Adopting our earlier terminology for Lie algebras, we will refer to $\mathcal{R}(H)$ as the **representation ring** of $H$. The following $\mathbb{K}$-algebra will be called the **representation algebra** of $H$:

$$\mathcal{R}_\mathbb{K}(H) = \mathcal{R}(H) \otimes_\mathbb{Z} \mathbb{K}.$$  

For functoriality, observe that if $\phi : H \to K$ is a map in HopfAlg$k$, then the group homomorphism $\mathcal{R}(\phi) : \mathcal{R}(K) \to \mathcal{R}(H)$, $[W] \mapsto [\phi^* W]$ (§1.5.5) is in fact a ring homomorphism. Thus, the contravariant functor $\mathcal{R}(\cdot) : \text{Alg}_k \to \text{AbGroups}$

---

$^3$For finite group algebras, the converse holds: if $PV \equiv P \mathbb{1} \otimes V$ for all $V \in \text{Irr}_k G$, then $k G$ has the Chevalley property [33]. The current proof of this fact uses the classification of finite simple groups.
restricts to a functor $\mathcal{R}(\cdot) : \text{HopfAlg}_k \to \text{Rings}$ and we also obtain a functor $\mathcal{R}_k(\cdot) : \text{HopfAlg}_k \to \text{Alg}_k$, both contravariant.

If $H$ is cocommutative, then the representation ring $\mathcal{R}(H)$ is commutative, because $V \otimes W \cong W \otimes V$ holds in this case. This is also true for almost cocommutative Hopf algebras (Exercise 10.1.2). In general, however, $\mathcal{R}(H)$ may well be non-commutative. Indeed, the $1$-dimensional representations $k_\alpha$ with $\alpha \in G(H^\circ)$ provide us with a $\mathbb{Z}$-independent family of elements $[k_\alpha] \in \mathcal{R}(H)$ whose $\mathbb{Z}$-span is isomorphic to the group ring of $G(H^\circ)$ over $\mathbb{Z}$; see Exercise 10.1.6:

$$\mathbb{Z}[G(H^\circ)] \hookrightarrow \mathcal{R}(H)$$

$$\alpha \longmapsto [k_\alpha]$$

**Example 10.20** (Duals of finite group algebras). Let $G$ be a finite group and let $H = (kG)^*$ be the dual Hopf algebra of the group algebra $kG$. As a $k$-algebra, $H$ is isomorphic to the direct product of $|G|$ copies of $k$ (Example 9.17); so all irreducible representation of $H$ are $1$-dimensional. Hence, the above embedding is an isomorphism. Moreover, $H^\circ = H^* \cong kG$ and so $G(H^\circ) \cong G$ (Example 9.2). Therefore,

$$\mathbb{Z}G \cong \mathcal{R}((kG)^*) \quad \text{and} \quad kG \cong \mathcal{R}_k((kG)^*)$$

### 10.3.2. The Character Map

Recall that the character $\chi_V$ of a representation $V \in \text{Rep}_{\text{fin}} H$ is the linear form on $H$ that is defined by

$$\langle \chi_V, h \rangle = \text{trace}(h_V) \quad (h \in H).$$

As we have observed in §§1.5.2 and 1.5.5 in the context of arbitrary algebras, all characters belong to the $k$-vector space $C(H)$ consisting of all finite trace forms $t \in H^\circ_{\text{trace}} = H^\circ \cap H^\circ_{\text{trace}}$ such that $\text{Ker} t$ contains some cofinite semiprime ideal of $H$ and there is a well-defined group homomorphism,

$$\chi : \mathcal{R}(H) \rightarrow C(H) \hookrightarrow H^\circ_{\text{trace}} \hookrightarrow H^\circ$$

$$[V] \longmapsto \chi_V$$

The $k$-linear extension of $\chi$ is an embedding $\chi_k : \mathcal{R}_k(H) \hookrightarrow C(H)$ in $\text{Vect}_k$, which is an isomorphism if $k$ is a splitting field for $H$ (Proposition 1.49). For Hopf algebras, more can be said:

**Proposition 10.21.**

(a) $H^\circ_{\text{trace}}$ is a $k$-subalgebra of $H^\circ$.

(b) The character map $\chi$ is multiplicative, giving a ring homomorphism $\chi : \mathcal{R}(H) \rightarrow H^\circ_{\text{trace}}$ and a monomorphism $\chi_k : \mathcal{R}_k(H) \hookrightarrow H^\circ_{\text{trace}}$ in $\text{Alg}_k$. 
(c) If \( \mathbb{k} \) is a splitting field for \( H \), then \( C(H) \) is a \( \mathbb{k} \)-subalgebra of \( H^\ast \) and 
\[ \chi_k: \mathcal{R}_k(H) \rightarrow C(H) \]
is an isomorphism in \( \text{Alg}_k \).

**Proof.** (a) The space of Lie commutators \([H, H]\) is easily seen to be a subcoalgebra of \( H \). It follows that \( H^\ast_{\text{trace}} \cong (H/[H, H])^\ast \) is a subalgebra of the convolution algebra \( H^* \), and hence \( H^\ast_{\text{trace}} \) is a subalgebra of \( H^\ast \).

(b) For \( V, W \in \text{Rep}_\text{fin} H \) and \( h \in H \), the calculation 
\[ \langle \chi_V \chi_W, h \rangle = \langle \chi_V, h_{(1)} \rangle \langle \chi_W, h_{(2)} \rangle \]  
has \( \chi \) is a ring homomorphism and \( \chi_k \) is a \( \mathbb{k} \)-algebra map, which is mono by Proposition 1.49.

(c) is now also an immediate consequence of Proposition 1.49. \( \square \)

**Example 10.22** (Duals of finite group algebras, revisited). Let \( H = (\mathbb{k}G)^\ast \) for a finite group \( G \). Then \( C(H) = H^\ast \cong \mathbb{k}G \), because \( H \cong \mathbb{k}^{|G|} \) as \( \mathbb{k} \)-algebra (Example 9.17). The isomorphism \( \chi_k: \mathcal{R}_k(H) \rightarrow C(H) \) in Proposition 10.21(c) is inverse to the isomorphism \( \mathbb{k}G \rightarrow \mathcal{R}_k(H) \) in Example 10.20 due to the obvious equality \( \chi_{k\alpha} = \alpha \) for any \( H \in \text{HopfAlg}_k \) and \( \alpha \in G(H^\ast) \).

### 10.3.3. The Representation Ring of a Finite Group Algebra

Much of the material developed thus far in this section was already discussed in Section 3.1 for the group algebra \( \mathbb{k}G \) of a finite group \( G \) over the field \( \mathbb{k} \), albeit in slightly more down-to-earth terms. This subsection makes the connection.

Put \( p = \text{char} \mathbb{k} \), with \( p = 0 \) being allowed, and let \( G_{p'} \) denote the set of all \( p \)-regular elements of \( G \), that is, the elements of \( G \) whose order is not divisible by \( p \). So \( G_{p'} = G \) if \( p \nmid |G| \), in particular for \( p = 0 \). Since \( G_{p'} \) closed under \( G \)-conjugation, we may consider the algebra

\[ \text{cf}_k(G_{p'}) \overset{\text{def}}{=} \{ \text{class functions } G_{p'} \rightarrow \mathbb{k} \} \]

consisting of all functions \( G_{p'} \rightarrow \mathbb{k} \) that are constant on \( G \)-conjugacy classes or, equivalently, the algebra of all \( \mathbb{k} \)-valued functions on the set of \( p \)-regular conjugacy classes of \( G \). In the following proposition, we view characters of \( \mathbb{k}G \) as class functions on \( G \) as in §3.1.5.

**Proposition 10.23.** Let \( G \) be a finite group and \( \mathbb{k} \) a field of characteristic \( p \geq 0 \). Then there is a monomorphism of \( \mathbb{k} \)-algebras \( \chi_k: \mathcal{R}_k(\mathbb{k}G) \hookrightarrow \text{cf}_k(G_{p'}) \); it is an isomorphism if \( \mathbb{k} \) is a splitting field for \( G \).

**Proof.** If \( p \nmid |G| \), then \( \mathbb{k}G \) is semisimple by Maschke’s Theorem (§3.4.1). Hence \( C(\mathbb{k}G) = (\mathbb{k}G)^{\text{trace}} \) is the algebra of all trace forms on \( \mathbb{k}G \), which is isomorphic to the algebra \( \text{cf}_k(G) \) of all \( \mathbb{k} \)-valued functions on the set of conjugacy classes of \( G \) (3.11). Thus, \( \chi_k \) gives a monomorphism \( \mathcal{R}_k(\mathbb{k}G) \hookrightarrow \text{cf}_k(G) \) in \( \text{Alg}_k \), which is an
isomorphism if \( k \) is a splitting field for \( G \) (Proposition 10.21). Thus, the proposition holds in this case.

Now suppose that \( p \) is a divisor of \(|G|\). If \( k \) is a splitting field for \( G \), then we know that \( C(kG) = \{ f \in \text{cf}_k(G) \mid f(g) = f(g_{p'}) \text{ for all } g \in G \} \), where \( g_{p'} \) is the \( p \)-regular part of \( g \) (3.13). Thus, restriction of functions from \( G \) to \( G_{p'} \) gives an isomorphism from \( C(kG) \) to the algebra \( \text{cf}_k(G_{p'}) \). Proposition 10.21 now gives the desired isomorphism \( \chi_k : \mathcal{R}_k(kG) \to \text{cf}_k(G_{p'}) \). For general \( k \), fix an algebraic closure \( \overline{k} \) and consider the embedding \( \overline{k} \otimes \mathcal{R}_k(kG) \to \mathcal{R}_k(\overline{k}G) \) (Lemma 1.48) and the isomorphism \( \chi_{\overline{k}} : \mathcal{R}_k(\overline{k}G) \to \text{cf}_k(G_{p'}) \) provided by the foregoing. The composite is an embedding \( \mathcal{R}_k(kG) \to \text{cf}_k(G_{p'}) \), which has image in \( \text{cf}_k(G_{p'}) \subseteq \text{cf}_k(G_{p'}) \), and this image coincides with the image of \( \chi_k \); see (1.60). This completes the proof of the proposition. \( \square \)

**Functoriality.** Let \( \phi : H \to G \) be a group homomorphism. This homomorphism gives rise to a homomorphism \( k\phi : kH \to kG \) in \( \text{HopfAlg}_k \) and then a “restriction” homomorphism of (commutative) rings (§10.3.1),

\[
\mathcal{R}(k\phi) : \mathcal{R}(kG) \to \mathcal{R}(kH).
\]

There is also an even more obvious map of (commutative) \( k \)-algebras,

\[
\phi^* : \text{cf}_k(G_{p'}) \to \text{cf}_k(H_{p'}),
\]

which is given by \( \phi^*(f) = f \circ \phi \) for a class function \( f : G_{p'} \to k \). With this and Proposition 10.23, we obtain the following version of (1.59):

\[
\begin{array}{ccc}
\mathcal{R}(kG) & \xrightarrow{\chi} & \text{cf}_k(G_{p'}) \\
\mathcal{R}(kH) & \xrightarrow{\chi} & \text{cf}_k(H_{p'}) \\
\end{array}
\]

(10.19)

In case \( H \) is a subgroup of \( G \) and \( \phi \) is the inclusion map, we will also write \( \text{Res}^G_H \) for both \( \mathcal{R}(k\phi) \) and \( \phi^* \).

**A Splitting Principle.** The following proposition, in some ways, reduces the problem of describing \( \mathcal{R}(kG) \) to the case where \( G \) is cyclic, and hence (over a large enough field) to the case where all irreducible representations of \( G \) are 1-dimensional.\(^4\) A finite group is called \( p \)-regular if its order is not divisible by \( p \) or, equivalently, all its elements are \( p \)-regular (which is always true for \( p = 0 \)).

**Proposition 10.24.** Let \( G \) be a finite group and \( k \) a field of characteristic \( p \geq 0 \). Let \( \mathcal{C}_{p'} \) denote the family of all \( p \)-regular cyclic subgroups of \( G \). Then the ring homomorphism \( \mathcal{R}(kG) \to \prod_{C \in \mathcal{C}_{p'}} \mathcal{R}(kC) \), \( x \mapsto (\text{Res}^G_C x) \) is injective.

\(^4\)An embedding such as the one provided by the proposition is often referred to as a “splitting principle” (e.g., Fulton and Lang [79] and Weibel [203]). See Exercise 10.3.3 for an application.
Proof. Let “res” denote the map in the proposition and also the corresponding map for \( R_k(G) \) and the map on function algebras, \( \text{cf}_k(C_p) \to \prod C \text{cf}_k(C_p) \), where \( C \) ranges over \( C_p \). Note that the latter map is injective, because \( G_p \) is the union of all \( C \in C_p \). Also, \( \text{cf}_k(C_p) = \mathbb{K}^C \), the algebra of all functions \( C \to \mathbb{K} \). Therefore, (10.19) and Proposition 10.23 give the following commutative diagram:

\[
\begin{array}{cccc}
\mathcal{R}(kG) & \overset{\text{can}}{\longrightarrow} & \mathcal{R}_k(kG) & \overset{\chi_k}{\longrightarrow} \\
\downarrow{\text{res}} & & \downarrow{\text{res}} & \downarrow{\text{res}} \\
\prod_C \mathcal{R}(kC) & \overset{\Pi C \text{ can}}{\longrightarrow} & \prod_C \mathcal{R}_k(kC) & \overset{\Pi C \chi_k}{\longrightarrow} \\
\end{array}
\]

Since \( \mathcal{R}(kG) \) is free abelian, the canonical map \( \mathcal{R}(kG) \to \mathcal{R}_k(kG) \) has kernel \( p \mathcal{R}(kG) \). Therefore, the kernel of the map in the proposition, the leftmost “res” in the diagram above, must be contained in \( p \mathcal{R}(kG) \). On the other hand, since the image of this map is free abelian, the kernel is a direct summand of \( \mathcal{R}(kG) \). This forces the kernel to vanish, proving the proposition. \( \square \)

10.3.4. Additional Structure

For an arbitrary \( H \in \text{HopfAlg}_\mathbb{K} \), the character map \( \chi \) provides a strong link between the representation ring \( \mathcal{R}(H) \) and the Hopf algebra \( H^\circ \). Certain structural features of \( \mathcal{R}(H) \) are naturally connected to various Hopf data of \( H^\circ \) via \( \chi \).

Dimension Augmentation and Counit. Since \( \dim \mathbb{K} (V \otimes W) = (\dim \mathbb{K} V)(\dim \mathbb{K} W) \), the dimension augmentation (§1.5.5) now is a ring homomorphism,

\[ \dim : \mathcal{R}(H) \to \mathbb{Z}, \quad [V] \mapsto \dim \mathbb{K} V. \]

This homomorphism corresponds to the counit \( \varepsilon_{H^\circ} = u^*|_{H^\circ} \) under the character map: the obvious formula \( \chi_V(1) = \dim \mathbb{K} V \cdot 1_\mathbb{K} \) translates into the commutative diagram

\[
\begin{array}{ccc}
\mathcal{R}(H) & \overset{\chi}{\longrightarrow} & H^\circ \\
\downarrow{\dim} & & \downarrow{\varepsilon_{H^\circ}} \\
\mathbb{Z} & \overset{\text{can.}}{\longrightarrow} & \mathbb{K}
\end{array}
\]

Duals and the Antipode. In view of exactness of the duality functor, we obtain a well-defined group endomorphism \( \cdot^* : \mathcal{R}(H) \to \mathcal{R}(H) \) by putting

\[ [V]^* := [V^*]. \]

In fact, \( \cdot^* \) is a ring anti-endomorphism of \( \mathcal{R}(H) \) by (10.12):

\[ ([V][W])^* = [W]^*[V]^*. \]

If the antipode \( S \) of \( H \) is bijective, then \( \cdot^* \) permutes the standard \( \mathbb{Z} \)-basis of \( \mathcal{R}(H) \), consisting of the classes \([S]\) with \( S \in \text{Irr}_{\text{fin}} H \); and if \( S^2 \) is an inner automorphism
of \( H \), then \( \cdot^* \) is an involution of \( \mathcal{R}(H) \), that is, a ring anti-automorphism of order \( \leq 2 \) (Lemma 10.2).

Taking traces in (10.10) and using (B.25), we obtain \( \chi_{V^*} = S^\circ(\chi_V) \). Thus, \( \cdot^* \) corresponds to the antipode \( S^\circ = S^\circ|_{H^\circ} \) of \( H^\circ \) under the character map, that is, the following diagram commutes:

\[
\begin{array}{ccc}
\mathcal{R}(H) & \xrightarrow{\chi} & H^\circ \\
\cdot^* & \downarrow & \downarrow S^* \\
\mathcal{R}(H) & \xrightarrow{\chi} & H^\circ 
\end{array}
\] (10.21)

The Regular Representation. Now assume that \( H \) is finite dimensional. Then we may consider the class of the regular representation, \([H_{\text{reg}}] \in \mathcal{R}(H)\). Since \( H_{\text{reg}} \) is self-dual (Theorem 10.9), we have

\[ [H_{\text{reg}}]^* = [H_{\text{reg}}]. \]

Moreover, Corollary 10.5 gives \( H_{\text{reg}} \otimes V \cong H_{\text{reg}}^{\otimes \dim_k V} \) for any \( V \in \text{Rep}_{\text{fin}} H \). Letting \( W \) denote the dual of the \( S^2 \)-twist of \( V \), it follows from Lemma 10.2 that \( V \cong W^* \).

Thus, we also have

\[ V \otimes H_{\text{reg}} \cong (H_{\text{reg}} \otimes W)^* \cong (H_{\text{reg}}^{\otimes \dim_k W})^* \cong H_{\text{reg}}^{\otimes \dim_k V}. \]

Therefore, for any \( x \in \mathcal{R}(H) \),

\[ [H_{\text{reg}}]x = x[H_{\text{reg}}] = (\dim x)[H_{\text{reg}}]. \] (10.22)

In the language of the upcoming Section 12.1, (10.22) states that \( [H_{\text{reg}}] \) is a left and right “integral” for \( \mathcal{R}(H) \) with respect to the dimension augmentation. We will see later (Corollary 12.17) that if \( H \) is semisimple and involutory, then the regular character \( \chi_{\text{reg}} = \chi_{H_{\text{reg}}} \) satisfies \( \chi_{\text{reg}}f = f \chi_{\text{reg}} = \langle \varepsilon H^*, f \rangle \chi_{\text{reg}} \) for all \( f \in H^* = H^\circ \), which corresponds to (10.22).

**Exercises for Section 10.3**

10.3.1 (\( \lambda \)-ring structure of \( \mathcal{R}(H) \) for cocommutative \( H \)). A commutative ring \( R \) is called a \( \lambda \)-ring if there is a family of operations \( \lambda^n : R \rightarrow R \) \((n \geq 0)\) such that \( \lambda^0 x = 1, \lambda^1 x = x \) and \( \lambda^n(x + y) = \sum_{i=0}^n \lambda^i(x)\lambda^{n-i}(y) \) for all \( x, y \in R \) and all \( n \).

Now let \( H \in \text{HopfAlg}_k \) be cocommutative and put \( R = \mathcal{R}(H) \). The purpose of this exercise is to equip \( R \) with the structure of a \( \lambda \)-ring. Exercise 10.2.1 will be crucial.

---

\(^5\)In some sources, \( \lambda \)-rings as defined above are called pre-\( \lambda \)-rings. Originally introduced by Grothendieck [93], the notion is fundamental in Riemann-Roch algebra and in \( K \)-theory (e.g., Fulton and Lang [79] and Weibel [203]). For applications of \( \lambda \)-rings in group representation theory, see Atiyah and Tall [7] and Kervaire [120].
(a) Let $0 \to U \to V \to W \to 0$ be a short exact sequence in $\mathsf{Rep}_{\text{fin}} H$. Show that $[\Lambda^n V] = \sum_{i=0}^n [\Lambda^i U \otimes \Lambda^{n-i} W]$ holds in $R$.

(b) For $V \in \mathsf{Rep}_{\text{fin}} H$, define a power series over $R$ by

$$
\lambda_t(V) \overset{\text{def}}{=} \sum_{n \geq 0} [\Lambda^n V] t^n \in \Lambda_R := 1 + tR[[t]].
$$

Observe that $\Lambda_R$ is a subgroup of the group of units $R[[t]]^\times$. Show that, if $0 \to U \to V \to W \to 0$ is exact in $\mathsf{Rep}_{\text{fin}} H$, then $\lambda_t(V) = \lambda_t(U)\lambda_t(W)$. Conclude that $\lambda_t : (R, +) \to (\Lambda_R, \cdot)$, $[V] \mapsto \lambda_t(V)$, yields a well-defined group homomorphism and that $\lambda^n[V] := [\Lambda^n V]$ yields operations $\lambda^n$ on $R$ that make $R$ into a $\lambda$-ring. Furthermore, $\lambda^n(x^r) = \lambda^n(x)^r$ for all $x \in R$.

(c) Observe that $R_+ := \{ [V] \mid V \in \mathsf{Rep}_{\text{fin}} H \}$ is stable under the operations $\lambda^n$ and that $\lambda^nx = 0$ if $x \in R_+$ and $n > \dim x$. Show that $\lambda^n x$ is a unit in $R$ for $x \in R_+$ with $n = \dim x$.

**10.3.2 (Adams operations on $\mathcal{H}(H)$ for cocommutative $H$).** This exercise is a continuation of Exercise 10.3.1. We use the same hypotheses and notation. For $x \in R = \mathcal{H}(H)$, define the power series

$$
\psi_t(x) \overset{\text{def}}{=} (\dim x) 1 - t \frac{d}{dt} \log \lambda_{-t}(x) = \sum_{n \geq 0} \psi^n(x) t^n \in R[[t]].
$$

Prove:

(a) $\psi^0(x) = \dim x$, $\psi^1(x) = x$, $\psi^2(x) = x^2 - 2\lambda^2(x)$ and, for $n \geq 3$,

$$
\psi^n(x) = \lambda^1(x)\psi^{n-1}(x) - \lambda^2(x)\psi^{n-2}(x) + \ldots + (-1)^n \lambda^{n-1}(x)\psi^1(x) + (-1)^n n\lambda^n(x).
$$

In particular, if $x \in R_+ = \{ [V] \mid V \in \mathsf{Rep}_{\text{fin}} H \}$ and $\dim x = 1$, then $\psi^n(x) = x^n$ for all $n \geq 0$.

(b) $\psi^n(x + y) = \psi^n(x) + \psi^n(y)$ for all $x, y \in R$.

(c) If all finite-dimensional irreducible representations of $H$ are 1-dimensional, then all $\psi^n$ are ring endomorphism, $\psi^n \circ \psi^m = \psi^{nm}$ for all $n$ and $m$, and the character $\chi_{\psi^n(x)} \in H^\hat{}$ for $x \in R$ is given by $\langle \chi_{\psi^n(x)}, h \rangle = \langle \chi_x, h^{[n]} \rangle$, where $h^{[n]} := h^{(1)} h^{(2)} \cdots h^{(n)}$ is the so-called $n^{\text{th}}$ **Sweedler power** of $h \in H$.

**10.3.3 (Adams operations for finite group algebras).** Let $G$ be a finite group. Use Proposition 10.24, Lemma 1.48 and Exercise 10.3.2 to show that the Adams operations $\psi^n$ are ring endomorphism of $\mathcal{H}(kG)$ satisfying $\psi^n \circ \psi^m = \psi^{nm}$ for all $n$ and $m$ and that $\langle \chi_{\psi^n(x)}, g \rangle = \langle \chi_x, g^n \rangle$ for $x \in \mathcal{H}(kG)$, $g \in G$.

---

*The maps $\psi^n : R \to R$ are called the **Adams operations**, having been introduced by J. F. Adams [3].
10.4. Actions and Coactions of Hopf Algebras on Algebras

The investigation of (co)actions of a Hopf algebra $H$ on other algebras is often referred to as noncommutative invariant theory. The central underlying notion is that of an $H$-(co)module algebra. For enveloping algebras $H = U\mathfrak{g}$ and group algebras $H = kG$, we have encountered $H$-module algebras earlier (§5.5.5) under the respective names $\mathfrak{g}$-algebras and $G$-algebras. In Chapter 11, we will use the general framework about to be developed in the present section to discuss actions of affine algebraic groups on algebras.

10.4.1. Module and Comodule Algebras

By definition, a left $H$-module algebra is an algebra $A \in \text{Alg}_k$ that is also a left $H$-module in such a way that the multiplication $m: A \otimes A \to A$ and the unit $u: k \to A$ are not merely $k$-linear but are in fact morphisms in the category $H\text{Mod} \equiv \text{Rep} H$. Here, $A \otimes A$ is equipped with the $H$-module structure (10.2) and $k$ is viewed as the trivial representation, $1$. Explicitly, denoting the $H$-action $H \otimes A \to A$ by $h \otimes a \mapsto h.a$ as usual, the following identities hold for all $h \in H$ and $a, b \in A$:

\begin{equation}
(10.23) \quad h.(ab) = (h(1).a)(h(2).b) \quad \text{and} \quad h.1_A = \langle \varepsilon, h \rangle 1_A.
\end{equation}

This definition can be phrased more compactly by stating that a left $H$-module algebra is the same as an algebra in the tensor category $H\text{Alg}$. Naturally, a homomorphism (or map) between $H$-module algebras, by definition, is an algebra map that is also an $H$-module map. We have thus assembled the ingredients of a category,

$H\text{Alg}$.

The category $U\mathfrak{g}\text{Alg}$ for $\mathfrak{g} \in \text{Lie}_k$ was denoted by $\mathfrak{g}\text{Alg}$ in §5.5.5. The above general definition does of course work for any bialgebra, and there is an obvious right-handed analog. The following lemma is clear by inspection of (10.23).

**Lemma 10.25.** Let $A \in H\text{Alg}$. Then the action homomorphism $H \to \text{End}_k(A)$ in $\text{Alg}_k$ restricts to maps $GH \to \text{Aut}_{\mathfrak{g}\text{Alg}}(A)$ in Groups and $LH \to \text{Der} A$ in $\text{Lie}_k$.

The dual notion is that of an $H$-comodule algebra. Stated in its right-sided version, this is an algebra $A$ in the category $\text{Mod}_H^R$ of all right $H$-comodules; so $A \in \text{Mod}_H^R$ and $A$ is equipped with a unit map $u: k = 1' \to A$ and a multiplication map $m: A \otimes A \to A$ in $\text{Mod}_H^R$ satisfying (1.1). Here, $1'$ has the trivial $H$-coaction and $A \otimes A \in \text{Mod}_H^R$ as in (10.5). This definition unfolds as follows: $A \in \text{Alg}_k$ via $m$ and $u$ and $A \in \text{Mod}_H^R$ via the $H$-coaction $\delta = \delta_A: A \to A \otimes H, a \mapsto a(0) \otimes a(1)$, which needs to satisfy

\begin{equation}
(10.24) \quad \delta(ab) = a(0)b(0) \otimes a(1)b(1) \quad \text{and} \quad \delta(1_A) = 1_A \otimes 1_H.
\end{equation}

With the obvious notion of morphisms of $H$-comodule algebras, we have a category $H\text{Alg}^R$. 

The Passage between Module and Comodule Algebras. The following proposition is the algebra analog of Proposition 9.12.

Proposition 10.26. (a) Every right $H$-comodule algebra is a left $H^\circ$-module algebra via Hom-$\otimes$ adjunction, and every morphism of right $H$-comodule algebras becomes a left $H^\circ$-module algebra morphism in this way.

(b) A left $H$-module algebra $A$ is a right $H^\circ$-comodule algebra via Hom-$\otimes$ adjunction if and only if $A$ is locally finite as $H$-module. Morphisms between locally finite left $H$-module algebras become right $A^\circ$-comodule algebra morphisms in this way.

Proof. For (a), let $A \in \text{Alg}^H$ be given, with $H$-coaction $A \to A \otimes H, a \mapsto a(0) \otimes a(1)$. Then the following action rule makes $A$ a module over the algebra $H^*$, and hence by restriction a module over the Hopf algebra $H^\circ$; see (9.24):

\[ h^* \cdot a := a(0) \langle h^*, a(1) \rangle \quad (h^* \in H^\circ, a \in A). \]

From (10.24), we obtain $h^* \cdot 1_A = 1_A \langle h^*, 1_H \rangle = \langle \varepsilon_H, h^* \rangle 1_A$ and

\[ h^* \cdot (ab) = a(0) b(0) \langle h^*, a(1) b(1) \rangle = a(0) b(0) \langle h^*_1, a(1) \rangle \langle h^*_2, b(1) \rangle \]

\[ = a(0) \langle h^*_1, a(1) \rangle b(0) \langle h^*_2, b(1) \rangle = (h^*_1 \cdot a)(h^*_2 \cdot b), \]

proving (10.23). Thus, the above action makes $A$ a left $H^\circ$-module algebra. We also know from Proposition 9.12(a) that, with this action, $H$-comodule maps become $H^*$-module maps, and hence $H^\circ$-module maps. Therefore, every morphism of right $H$-comodule algebras becomes a left $H^\circ$-module algebra morphism. This proves (a). The proof of (b) is analogous, using (9.25) and part (b) of Proposition 9.12 in place of (9.24) and part (a). □

Of course, local finiteness is automatic for $H$-modules if $H$ is finite dimensional, and we also have $H^\circ = H^*$ and $H^{**} \cong H$ in this case. Thus:

Corollary 10.27. If $H$ is finite dimensional, then $H \text{Alg} \cong \text{Alg}^H$.

10.4.2. Examples

As was mentioned at the start of this section, the standard examples of $H$-module algebras arise from Lie algebras and from groups. In this subsection, we present some examples of $H$-module algebras and $H$-comodule algebras for other Hopf algebras and bialgebras.

Adjoint Actions. Let $A \in \text{Alg}_k$ and let $f : H \to A$ be a map in $\text{Alg}_k$. Then $A \in _H \text{Mod}^H_H$ by restriction of the regular $(A, A)$-bimodule on $A$ (Example 1.2). The adjoint action (10.6) now becomes the following $H$-action on $A$:

\[ (10.25) \quad h \cdot a = f(h(1))a f(S(h(2))) \quad (h \in H, a \in A). \]
This action satisfies (10.23):

\[ h.1_A = f(h(1))1_A f(S(h(2))) = f(h(1)S(h(2)))1_A = \langle \varepsilon, h \rangle 1_A. \]

and

\[ h.(ab) = f(h(1))ab f(S(h(2))) = f(h(1))a\langle \varepsilon, h(2) \rangle b f(S(h(3))) \]
\[ = f(h(1))a f(S(h(2))) f(h(3))b f(S(h(4))) = (h(1),a)(h(2),b). \]

Thus, \( A \) becomes an \( H \)-module algebra via the adjoint action; we will adopt the notation \( A_{\text{ad}} \) from §10.1.1.

**Example 10.28** (\( H_{\text{ad}} \) as \( H \)-module algebra). With \( \text{Id}_H : H \to H \), we obtain \( H_{\text{ad}} \in H_{\text{Alg}} \). The given map \( f \) in \( \text{Alg}_k \) satisfies \( f(h,k) = h.f(k) \); so \( f : H_{\text{ad}} \to A_{\text{ad}} \) is in fact a morphism in \( H_{\text{Alg}} \).

**Example 10.29** (\( H \)-module algebras from representations). If \( V \in \text{Rep} \, H \), then \( \text{End}_V(V)_{\text{ad}} \in H_{\text{Alg}} \) via the map \( H \to \text{End}_V(V), h \mapsto h_V \), which is a map in \( H_{\text{Alg}} \) by the previous example. The \( H \)-action (10.25) on \( \text{End}_V(V)_{\text{ad}} \) is the same as (10.8).

**Convolution Algebras.** For any \( A \in \text{Alg}_k \), the convolution algebra \( \text{Hom}_k(H, A) \) (§9.1.4) becomes a left and right \( H \)-module algebra using the \( H \)-actions \( \cdot \) and \( \langle \cdot, \cdot \rangle \) from (2.19): \( \langle h \cdot f \cdot h', k \rangle = \langle f, h'kh \rangle \) for \( h, h', k \in H \) and \( f \in \text{Hom}_k(H, A) \). To check the requisite identity \( h \cdot (f \cdot g) = (h(1) \cdot f) \star (h(2) \cdot g) \) in (10.23) for the left action \( \cdot \), we calculate

\[ \langle h \cdot (f \cdot g), k \rangle = \langle f \cdot g, kh \rangle = \langle f, k(1)h(1) \rangle \langle g, k(2)h(2) \rangle \]
\[ = \langle h(1), f, k(1) \rangle \langle h(2), g, k(2) \rangle = \langle (h(1) \cdot f), (h(2) \cdot g), k \rangle. \]

The identity \( h \cdot (u \circ \varepsilon) = \langle \varepsilon, h \rangle (u \circ \varepsilon) \) is equally routine to check as are the corresponding identities for \( \langle \cdot, \cdot \rangle \). In particular, the dual algebra \( H^* \) becomes a right and left \( H \)-module algebra in this way; it is isomorphic to an \( H \)-module subalgebra of \( \text{Hom}_k(H, A) \) if \( A \neq 0 \). See (10.16) for one of the identities (10.23) for \( \langle \cdot, \cdot \rangle \).

**Tensor, Symmetric and Exterior Algebras.** As we have seen (§10.2.3), all tensor powers \( V \otimes_k \) of a given representation \( V \in \text{Rep} \, H \) become representations of \( H \) as well. Hence, \( TV = \bigoplus_{k \geq 0} V \otimes^k \in \text{Rep} \, H \). Since \( H \) acts on \( V \otimes^k \) \((k \geq 2)\) through the map \( \Delta_k : H \to H \otimes^k \) and \( V \otimes^0 = 1 \), conditions (10.23) are satisfied. Thus, the tensor algebra \( TV \) is in fact an \( H \)-module algebra. In this way, we obtain a functor,

\[ T : \text{Rep} \, H \to H_{\text{Alg}}. \]

This functor is left adjoint to the forgetful functor \( \text{Rep}_H : H_{\text{Alg}} \to \text{Rep} \, H \) that forgets the multiplication and unit of a given \( H \)-module algebra. All this was discussed in §§5.5.4 and 5.5.5 in the context of enveloping algebras and the material extends in a straightforward manner to general Hopf algebras. Moreover, generalizing earlier observations for enveloping algebras and for group algebras, one shows that if \( H \) is cocommutative, then the \( H \)-action on \( TV \) passes down to actions on the symmetric algebra \( \text{Sym} \, V \) and on the exterior algebra \( \Lambda \, V \). For example, for the
symmetric algebra, recall that \(\text{Sym} V = (TV)/I\) with \(I\) the ideal that is generated by the Lie commutators \([v,v'] = v \otimes v' - v' \otimes v\) \((v,v' \in V)\). In view of (10.23), it suffices to show that \(h \cdot [v,v'] \in I\) for all \(h \in H\). But \(h_{(1)} \otimes h_{(2)} = h_{(2)} \otimes h_{(1)}\) by hypothesis on \(H\), and so

\[
\begin{align*}
  h \cdot [v,v'] &= h_{(1)} \cdot (v \otimes h_{(2)} \cdot v' - h_{(2)} \cdot v' \otimes h_{(1)} \cdot v) \\
  &= h_{(1)} \otimes h_{(2)} \cdot v' - h_{(2)} \otimes h_{(1)} \cdot v = [h_{(1)} \cdot v, h_{(2)} \cdot v'],
\end{align*}
\]

proving the claim for \(\text{Sym} V\). The verification for \(\Lambda V\) is similar. See also Exercise 10.2.1 for more on \(\text{Sym} V\) and \(\Lambda V\) as \(H\)-representations.

**Regular Comodule Algebras.** As we had seen in Example 9.10, the comultiplication \(\Delta\) makes any Hopf algebra \(H\) a right (and left) comodule over itself. Since condition (10.24) also holds, because \(\Delta\) is a map in \(\text{Alg}_{k^M}\), we obtain the regular comodule algebra, \(H \in \text{Alg}^{k^H}\). By Proposition 10.26, we may view \(H\) as a left \(H^\circ\)-module algebra with action

\[
h^\circ \cdot h = h_{(1)} \langle h^\circ, h_{(2)} \rangle \quad (h^\circ \in H^\circ, h \in H).
\]

**Comodule Algebras over Monoid Algebras.** Let \(M\) be a monoid and consider the monoid algebra \(kM\) with it standard bialgebra structure. The category \(\text{Mod}^{kM}\) of \((\text{right or left})\) \(kM\)-comodules is equivalent to the category \(\text{Vect}_k^M\) of \(M\)-graded \(k\)-vector spaces \(V\) (Example 9.11). Explicitly, an \(M\)-grading \(V = \bigoplus_{x \in M} V^x\) corresponds to the \(kM\)-coaction \(\delta: V \to V \otimes kM\) that is given by \(\delta v = \sum_{x \in M} v^x \otimes x\), where \(v^x \in V^x\) is the \(x\)-homogeneous component of \(v\). If \(V = A\) is also equipped with a \(k\)-algebra structure, then conditions (10.24) state that \(A^x A^y \subseteq A^{xy}\) for all \(x,y \in M\) and \(1_A \in A_1\). Thus, \(A\) is a \(kM\)-comodule algebra precisely if \(A\) is an \(M\)-graded \(k\)-algebra as in Exercise 1.1.11. In sum,

\[
\text{Alg}^{kM} = \text{Alg}^M_k.
\]

10.4.3. Smash Products and Invariants

Associated to any \(A \in H\text{Alg}\), there are two important \(k\)-algebras: the *algebra of invariants*, \(A^H\), and the *smash product*, \(A \# H\). The former is nothing but the space of \(H\)-invariants in \(A\) (§10.1.1), which is easily seen to be a \(k\)-subalgebra of \(A\) by virtue of (10.23):

\[
A^H \overset{\text{def}}{=} \{ a \in A \mid h \cdot a = \langle e, h \rangle a \text{ for all } h \in H \}.
\]

The smash product \(A \# H\), as a \(k\)-vector space, is defined to be \(A \otimes H\). The multiplication of \(A \# H\) is given by the rule

\[
(a \otimes h)(b \otimes k) = a(h_{(1)}, b) \otimes h_{(2)} k.
\]

A routine verification shows that this definition makes \(A \# H\) into a \(k\)-algebra and that \(A\) and \(H\) are \(k\)-subalgebras of \(A \# H\) via the identifications \(A \equiv A \otimes 1\) and
\( H \cong 1 \otimes H \). Using these identifications, the element \( a \otimes h \) can simply be written as \( ah \) and the essence of the multiplication rule in \( A \# H \) becomes

\[
(10.28) \quad ha = (h_{(1)}, a)h_{(2)} \quad (h \in H, a \in A).
\]

For more on smash products, see Exercises 5.5.2 (with \( H = U_\emptyset \)) and 12.1.5 (with \( H \in \text{HopfAlg}_k \) finite dimensional).

**Outlook: The Structure of Cocommutative Hopf Algebras.** For background, we mention the following theorem, which is generally attributed to Cartier, Gabriel and Kostant. Proofs may be found in [194, Section 13.1] and [153, Section 5.6].

For any \( H \in \text{HopfAlg}_k \), the Lie algebra \( g = L H \) of primitive elements is easily seen to be stable under the adjoint (conjugation) action of the group \( G = GH \) of group-like elements on \( H \). Since the action \( G \subseteq \text{H} \) is by algebra automorphisms, the action \( G \subseteq g \) is by Lie algebra automorphisms. Thus, \( G \) acts by \( k \)-algebra automorphisms on the enveloping algebra \( Ug \), which is therefore a \( G \)-algebra. Hence, we may form the smash product \( Ug \#_k G \). The embeddings \( g \hookrightarrow H_{\text{lie}} \) and \( G \hookrightarrow H^* \) extend uniquely to algebra maps \( Ug \rightarrow H \) and \( k G \rightarrow H \), which combine to give a homomorphism \( Ug \#_k G \rightarrow H \) in \( \text{Alg}_k \). For cocommutative Hopf algebras over an algebraically closed field of characteristic 0, this is an isomorphism:

**Cartier-Gabriel-Kostant Theorem.** Let \( H \in \text{HopfAlg}_k \) be cocommutative and let \( k \) be algebraically closed with \( \text{char} \, k = 0 \). Then \( H \cong U_\emptyset \#_k G \), where \( g = LH \) and \( G = GH \).

### 10.4.4. Adjoint Action and Chevalley Property

In this subsection, we will take a closer look at the adjoint action (10.25) of finite-dimensional Hopf algebra \( H \) on its semisimplification, \( H^{sp} = H/\text{rad} \, H \).

**Proposition 10.30.** Let \( H \in \text{HopfAlg}_k \) be finite dimensional.

(a) There is a monomorphism \( (H^{sp})_{\text{ad}} \hookrightarrow \bigoplus_{S \in \text{Irr} \, H} S \otimes S^* \) in \( \text{Rep} \, H \); it is an isomorphism if \( k \) is a splitting field for \( H \).

(b) If \( H \) is involutory and \( (H^{sp})_{\text{ad}} \) is completely reducible, then \( \text{char} \, k \) does not divide the degree of any absolutely irreducible representation of \( H \).

**Proof.** (a) The embedding follows from the isomorphism \( S \otimes S^* \cong \text{End}_k(S) \) in \( \text{Rep} \, H \) combined with the Artin-Wedderburn isomorphism (Corollary 1.34),

\[
\begin{array}{cccc}
H^{sp} & \sim & \prod_{S \in \text{Irr} \, H} \text{End}_{D(S)}(S) & \hookrightarrow \prod_{S \in \text{Irr} \, H} \text{End}_k(S) \\
\wedge & & \downarrow & \nwarrow \\
& & \text{End}_{D(S)}(S) & \hookrightarrow \text{End}_k(S) \\
x & \mapsto & (x_S) &
\end{array}
\]

Indeed, the action maps \( H^{sp} \rightarrow \text{End}_k(S), x \mapsto x_S \), are equivariant for the adjoint \( H \)-action on \( H^{sp} \) (Example 10.29). If \( k \) is a splitting field, then \( D(S) = k \) for all \( S \), and hence the embedding is an isomorphism.
(b) Let \( S \in \text{Irr } H \) be absolutely irreducible; so \( D(S) = \text{End}_H(S) = k \) (Proposition 1.36). In view of the above epimorphism \((H^{s.p.})_{\text{ad}} \rightarrow S \otimes S^*\), we know that \( S \otimes S^* \) is completely reducible. Thus, (b) follows from Proposition 10.15(a). \( \square \)

Recall that \( H \) is said to have the Chevalley property if \( S \otimes S' \) is completely reducible for any two \( S, S' \in \text{Irr } H \) (§10.2.5). In particular, \( S \otimes S^* \) is completely reducible in this case (Lemma 10.2). Thus, for any involutory finite-dimensional Hopf algebra \( H \) over a splitting field \( k \) of characteristic \( p \geq 0 \), we have the following implications:

<table>
<thead>
<tr>
<th>( H ) has the Chevalley property</th>
<th>((H^{s.p.})_{\text{ad}}) is completely reducible</th>
<th>all irreducible representations of ( H ) have ( p')-degree</th>
</tr>
</thead>
<tbody>
<tr>
<td>\implies | | | \implies</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The Case of Finite Group Algebras. It turns out that the three conditions above are all equivalent for any finite group algebra \( kG \) over a splitting field \( k \). To see this, we may of course assume that \( \text{char } k = p > 0 \). By a result of Michler [150, Theorem 2.4], it is known that if \( p \) does not divide the degree of any irreducible representation of \( kG \), then \( G \) has a normal Sylow \( p \)-subgroup, which we have seen to be equivalent to the Chevalley property of \( kG \) (Example 10.19). The proof of Michler’s Theorem relies on the classification of finite simple groups. It is a rather more elementary fact that \((kG)_{\text{ad}}\) is completely reducible if and only if \( G \) has a central Sylow \( p \)-subgroup (Exercise 3.4.10). To summarize, adding two obvious vertical implications,

\[
\begin{align*}
(kG)_{\text{ad}} \text{ is completely reducible} & \iff G \text{ has a central Sylow } p \text{-subgroup} \\
(kG^{s.p.})_{\text{ad}} \text{ is completely reducible} & \iff G \text{ has a normal Sylow } p \text{-subgroup}
\end{align*}
\]

10.4.5. Finite Generation of Invariants

Classical invariant theory, for the most part, studies the action of a group \( G \) on some affine commutative \( A \in kG\text{Alg} \). One of the main concerns is to try and prove that the invariant subalgebra \( A^G = \{ a \in A \mid g.a = a \text{ for all } g \in G \} \) is again affine and to possibly exhibit an explicit finite set of algebra generators. Hilbert’s fourteenth problem, posed in 1900, considers the natural action of an algebraic subgroup of \( G \leq \text{GL}_n(k) \) on the polynomial algebra \( k[x_1, \ldots, x_n] \) (“by linear substitution of the variables”) and asks whether the ring of invariants \( k[x_1, \ldots, x_n]^G \) will always be affine. While the answer is negative in general—the first counterexample was found by Nagata [155] in 1958—many positive results have been proven under additional hypotheses. Notably, Noether’s Finiteness Theorem [158] states that
if $G$ is finite, then $A^G$ is affine for any affine commutative $A \in \mathcal{G} \mathcal{G}\text{-Alg}$. In this subsection, we present a more general result, replacing $\mathcal{G} \mathcal{G}$ by any finite-dimensional cocommutative Hopf algebra (Theorem 10.32). The result is originally due to Grothendieck (see [54]), but our presentation follows Ferrer Santos [72], who rediscovered it independently.

**Characteristic Polynomials.** Let $A \in H\text{-Alg}$ and let $V, W \in A^H\text{-Mod}$, where $A^H$ is the smash product ($\S 10.4.3$). Viewing $V$ and $W$ as representations of $A$ and of $H$ by restriction, $\text{Hom}_{k}(V, W)$ may be regarded as a representation of $H$ (10.8) and we may consider $\text{Hom} \# (V, W) \subseteq \text{Hom}_{k}(V, W)$. If $H$ is cocommutative, then the following computations show that $\text{Hom} \# (V, W)$ is a subrepresentation of $\text{Hom}_{k}(V, W)$: for $h \in H$, $f \in \text{Hom} \# (V, W)$, $a \in A$ and $v \in V$,

$$(h.f)(a.v) = h(1).f((S(h_{(3)})a)S(h_{(2)}).v)$$

$$= h(1).f((S(h_{(3)})a)S(h_{(2)}).v)$$

$$= (S(h_{(3)})a)S(h_{(2)}).f(S(h_{3})).v)$$

$$= (h_{(1)}S(h_{(3)})a)(h_{(2)}).f(v)$$

$$= (a.(h_{(2)}).f)(v),$$

where the second-to-last equality uses cocommutativity of $H$. Note also that

$$\text{Hom}_{A^H}(V, W) = \text{Hom}_{A}(V, W) \cap \text{Hom}_{H}(V, W)$$

$$= \text{Hom}_{A}(V, W) \cap \text{Hom}_{k}(V, W)^{H}$$

$$= \text{Hom}_{A}(V, W)^{H}. $$

If $A$ is commutative, then we may view the tensor product $V \otimes_{A} W$ as the quotient of $V \otimes W$ by the subspace that is generated by the tensors $a.v \otimes w - v \otimes a.w$. A straightforward calculation, again using (10.28) and cocommutativity of $H$, shows that this subspace is stable under the $H$-action (10.2) on $V \otimes W$. In this way, $V \otimes_{A} W$ becomes a left $A^H$-module. In particular, the tensor powers $V \otimes_{A}^{n}$ become left $A^H$-modules, and so do the symmetric and exterior powers $\text{Sym}_{A}^{n} V$ and $\Lambda_{A}^{n} V$ by the arguments in $\S 10.4.2$. These constructions are easily seen to be functorial, giving functors $\otimes_{A}^{n}$, $\text{Sym}_{A}^{n}$, $\Lambda_{A}^{n}$ : $A^H\text{Mod} \rightarrow A^H\text{Mod}$. After these preparations, we are now ready for the following result due to Ferrer Santos [72, Theorem 2.1], which is the technical core of this subsection.
Proposition 10.31. Assume that $H$ is cocommutative and let $A \in H\text{Alg}$ be cocommutative. Let $M \in A \# H\text{Mod}$ be such that $M \downarrow A$ is free of finite rank. Then the characteristic polynomial over $A$ of any operator in $\text{End}_{A \# H}(M)$ has coefficients in the invariant subalgebra $A^H$.

Proof. Let $f \in \text{End}_{A \# H}(M)$ and let $c_f \in A[t]$ denote the characteristic polynomial of $f$, viewed as an endomorphism of $\text{End}_A(M) \cong \text{Mat}_n(A)$, where we have put $n = \text{rank } M \downarrow A$. Thus, considering the $A[t]$-module $\mathbb{k}[t] \otimes M = M[t]$ and the endomorphism $\tilde{f} = t \text{Id}_{M[t]} - \text{Id}_{\mathbb{k}[t]} \otimes f \in \text{End}_{A[t]}(M[t]) \cong \text{Mat}_n(A[t])$, we have

$$c_f = \det \tilde{f} = \text{trace } \Lambda^n_{A[t]} \tilde{f}.$$ 

Putting $h.t^i = (e, h)t^i$, the polynomial algebra $A[t]$ becomes an $H$-module algebra with smash product $A[t] \# H \cong (A \# H)[t]$ and invariant algebra $A[t]^H = A^H[t]$. Furthermore, $M[t] \in A[t] \# H\text{Mod}$ and $\tilde{f} \in \text{End}_{A[t] \# H}(M[t])$. Therefore, by our remarks before the proposition,

$$\Lambda^n_{A[t]} \tilde{f} \in \text{End}_{A[t] \# H}(\Lambda^n_{A[t]} M[t]) = \text{End}_{A[t]}(\Lambda^n_{A[t]} M[t])^H.$$ 

Now $V := \Lambda^n_{A[t]} M[t]$ is free of rank 1 over $A[t]$ and so $\text{End}_{A[t]}(V) \cong A[t]$ via the trace. Explicitly, fixing a free generator $v$ of $V$ over $A[t]$, we have trace $\phi = b_\phi$ for $\phi \in \text{End}_{A[t]}(V)$, where $\phi(v) = b_\phi \cdot v$ and so $\phi(wv) = b_\phi \cdot wv$ for all $w \in V$. Finally, the trace is $H$-equivariant as the following calculation shows:

$$(h.\phi)(v) = h(1).\phi(S(h(2)).v) = h(1)b_\phi S(h(2)).v \overset{(10.28)}{=} (h.b_\phi).v.$$ 

Therefore, $c_f = \text{trace } \Lambda^n_{A[t]} \tilde{f} \in A[t]^H = A^H[t]$, finishing the proof. \qed

Integrality and Finite Generation of Invariants. We now come to the main result of this subsection. The reader may wish to refer to §2.2.7 for the notion of an integral ring extension. We note that Theorem 10.32 does not hold for arbitrary finite-dimensional Hopf algebras; see [209].

Theorem 10.32. Let $H$ be finite dimensional and cocommutative and let $A \in H\text{Alg}$ be cocommutative. Then $A$ is integral over the invariant subalgebra $A^H$. In particular, if $A$ is affine over $\mathbb{k}$, then so is $A^H$.

Proof. Consider the regular left $A \# H$-module, $M = (A \# H)_{\text{reg}}$, and let $r_a \in \text{End}_{A \# H}(M)$ denote the endomorphism that is given by right multiplication with a given $a \in A$. Since $M$ is free of rank $\dim_{\mathbb{k}} H$ as left $A$-module, we may apply Proposition 10.31 to obtain that the characteristic polynomial $c_{r_a}$ belongs to $A^H[t]$ and the Cayley-Hamilton Theorem further gives $0 = c_{r_a}(r_a)$. It follows that $c_{r_a}(a) = 0$, showing that $a$ satisfies a monic polynomial over $A^H$. This proves the first assertion. The second assertion now follows from standard facts from commutative algebra: if $A \in \text{Alg}_{\mathbb{k}}$ is affine commutative and integral over some subalgebra $B \subseteq A$, then
A is finitely generated as $B$-module (e.g., [65, Corollary 4.5]) and the Artin-Tate Lemma (Exercise 1.1.7) further tells us that $B$ is affine as well. □

### Exercises for Section 10.4

10.4.1 (Some identities for $H$-module algebras). Let $A \in H\text{Alg}$. Prove the identities $(h.a)b = h_{(1)}(a(S(h_{(2)}).b))$ and $a(h.b) = h_{(2)}((S^{-1}(h_{(1)}).a)b)$ for $a, b \in A$ and $h \in H$, assuming the composition inverse $S^{-1}$ exists for the second identity.

10.4.2 (The adjoint representation of the Sweedler algebra). Let $H = H_{-1}$ be the Sweedler algebra over a field $k$ of characteristic $\neq 2$ (Example 9.23). Recall that $H$ has $k$-basis $1, g, x, gx$ and algebra relations $g^2 = 1, x^2 = 0$ and $xg = -gx$. The coalgebra structure and antipode of $H_{-1}$ are given by $\Delta g = g \otimes g$, $\langle e, g \rangle = 1$, $Sg = g$ and $\Delta x = x \otimes 1 + g \otimes x$, $\langle e, x \rangle = 0$, $Sx = -gx$. Show:

(a) $\text{rad } H = xH = Hx$ and $H^{k,p} \cong k\langle g \rangle$, the group algebra of $\langle g \rangle \cong C_2$.

(b) $\text{Irr } H = \{1, k_g\}$, where $\alpha : H_4 \rightarrow k$ is given by $\langle \alpha, g \rangle = -1$, $\langle \alpha, x \rangle = 0$.

(c) $(H^{k,p})_{ad} \cong \mathbf{1} \otimes 2$ in $\text{Rep } H$.

(d) $H_{ad} = k1 \oplus kx \oplus V \cong \mathbf{1} \oplus k_{a'}/k \oplus P1$ with $V = k_{g} \oplus k_{gx} \cong P1$, the projective cover of $\mathbf{1}$. The representation $V$ is not semisimple and not self-dual, and hence neither is $H_{ad}$. Finally, $\text{Ker } H_{ad} = k\Lambda$ with $\Lambda = x + gx$.

10.4.3 ($H$-ideals and $H$-cores). Let $A \in H\text{Alg}$. An ideal $I$ of $A$ is called an $H$-ideal if $H.I \subseteq I$. It is easy to see that arbitrary intersections and sums as well as finite products of $H$-ideals are again $H$-ideals. For an arbitrary ideal $I$ of $A$, the sum of all $H$-ideals of $A$ that are contained in $I$ will be denoted by $I:H$ and called the $H$-core of $I$. Similarly, we define the $H$-hull of $I$ to be the intersection of all $H$-ideals that contain $I$. Show:

(a) $I:H = \{a \in A \mid H.a \subseteq I\}$.

(b) If $S$ is bijective, then the $H$-hull of $I$ is $H.I$. (Use Exercise 10.4.1.)

10.4.4 ($H$-primes). An $H$-module algebra $A$ is said to be $H$-prime if $A \neq 0$ and the product of any two nonzero $H$-ideals of $A$ is again nonzero. An $H$-ideal $I$ of $A$ is called $H$-prime if $A/I \in H\text{Alg}$ is $H$-prime. Show:

(a) $H$-cores of prime ideals of $A$ are $H$-prime.

(b) The converse of (a) holds if the ideal $(H.a)$ of $A$ is finitely generated for each $a \in A$: every $H$-prime ideal $I$ of $A$ has the form $I = P.H$ for some $P \in \text{Spec } A$ in this case. In particular, this holds whenever $A$ satisfies the maximum condition on ideals or the $H$-action on $A$ is locally finite.

10.4.5 ($H$-module algebras and convolution algebras). Let $A \in H\text{Alg}$ and consider the convolution algebra $\text{Hom}_k(H, A)$. Recall that $\text{Hom}_k(H, A) \in H\text{Alg}$. via
(§10.4.2) and that $\varepsilon^*: A \cong \text{Hom}_k(\mathbb{Z}, A) \to \text{Hom}_k(H, A)$, $\varepsilon^*(a) = (h \mapsto \langle \varepsilon, h \rangle a)$, is an algebra homomorphism (§9.1.4).

(a) Show that $\rho: A \to \text{Hom}_k(H, A)$, $\rho(a) = (h \mapsto h\cdot a)$, is a map in $\mathcal{H}\text{Alg}$.

(b) For any ideal $I$ of $A$, consider the canonical epimorphism $\pi_I: A \to A/I$ and the algebra map $(\pi_I)_*: \text{Hom}_k(H, A) \to \text{Hom}_k(H, A/I)$ (Exercise 9.2.1). Show that the kernel of $(\pi_I)_* \circ \rho: A \to \text{Hom}_k(H, A/I)$ is the $H$-core $I:H$.

(c) Define $\Phi: \text{Hom}_k(H, A) \to \text{Hom}_k(H, A)$ by $\Phi(f) = (h \mapsto h_{(1)} \cdot f(h_{(2)}))$. Show: $\Phi \circ \varepsilon^* = \rho$; $\Phi$ is invertible; and if $H$ is cocommutative, then $\Phi$ is an algebra automorphism.

10.4.6 (Comodule algebras and centralizing extensions). Let $A \in \text{Alg}^H_k$ with coaction $\delta: A \to A \otimes H$, $a \mapsto a_{(0)} \otimes a_{(1)}$. Show:

(a) The map $A \otimes H \to A \otimes H$, $a \otimes h \mapsto a_{(0)} \otimes a_{(1)} h$, is an isomorphism in $\text{Vect}_k$.

(b) If $H$ is commutative, then $\delta$ is a centralizing map in $\text{Alg}_k$ (Exercise 1.3.5).
This chapter merely gives a glimpse into what is a vast and interesting field. Aiming for little more than an explanation of the basic notions concerning affine algebraic groups along with a few examples, we will take a functorial approach, starting with affine group schemes rather than with affine algebraic varieties. This will allow us to make the most direct use of the material on Hopf algebras developed thus far. The connection between commutative Hopf algebras and affine algebraic groups laid out below is what gave rise to the moniker “quantum groups” for certain non-commutative Hopf algebras. As applications, we will prove a theorem of Chevalley, which mentioned earlier in connection with the Chevalley property (§10.2.5), and we will also briefly discuss algebraic group actions on noncommutative spectra.

For a more in-depth study of algebraic groups, the reader may wish to consult Borel [18], Hochschild [100], Humphreys [102] or Springer [189]. The short monograph [202] by Waterhouse gives an accessible and concise introduction to affine group schemes, while Jantzen [113] is an excellent source for more advanced material on the the representation theory of algebraic groups.

11.1. Affine Group Schemes

11.1.1. Group Functors from Hopf Algebras

Let $H \in \text{HopfAlg}_k$, with counit $\varepsilon$ and antipode $S$. By (9.35) and (9.30), the group-like elements of the Hopf algebra $H^\circ$ are exactly the maps $H \to k$ in $\text{Alg}_k$ and $G(H^\circ)$ is a subgroup of the group of units $(H^\circ)^\times$:

$$G(H^\circ) = \text{Hom}_{\text{Alg}}(H, k) \leq (H^\circ)^\times \leq (H^*)^\times.$$ 

Thus, the binary operation of $G(H^\circ)$ is given by restricting the convolution multiplication of the algebra $H^\circ$; the identity element is $1_{H^\circ} = \varepsilon$; and the inverse of
The group homomorphism. We have thus constructed a "group functor,"
\[ \Gamma \:
\text{scheme} \rightarrow \text{sets} \]

isomorphism of Hopf algebras (Exercise 11.1.1). Therefore, for any affine group of Hopf algebras, and if \( H \)
for \( g \in \text{Hom}_{\text{CommAlg}}(H, R) \) \( (\text{Exercise } 9.3.1) \). Thus, for any \( R \in \text{CommAlg}_{\mathbb{k}} \), we have defined a group,

\[ \Gamma_H(R) \overset{\text{def}}{=} \text{Hom}_{\text{CommAlg}}(H, R) \]

If \( f : R \rightarrow R' \) is a map in \( \text{CommAlg}_{\mathbb{k}} \), then \( f_* : \Gamma_H(R) \rightarrow \Gamma_H(R') \), \( g \mapsto f \circ g \), is a group homomorphism. We have thus constructed a "group functor,"

\[ \Gamma_H = \text{Hom}_{\text{CommAlg}}(H, \cdot) : \text{CommAlg}_{\mathbb{k}} \rightarrow \text{Groups} \]

The group \( \Gamma_H(R) \) is called the group of \textit{\textbf{R-points}} of the functor \( \Gamma_H \).

Defining \( \Gamma_A(R) \) in the same way for an arbitrary \( A \in \text{Alg}_{\mathbb{k}} \) results in a functor \( \Gamma_A = \text{Hom}_{\text{Alg}}(A, \cdot) : \text{CommAlg}_{\mathbb{k}} \rightarrow \text{Sets} \). Clearly, every algebra map \( A \rightarrow R \) with \( R \in \text{CommAlg}_{\mathbb{k}} \) vanishes on the ideal \( [A, A]A = A[A, A] \) of \( A \), and hence it factors through the canonical map \( A \rightarrow A^{ab} \), where \( A^{ab} \in \text{Alg}_{\mathbb{k}} \) is the abelianization of \( A \):

\[ A^{ab} \overset{\text{def}}{=} A/[A, A]A \]

In this way, we obtain an isomorphism of functors, \( \Gamma_A \cong \Gamma_A^{ab} (\S A.3.2) \). So there is no loss in replacing \( A \) by the commutative algebra \( A^{ab} \). If \( A = H \) is a Hopf algebra, then \( [H, H]H \) is in fact a Hopf ideal of \( H \), because \( [H, H] \) is a coideal. Therefore, \( H^{ab} \) is a commutative Hopf algebra. Consequently, in studying the group functors \( \Gamma_H \), we may assume that \( H \in \text{CommHopfAlg}_{\mathbb{k}} \), the category of commutative Hopf \( \mathbb{k} \)-algebras.

\subsection{11.1.2. Affine Group Schemes}

Functors \( \Gamma : \text{CommAlg}_{\mathbb{k}} \rightarrow \text{Sets} \) that are isomorphic to functors of the form \( \Gamma_A \) with \( A \in \text{CommAlg}_{\mathbb{k}} \) are called \textit{affine \( \mathbb{k} \)-schemes}. Our main interest is in \textit{affine group schemes} over \( \mathbb{k} \), that is, functors \( \Gamma : \text{CommAlg}_{\mathbb{k}} \rightarrow \text{Groups} \) such that \( \Gamma \cong \Gamma_H \) for some \( H \in \text{CommHopfAlg}_{\mathbb{k}} \). The Hopf algebra \( H \) is said to \textit{represent} the functor \( \Gamma \). In fact, the functor determines the Hopf algebra. For, if \( \alpha : \Gamma_H \Rightarrow \Gamma_K \) is a natural transformation of functors (\S A.3.1) with \( H, K \in \text{CommHopfAlg}_{\mathbb{k}} \), then \( \alpha \) gives rise to the group homomorphism \( a_H : \Gamma_H(H) \rightarrow \Gamma_K(H) \). The image of \( \text{Id}_H \in \Gamma_H(H) \) is a \( \mathbb{k} \)-algebra map \( a_H(\text{Id}_H) : K \rightarrow H \) that is actually a homomorphism of Hopf algebras, and if \( \alpha \) is an isomorphism of functors, then \( a_H(\text{Id}_H) \) is an isomorphism of Hopf algebras (Exercise 11.1.1). Therefore, for any affine group scheme \( \Gamma : \text{CommAlg}_{\mathbb{k}} \rightarrow \text{Groups} \), the representing commutative Hopf algebra \( H \)
is determined up to isomorphism; it will be denoted by
\[ O(\Gamma) \in \text{CommHopfAlg}_k. \]

Properties of \( \Gamma \) may be defined though properties of the algebra \( O(\Gamma) \). For example, the following definitions are standard:

- \( \Gamma \) is said to be \textit{finite} \( \iff \) \( O(\Gamma) \) is finite dimensional ("finite")
- \( \Gamma \) is said to be \textit{algebraic} \( \iff \) \( O(\Gamma) \) is affine (finitely generated)
- \( \Gamma \) is said to be \textit{reduced} \( \iff \) \( O(\Gamma) \) is reduced
- \( \Gamma \) is said to be \textit{integral} \( \iff \) \( O(\Gamma) \) is an integral domain

Recall that a commutative \( k \)-algebra is said to be \textit{reduced} if it has no nonzero nilpotent elements.

### 11.1.3. Categorical View

We can form a category, \( \text{AffineGroupSchemes}_k \), by taking all affine group schemes over \( k \) as objects and defining the morphisms to be the same as natural transformations of functors,

\[
\begin{array}{ccc}
\text{CommAlg}_k & -\ni- & \text{Groups} \\
\Gamma & \downarrow & \alpha \\
\Gamma' & \nearrow & 
\end{array}
\]

Associating the affine group scheme \( \Gamma_H \) to the commutative Hopf algebra \( H \) and the natural transformation \( f^* : \Gamma_H = \text{Hom}_{\text{Alg}}(H, \cdot) \Rightarrow \text{Hom}_{\text{Alg}}(K, \cdot) = \Gamma_K \) to a Hopf algebra homomorphism \( f : K \to H \), we obtain an equivalence of categories (see Exercise 11.1.1),

\begin{equation}
\text{CommHopfAlg}_k \equiv \left( \text{AffineGroupSchemes}_k \right)^{\text{op}}
\end{equation}

### 11.1.4. Some Examples of Affine Group Schemes

**Example 11.1** (The affine group scheme defined by a group algebra). Taking \( H \) to be the group algebra \( kG \) of an arbitrary (for now) group \( G \) in the foregoing, the resulting functor is

\[
\Gamma_{kG}(R) = \text{Hom}_{\text{Alg}}(kG, R) \cong \text{Hom}_{\text{Groups}}(G, R^\times) \cong \text{Hom}_{\text{Groups}}(G^{\text{ab}}, R^\times),
\]

where \( G^{\text{ab}} = G/[G, G] \) is the abelianization of \( G \). Thus, \( (kG)^{\text{ab}} = k[G^{\text{ab}}] \) and \( \Gamma_{kG} \cong \Gamma_{k[G^{\text{ab}}]} \). Let us consider some special cases in detail. For \( G = C_\infty = \langle x \rangle \), the infinite cyclic group, the group algebra \( kC_\infty \) is isomorphic to the Laurent polynomial

---

\( ^1 \)This definition of a category may require some extra care with regard to the foundational aspects of mathematics. Instead of working with all commutative \( k \)-algebras, we should perhaps only consider those in some fixed "universe" (as in Demazure and Gabriel [54]). We shall ignore these subtleties below.
algebra $\mathbb{k}[x^{\pm 1}]$ and $\text{Hom}_{\text{Groups}}(C_{\infty}, R^x) \cong R^x$ via $f \leftrightarrow fx$. In this way, we obtain the \textit{multiplicative group} of units,
\[ G_m \overset{\text{def}}{=} R^x \cong \Gamma_{\mathbb{k}[x^{\pm 1}]} . \]
For the cyclic group of order $n$, $\text{Hom}_{\text{Groups}}(C_n, R^x) \cong \mu_n(R) = \{ r \in R \mid r^n = 1 \}$, the subgroup of $n^{\text{th}}$ \textit{roots of unity} of $R^x$. Thus,
\[ \Gamma_{\mathbb{k}C_n} \cong \mu_n . \]

The canonical group epimorphism $C_{\infty} \to C_n$ lifts uniquely to an epimorphism $\mathbb{k}C_{\infty} \to \mathbb{k}C_n$ of Hopf $\mathbb{k}$-algebras, and this map in turn corresponds to the morphism of functors, $\mu_n \cong \Gamma_{\mathbb{k}C_n} \to \Gamma_{\mathbb{k}C_n} \cong R^x$, which is given by the inclusions $\mu_n(R) \hookrightarrow R^x$ for $R \in \text{CommAlg}_\mathbb{k}$. Since these inclusions are monomorphism of groups, one says that $\mu_n$ is a \textit{subgroup scheme} of $R^x$.

\textbf{Example 11.2} (The affine group scheme defined by an enveloping algebra). The enveloping algebra $U_\mathfrak{g}$ of a Lie algebra $\mathfrak{g}$ gives the functor
\[ \Gamma_{U_\mathfrak{g}}(R) = \text{Hom}_{\text{Alg}_R}(U_\mathfrak{g}, R) \overset{(5.15)}{=} \text{Hom}_{\text{Lie}_R}(\mathfrak{g}, R_{\text{Lie}}) \equiv \text{Hom}_R(\mathfrak{g}_{\text{ab}}, R_{\text{Lie}}), \]
where $\mathfrak{g}_{\text{ab}} = \mathfrak{g}/[\mathfrak{g}, \mathfrak{g}]$ is the abelianization of $\mathfrak{g}$. Thus, $(U_\mathfrak{g})_{\text{ab}} = U(\mathfrak{g}_{\text{ab}})$. The enveloping algebra of the 1-dimensional Lie algebra $\mathfrak{g} = \mathbb{k}x$ is isomorphic to the polynomial algebra $\mathbb{k}[x]$ and $\text{Hom}_{\text{Lie}_R}(\mathfrak{g}, R_{\text{Lie}}) \equiv (R, +)$ via $f \leftrightarrow fx$. This leads to the \textit{additive group} functor,
\[ G_a \overset{\text{def}}{=} (., +) \equiv \Gamma_{\mathbb{k}[x]} . \]

\textbf{Example 11.3} (GL$_n$ and GL$_F$). In Example 9.20, we had defined the commutative Hopf algebra
\[ O(\text{GL}_n) = \mathbb{k}[X_{ij} \mid 1 \leq i, j \leq n][D^{-1}] , \]
where $D = \sum_{s \in S_n} \text{sgn}(s) X_{1s(1)} X_{2s(2)} \ldots X_{ns(n)}$ is the determinant. The comultiplication and counit are given by $\Delta X_{ij} = \sum_k X_{ik} \otimes X_{kj}$ and $\langle e, X_{ij} \rangle = \delta_{ij}$, respectively, and the antipode by $\text{SX} = X^{-1} = D^{-1}C^r$, where $X = (X_{ij})_{i,j}$ is the generic $n \times n$-matrix and $C$ is the matrix of its cofactors. The resulting affine group scheme $\Gamma_{O(\text{GL}_n)}$ is isomorphic to the \textit{general linear group} functor, $\text{GL}_n$, thereby ensuring consistency of our notations. Indeed, any $f \in \Gamma_{O(\text{GL}_n)}(R) = \text{Hom}_{\text{Alg}_R}(O(\text{GL}_n), R)$ is determined by the matrix $(f(X_{ij}))_{i,j} \in \text{Mat}_n(R)$ and the only constraint is that the determinant of this matrix be invertible in $R$. Furthermore, the comultiplication, counit and antipode of $O(\text{GL}_n)$ translate into matrix multiplication, the identity matrix and matrix inversion, respectively. Thus, for any commutative $\mathbb{k}$-algebra $R,$
\[ \Gamma_{O(\text{GL}_n)}(R) \equiv \text{GL}_n(R) = \{ M \in \text{Mat}_n(R) \mid \det M \in R^x \} . \]
For \( n = 1 \), we obtain the multiplicative group functor (Example 11.1): \( \text{GL}_1 \cong \mathbb{G}_m \). More generally, for any finite-dimensional \( V \in \text{Vect}_k \), we obtain an affine group scheme \( \text{GL}_V \) over \( k \) by putting
\[
\text{GL}_V(R) = \text{Aut}_R(V \otimes R),
\]
the automorphism group of \( V \otimes R \in R\text{-Mod} \). Any choice of \( k \)-basis for \( V \) results in an isomorphism of functors \( \text{GL}_V \cong \text{GL}_n \) with \( n = \dim_k V \).

**Example 11.4 (\( \text{SL}_n \))**. Similarly, the quotient \( \text{O}(\text{SL}_n) = \text{O}(\text{GL}_n)/(D-1) \) of \( \text{O}(\text{GL}_n) \) (Example 9.21) leads to the **special linear group** functor \( \text{SL}_n \):
\[
\Gamma_{\text{O}(\text{SL}_n)}(R) = \text{SL}_n(R) = \{ M \in \text{Mat}_n(R) \mid \det M = 1_R \}.
\]
The Hopf algebra epimorphism \( \text{O}(\text{GL}_n) \twoheadrightarrow \text{O}(\text{SL}_n) \) corresponds to a morphism of group schemes, \( \text{SL}_n \to \text{GL}_n \), exhibiting \( \text{SL}_n \) as a subgroup scheme of \( \text{GL}_n \). The Hopf map \( kC_\infty \to \text{O}(\text{GL}_n) \) that is given by sending a fixed generator of \( C_\infty \) to the group-like element \( D \in \text{O}(\text{GL}_n) \) corresponds to a morphism of group schemes, \( \det : \text{GL}_n \to \mathbb{G}_m \); see also Example A.1.

**Example 11.5 (Constant group schemes)**. Let \( G \) be a finite group and consider the Hopf algebra \( H = (\mathbb{G}_k)^G \) with its standard \( k \)-basis \( (\delta_x)_{x \in G} \) consisting of orthogonal idempotents such that \( \sum_{x \in G} \delta_x = 1 \) (Example 9.17). A commutative algebra \( R \in \text{CommAlg}_k \) is said to be **connected** if 0 and 1 are the only idempotents of \( R \) (see Exercise 11.1.5). In this case, for any \( f \in \Gamma_{(\mathbb{G}_k)^G}(R) = \text{Hom}_{\text{Alg}_k}(\mathbb{G}_k, R) \), we must have \( f(\delta_x) = 1_R \) for exactly one \( x \in G \) and \( f(\delta_y) = 0 \) for all \( y \neq x \). This gives a bijection between \( \Gamma_{(\mathbb{G}_k)^G}(R) \) and \( G \), which is easily seen to be a group homomorphism. Thus, for any connected \( R \in \text{CommAlg}_k \),
\[
\Gamma_{(\mathbb{G}_k)^G}(R) \cong G.
\]
If \( R \cong \prod_{i \in I} R_i \) for connected commutative \( k \)-algebras \( R_i \), then \( \Gamma_{(\mathbb{G}_k)^G}(R) = \prod_{i \in I} \Gamma_{(\mathbb{G}_k)^G}(R_i) = G^{|I|} \). The last observation also shows that no affine \( k \)-scheme can truly be constant.

**Exercises for Section 11.1**

11.1.1 (Yoneda Lemma). (a) Let \( \Gamma_A, \Gamma_B : \text{CommAlg}_k \to \text{Sets} \) be functors represented by \( A, B \in \text{CommAlg}_k \). Show that every algebra map \( f : B \to A \) gives rise to a natural transformation \( f^* : \Gamma_A \Rightarrow \Gamma_B \). Moreover, every natural transformation \( \Gamma_A \Rightarrow \Gamma_B \) arises in this way. Finally, \( f^* \) is an isomorphism of functors if and only if \( f \) is an isomorphism of algebras.

(b) Now let \( H \) and \( K \) be commutative Hopf \( k \)-algebras and consider the functors \( \Gamma_H, \Gamma_K : \text{CommAlg}_k \to \text{Groups} \). Show that the natural transformations \( \Gamma_H \Rightarrow \Gamma_K \) correspond as in (a) to Hopf algebra maps \( K \to H \).
11. Affine Algebraic Groups

11.1.2 (Products). (a) Let $H, K \in \text{HopfAlg}_k$. Show that $H \otimes K$ becomes a Hopf algebra with the usual algebra structure (1.3), the coalgebra structure of Exercise 9.1.11, and the antipode $S_{H \otimes K} = S_H \otimes S_K$.

(b) Let $A, B \in \text{CommAlg}_k$. Show that $\Gamma_{A \otimes B} \cong \Gamma_A \times \Gamma_B$, where $\Gamma_A \times \Gamma_B$ is defined by the direct product of sets: $(\Gamma_A \times \Gamma_B)(R) := \Gamma_A(R) \times \Gamma_B(R)$ for $R \in \text{CommAlg}_k$. If $A$ and $B$ are Hopf algebras, then this is an isomorphism of groups.

11.1.3 (Another example). Let $\Gamma : \text{CommAlg}_k \to \text{Groups}$ be the functor that is given by the semidirect product $\Gamma(R) = R \rtimes R^\times$ with $R^\times$ acting by multiplication on $R$; so $(r, u)(r', u') = (r + ur', uu')$ for $r, r' \in R$ and $u, u' \in R^\times$. Show that $\Gamma$ is an affine group scheme and find the representing Hopf algebra $H \in \text{CommHopfAlg}_k$.

11.1.4 (A non-example). Let $\mu : \text{CommAlg}_k \to \text{Groups}$ be the functor that is defined by $\mu(R) = \{ r \in R \mid r^n = 1 \text{ for some } n \in \mathbb{N} \}$. Show that $\mu$ is not an affine group scheme over $k$.

11.1.5 (Connectedness). Let $R$ be a commutative $k$-algebra. Show that $\text{Spec } R$ is connected for the Jacobson-Zariski topology if and only if $R$ has only the trivial idempotents, 0 and 1. (Use Lemma 2.6.)

11.2. Affine Algebraic Groups

For the remainder of this chapter, with the exception of §11.6.5, the base field $k$ is understood to be algebraically closed.

11.2.1. The Definition

An affine algebraic $k$-group, by definition, is the group $\Gamma(k)$ of $k$-points of an affine group scheme $\Gamma$ over $k$ that is assumed to be algebraic. To unfold this definition, recall that our hypothesis on $\Gamma$ means that $\Gamma \cong \Gamma_H = \text{Hom}_{\text{Alg}_k}(H, \cdot)$ with $H = O(\Gamma) \in \text{CommHopfAlg}_k$ being affine, that is, finitely generated as $k$-algebra. The isomorphism of group functors $\Gamma \cong \Gamma_H$ gives rise to a group isomorphism of $k$-points, $\Gamma(k) \cong \Gamma_H(k)$. Thus, an affine algebraic $k$-group, $G$, has the following form, for an affine $H \in \text{CommHopfAlg}_k$:

$$G = \Gamma(k) \cong \Gamma_H(k) = (9.35) \Gamma(H)$$

Whereas the group scheme $\Gamma$ determines the isomorphism type of the representing Hopf algebra $O(\Gamma)$ as long as it is commutative, its group of $k$-points $\Gamma(k)$ does not. In fact, in addition to our hypothesis that $\Gamma$ is affine algebraic, we may also assume that $\Gamma$ is reduced, that is, $O(\Gamma)$ has no nonzero nilpotent elements, without altering $\Gamma(k)$:

**Proposition 11.6.** Let $G$ be an affine algebraic $k$-group. Then $G = \Gamma(k)$ for an affine algebraic $k$-group scheme $\Gamma$ that is reduced: $O(\Gamma)$ embeds into $(kG)^\circ$. 
Proof. Recall from Example 9.17 that \((kG)\circ\), the Hopf algebra of representative functions of \(G\), is a subalgebra of the algebra \(kG\) of all functions \(G \rightarrow k\) with pointwise addition and multiplication. Consequently, \((kG)\circ\) is reduced.

By definition, \(G \cong G(H)\), where \(H = O(\Gamma) \in \text{CommHopfAlg}_k\) is affine. Thus, the group algebra \(kG\) embeds as a Hopf subalgebra into \(H\) (§9.3.4). By Proposition 9.16, this embedding corresponds to a Hopf algebra homomorphism \(H \rightarrow (kG)\circ\). Explicitly, viewing elements \(g \in G\) as algebra maps \(H \rightarrow k\), \(h \mapsto \langle g, h \rangle\), by means of the isomorphism \(G \cong G(H) = \text{Hom}_{\text{Alg}}(H, k)\), this map is given by
\[
H \rightarrow (kG)\circ \hookrightarrow k^G, \quad h \mapsto \langle (g, h) \rangle_{g \in G}.
\]
The image of \(H\) is a Hopf subalgebra \(\overline{H} \subseteq (kG)\circ\) and all \(g \in G\) factor through the canonical epimorphism \(f : H \rightarrow \overline{H}\). Thus, we have a group isomorphism \((f^*)_k : \Gamma \overline{H}(k) \cong \Gamma(k) = G\) coming from the natural transformation \(f^* : \Gamma \overline{H} \Rightarrow \Gamma_H \cong \Gamma\). Replacing \(\Gamma\) by subfunctor \(\Gamma \subseteq \Gamma\) that is given by \(f^*\), we obtain a reduced affine \(k\)-group scheme \(\Gamma\) with \(G = \overline{\Gamma}(k)\) and a Hopf algebra embedding \(\overline{H} \cong O(\Gamma) \hookrightarrow (kG)\circ\).

From now on, we assume that every affine algebraic \(k\)-group is given as \(G = \Gamma(k)\), the group of \(k\)-points of an affine \(k\)-group scheme \(\Gamma\) that is both algebraic and reduced. More generally, if a group \(G\) is given along with a group isomorphism \(G \cong \Gamma(k)\), then \(G\) will also be regarded as an affine algebraic \(k\)-group via this isomorphism.

11.2. Affine Algebraic Groups

A homomorphism (or map) of affine algebraic \(k\)-groups is defined to be a group homomorphism \(\phi_k : G = \Gamma(k) \rightarrow D = \Delta(k)\) that arises from a natural transformation \(\phi : \Gamma \Rightarrow \Delta\) of reduced algebraic affine \(k\)-group schemes. By (11.1), any such \(\phi\) corresponds to a Hopf algebra map \(f : O(\Delta) \rightarrow O(\Gamma)\) via
\[
\phi : \Gamma \cong \text{Hom}_{\text{Alg}}(O(\Gamma), \cdot) \Rightarrow \text{Hom}_{\text{Alg}}(O(\Delta), \cdot) \cong \Delta.
\]
We have thus arrived at yet another category, \(\text{AffineAlgebraicGroups}_k\).

The composite of the group-like functor \(G : \text{HopfAlg}_k \rightarrow \text{Groups}\) with the finite dual functor \(^\circ : \text{HopfAlg}_k \rightarrow (\text{HopfAlg}_k)^{\text{op}}\) is a functor \(G \circ \circ : \text{HopfAlg}_k \rightarrow (\text{Groups})^{\text{op}}\). Restriction to the full subcategory \(\text{AffineReducedCommHopfAlg}_k\) of \(\text{HopfAlg}_k\) gives the functor
\[
F = G \circ \circ : \text{AffineReducedCommHopfAlg}_k \rightarrow (\text{AffineAlgebraicGroups}_k)^{\text{op}}.
\]
Indeed, \(FH = \Gamma_H(k)\) is an affine algebraic \(k\)-group for each affine reduced commutative Hopf \(k\)-algebra \(H\) and, for any map \(f : H \rightarrow K\) in \(\text{AffineReducedCommHopfAlg}_k\), the map \(Ff = (f^*)_k : FK = G(K^\circ) \rightarrow FH = G(H^\circ)\) is a map of affine algebraic \(k\)-groups.
Lemma 11.7. The functor $F = G \circ \cdot^\circ$ gives an equivalence of categories,
\[ \text{AffineReducedCommHopfAlg}_k \cong (\text{AffineAlgebraicGroups}_k)^{\text{op}}. \]

Proof. Let us abbreviate the category on the left by $\text{AH}$ and the one on the right by $\text{AG}^{\text{op}}$. The functor $F: \text{AH} \to \text{AG}^{\text{op}}$ is essentially surjective: every $G \in \text{AG}$ is isomorphic in $\text{AG}$ to an object of the form $FH = G(H')$ with $H' \in \text{AH}$. Furthermore, for each pair of objects $H, K \in \text{AH}$, the map $\text{Hom}_{\text{AH}}(H, K) \to \text{Hom}_{\text{AG}}(FK, FH)$ that is given by $F$ is surjective, by the definition of maps in $\text{AG}$. To establish the desired category equivalence, it suffices to show that this map is also injective (§A.3.3).

But, for distinct $f, f' \in \text{Hom}_{\text{AH}}(H, K)$, there is some $x \in H$ so that $fx \neq f'x$ in $K$. Since $K$ is a reduced affine commutative $\mathbb{k}$-algebra and $\mathbb{k}$ is algebraically closed, the Nullstellensatz (Section C.1) implies that $\bigcap g \ker g = 0$, where $g$ runs over $FK = \text{Hom}_{\text{Alg}}(K, \mathbb{k})$. Therefore, for some $g$, we must have $(g \circ f)x \neq (g \circ f')x$. Thus, $Ff(g) \neq Ff'(g)$ and so $Ff \neq Ff'$, as was to be shown. \(\square\)

Let $O$ denote a quasi-inverse for the functor $F = G \circ \cdot^\circ$. Thus, for any affine algebraic $\mathbb{k}$-group $G$, we have a reduced affine commutative Hopf algebra, $O(G)$, and
\[ G \equiv F(O(G)) = G(O(G)^\circ) = \text{Hom}_{\text{Alg}}(O(G), \mathbb{k}). \]

Each element $g \in G$ may thus be viewed as a $\mathbb{k}$-valued function on $O(G)$, which we will generally write as $h \mapsto \langle g, h \rangle$ ($h \in O(G)$). We also have an embedding
\[ \mathbb{k}G \cong \mathbb{k}[G(O(G)^\circ)] \hookrightarrow O(G)^\circ \hookrightarrow O(G)^\circ. \]

By Proposition 9.16, the map $\mathbb{k}G \hookrightarrow O(G)^\circ$ corresponds to a Hopf algebra map $O(G) \to (\mathbb{k}G)^\circ$, which is in fact injective by the Nullstellensatz, being given by
\[ O(G) \hookrightarrow (\mathbb{k}G)^\circ \hookrightarrow \mathbb{k}G, \quad h \mapsto \langle (g, h) \rangle_{g \in G}. \]

Thus, we may also regard the Hopf algebra $O(G)$ as a subalgebra of the algebra $\mathbb{k}G$ of all functions $G \to \mathbb{k}$. Taking the latter point of view, we will write $h(g) = \langle g, h \rangle \in \mathbb{k}$. A homomorphism $\varphi: G \to E$ of affine algebraic $\mathbb{k}$-groups then corresponds to the Hopf algebra map $O\varphi = \varphi^\circ: O(E) \to O(G)$, $h \mapsto h \circ \varphi$. It can happen that a homomorphism $\varphi$ of affine algebraic $\mathbb{k}$-groups is an isomorphism in Groups but the inverse $\varphi^{-1}$ is not a homomorphism of affine algebraic $\mathbb{k}$-groups (Exercise 11.2.1).

11.2.3. First Examples

Example 11.8 ($\text{GL}_n(\mathbb{k})$ and $\text{SL}_n(\mathbb{k})$). The Hopf algebras $O(\text{GL}_n)$ and $O(\text{SL}_n)$ in Examples 9.20/11.3 and 9.21/11.4 are affine commutative, and hence the group schemes $\text{GL}_n \cong \Gamma_{O(\text{GL}_n)}$ and $\text{SL}_n \cong \Gamma_{O(\text{SL}_n)}$ are algebraic. Moreover, $O(\text{GL}_n)$ and $O(\text{SL}_n)$ are reduced; in fact, they are both integral domains.\(^2\) Thus, $\text{GL}_n(\mathbb{k})$ is an

\(^2\)For $O(\text{SL}_n)$, this amounts to irreducibility of the polynomial $D - 1 \in \mathbb{k}[X_{ij} \mid 1 \leq i, j \leq n]$, where $D$ is the determinant, which follows from irreducibility of $D$. See [17, Section 61].
affine algebraic \( \mathbb{k} \)-group with \( O(\text{GL}_n(\mathbb{k})) \equiv O(\text{GL}_n) \) and likewise for \( \text{SL}_n(\mathbb{k}) \). The inclusion map \( \text{SL}_n(\mathbb{k}) \hookrightarrow \text{GL}_n(\mathbb{k}) \) comes from the Hopf algebra map \( O(\text{GL}_n) \to O(\text{SL}_n) = O(\text{GL}_n)/(D - 1) \); so this inclusion is a homomorphism of affine algebraic \( \mathbb{k} \)-groups. The determinant \( \det : \text{GL}_n(\mathbb{k}) \to \mathbb{k}^\times \equiv \text{GL}_1(\mathbb{k}) \) is a homomorphism of affine algebraic \( \mathbb{k} \)-groups, because it comes from the Hopf map \( \det^* : O(\text{GL}_1) \equiv \mathbb{k}[x^\pm 1] \hookrightarrow O(\text{GL}_n), x \mapsto D \). More generally, for any finite-dimensional \( \mathbb{V} \in \text{Vec}_\mathbb{k} \), the group \( \text{GL}_n(\mathbb{k}) = \text{GL}(\mathbb{V}) \) is affine algebraic and a choice of basis for \( \mathbb{V} \) gives an isomorphism \( \text{GL}(\mathbb{V}) \equiv \text{GL}_n(\mathbb{k}) \).

**Example 11.9** (Finite groups). For any finite group \( G \), the Hopf algebra \( H = (\mathbb{k}G)^* \) is evidently affine, commutative and reduced. Furthermore, as was explained in Example 11.5, \( \Gamma_H(\mathbb{k}) \equiv G \). Indeed, the isomorphism \( H^\ast = (\mathbb{k}G)^\ast \equiv \mathbb{k}G \) gives \( \Gamma_H(\mathbb{k}) = G(H^\ast) \equiv G(\mathbb{k}G) = G \). Thus, \( G \) is an affine algebraic \( \mathbb{k} \)-group with \( O(G) \equiv (\mathbb{k}G)^\ast \). Furthermore, any homomorphism \( \varphi : G \to \mathbb{k}E \) of finite groups gives rise to the Hopf algebra maps \( \mathbb{k}\varphi : \mathbb{k}G \to \mathbb{k}E \) and \( (\mathbb{k}E)^\ast \to (\mathbb{k}G)^\ast \). The latter homomorphism is the map denoted \( O\varphi = \varphi^* \) above. So arbitrary homomorphism of finite groups are in fact maps in the category of affine algebraic \( \mathbb{k} \)-groups.

**Example 11.10** (Algebraic tori). The group algebra \( H = \mathbb{k}L \) of the lattice \( L = \mathbb{Z}^n \) is an affine commutative Hopf algebra that is a domain. The affine algebraic group associated to \( H \) has the following form (Example 11.1):

\[
G(H) \equiv \text{Hom}_{\text{Groups}}(L, \mathbb{k}^\times) \equiv (\mathbb{k}^\times)^n.
\]

Thus, the group \( G = (\mathbb{k}^\times)^n \) is affine algebraic, called the **algebraic n-torus** over \( \mathbb{k} \), and \( O(G) \equiv \mathbb{k}L \equiv \mathbb{k}[x_1^{\pm 1}, \ldots, x_n^{\pm 1}] \). Viewing elements of \( \mathbb{k}L \) as a \( \mathbb{k} \)-valued functions on \( G \) as in (11.4), the variable \( x_i \) is the \( i \)th coordinate function, \( x_i(g) = \gamma_i \) for \( g = (\gamma_1, \ldots, \gamma_n) \in G \), and \( \lambda = (z_1, \ldots, z_n) \in L \) is the function given by \( \lambda(g) = \prod_i \gamma_i^{z_i} \).

**Example 11.11** (Direct products). Let \( G \) and \( E \) be affine algebraic \( \mathbb{k} \)-groups. Then \( O(G) \otimes O(E) \) carries a natural Hopf algebra structure and there is a natural group isomorphism (Exercise 11.1.2),

\[
\Gamma_{O(G) \otimes O(E)}(\mathbb{k}) \equiv \Gamma_{O(G)}(\mathbb{k}) \times \Gamma_{O(E)}(\mathbb{k}) \equiv G \times E.
\]

Furthermore, as we shall prove in Lemma 11.19, \( O(G) \otimes O(E) \) is reduced. Therefore, the direct product \( G \times E \) is an affine algebraic \( \mathbb{k} \)-group with \( O(G \times E) \equiv O(G) \otimes O(E) \). As a function on \( G \times E \), the tensor \( h \otimes k \in O(G) \otimes O(E) \) is given by

\[
(h \otimes k)(g, e) = h(g)k(e) \quad (g \in G, e \in E).
\]

For example, the \( n \)-torus \( (\mathbb{k}^\times)^n \) is isomorphism to the \( n \)-fold direct product of the multiplicative group \( \mathbb{k}^\times \equiv \text{GL}_1(\mathbb{k}) \) with itself via the Hopf algebra isomorphism \( O((\mathbb{k}^\times)^n) \equiv \mathbb{k}[x_1^{\pm 1}, \ldots, x_n^{\pm 1}] \equiv \mathbb{k}[x_1^{\pm 1}]^\otimes n \) (Example 11.10).
Exercises for Section 11.2

Recall that the base field \( k \) is understood to be algebraically closed.

11.2.1 (A non-invertible bijective endomorphism). Assume that \( \text{char } k = p > 0 \) and consider the map \( \varphi : k \to k, \lambda \mapsto \lambda^p \); this is an automorphism of \((k, +)\) in Groups. Show that \( \varphi \) is an endomorphism of the affine algebraic \( k \)-group \( G_a(k) = (k, +) \) (Example 11.2) but not \( \varphi^{-1} \).

11.2.2 (Additive and multiplicative group). Consider the affine algebraic \( k \)-groups \( G_m(k) = k^\times \) and \( G_a(k) = (k, +) \) (Examples 11.1 and 11.2). Show that \( \text{Aut } G_m(k) \cong C_2 \) whereas \( \text{Aut } G_a(k) \cong k^\times \), where \( \text{Aut } = \text{Aut } \text{AffineAlgebraicGroups}_k \).

11.2.3 (Further examples). Show that the following subgroups of \( \text{GL}_2(k) \) are affine algebraic \( k \)-groups: the group \( G \) consisting of all matrices \( \begin{pmatrix} \lambda & \mu \\ 0 & \lambda' \end{pmatrix} \) with \( \lambda, \lambda' \in k^\times \) and \( \mu \in k \); the subgroup \( G_1 \leq G \) consisting of the matrices with \( \lambda = \lambda' \); the subgroup \( G_2 \leq G \) consisting of all matrices with \( \lambda' = 1 \); and the subgroup \( G_3 \leq G \) consisting of all matrices with \( \lambda = \lambda' = 1 \). (See Exercise 11.1.3 for \( G_2 \).)

11.2.4 (Some general maps). Let \( G \) be an affine algebraic \( k \)-group. Show:

(a) The opposite group \( G^{\text{op}} \) is an affine algebraic \( k \)-group. (See Exercise 9.1.7.)

(b) The map \( G^{\text{op}} \to G, g \mapsto g^{-1} \), is an isomorphism in \( \text{AffineAlgebraicGroups}_k \) and \( G \to G \times G, g \mapsto (g, g) \) is also a map in \( \text{AffineAlgebraicGroups}_k \).

11.2.5 (The reduced quotient of a commutative Hopf algebra). Let \( H \in \text{HopfAlg}_k \) be commutative and let \( \sqrt{0} \) denote the ideal of \( H \) consisting of all nilpotent elements of \( H \). Show that \( \sqrt{0} \) is a Hopf ideal of \( H \); so \( H^{\text{red}} = H/\sqrt{0} \) is a reduced Hopf algebra.

11.3. Representations and Actions

We now turn to representations of an affine algebraic \( k \)-group \( G \) and to actions of \( G \) on \( k \)-algebras (by automorphisms). Inasmuch as \( G \) originates from the Hopf algebra \( O(G) \), being isomorphic to \( G(O(G)^\circ) \), it is only natural that the “rational” representations and actions of \( G \) that feature foremost here also come from \( O(G) \).

Throughout this section, \( G \) is an affine algebraic \( k \)-group and \( k \) is algebraically closed.

11.3.1. Rational Representations

A representation \( V \in \text{Rep } kG \) is said to be rational if the \( G \)-action on \( V \) arises by restriction from a right \( O(G) \)-comodule structure on \( V \). In detail, by Proposition 9.12, a right \( O(G) \)-comodule structure makes \( V \) a (locally finite) left module over the algebra \( O(G)^\circ \), and hence \( V \) is a left \( kG \)-module via the embedding \( kG \hookrightarrow O(G)^\circ \).
(11.3). Writing the right $O(G)$-coaction $\delta: V \to V \otimes O(G)$ as $v \mapsto v_0 \otimes v_1$ as usual, (9.24) gives the following formula for the action $G \subset V$:

$$g \cdot v = v_0(g, v_1) \quad (g \in G, v \in V).$$

(11.5)

**Proposition 11.12.**

(a) Let $V, W \in \text{Rep} \mathbb{k}G$ be rational. A $\mathbb{k}$-linear map $f: V \to W$ is a map in $\text{Rep} \mathbb{k}G$ if and only if it is a map in $\text{Mod}^{O(G)}$.

(b) All subrepresentations and all homomorphic images in $\text{Rep} \mathbb{k}G$ of rational representations are rational. Furthermore, tensor products and arbitrary direct sums of rational representations are rational.

(c) All rational representations of $G$ are locally finite.

**Proof.** (a) The map $f$ is a morphism in $\text{Rep} \mathbb{k}G$ if and only if $f(g \cdot v) = g \cdot f(v)$ for all $g \in G$ and $v \in V$, which by (11.5) can be written as

$$f(v_0)\langle g, v_1 \rangle = f(v)\langle g, f(v)_1 \rangle \quad (g \in G, v \in V).$$

Similarly, $f$ is a morphism in $\text{Mod}^{O(G)}$ if and only if

$$f(v_0) \otimes v_1 = f(v) \otimes f(v)_1 \quad (v \in V).$$

Clearly, the latter condition implies the former. For the converse, note that the first condition states that the images of $f(v_0)\otimes v_1$, $f(v_0) \otimes f(v)_1 \in W \otimes O(G)$ under the embedding $W \otimes O(G) \hookrightarrow W \otimes \mathbb{k}G$ given by (11.4) agree, and hence the two elements must be the same.

(b) Let $V \in \text{Rep} \mathbb{k}G$ be rational, with $G \subset V$ given by the coaction $\delta v = v_0 \otimes v_1$ as above, and let $U \subset V$ be a subrepresentation. We need to show that $\delta u = u_0 \otimes u_1 \in U \otimes O(G)$ for all $u \in U$. Consider the following commutative diagram of inclusions, the vertical inclusions coming from (11.4):

$$
\begin{array}{ccc}
U \otimes O(G) & \hookrightarrow & V \otimes O(G) \\
\downarrow & & \downarrow \\
U \otimes \mathbb{k}G & \hookrightarrow & V \otimes \mathbb{k}G
\end{array}
$$

Then $\delta u \in V \otimes O(G)$ and also $\delta u \in U \otimes \mathbb{k}G$, because $U$ is stable under the $G$-action on $V$ and so $u_0(g, u_1) \in U$ for all $g \in G$. Therefore, $\delta u \in (V \otimes O(G)) \cap (U \otimes \mathbb{k}G) = U \otimes O(G)$ as desired.

Next, let $\pi: V \to \overline{V}$ be an epimorphism in $\text{Rep} \mathbb{k}G$. We have just seen that $\text{Ker} \pi$ is a $O(G)$-subcomodule of $V$. Therefore, the $O(G)$-coaction on $V$ passes down to a coaction $\overline{V} \to \overline{V} \otimes O(G)$, $\pi(v) \mapsto \pi(v_0) \otimes v_1$. This coaction produces the given $G$-action on $\overline{V}$, whence $\overline{V}$ is rational:

$$g \cdot \pi(v) = \pi(g \cdot v) = \pi(v_0)(g, v_1).$$
As for tensor products $V \otimes W$ with $V, W \in \text{Rep} \mathbb{k}G$ both rational, observe that the $G$-action on $V \otimes W$ comes from the standard $O(G)$-coaction (10.5) on $V \otimes W$:

$$g \cdot (v \otimes w) = g.v \otimes g.w \overset{(11.5)}{=} v_{(0)} \otimes w_{(0)} \langle g, v_{(1)} \rangle \langle g, w_{(1)} \rangle$$

$$= v_{(0)} \otimes w_{(0)} \langle g, v_{(1)} \rangle \langle g, w_{(1)} \rangle.$$

Thus, $V \otimes W$ is rational. Finally, if $V_i \in \text{Rep} \mathbb{k}G$ is a family of rational representations, with $G$-actions coming from the coactions $\delta_i : V_i \to V_i \otimes O(G)$, then the standard action $G \C V := \bigoplus_i V_i$ comes from the coaction

$$\delta = \bigoplus_i \delta_i : \bigoplus_i V_i \to \bigoplus_i V_i \otimes O(G) \cong V \otimes O(G),$$

proving that $V$ is rational.

(c) We have already remarked above that every rational representation $V$ is a locally finite left right $O(G)^\circ$-comodule, and hence it is certainly a locally finite left $\mathbb{k}G$-module. More directly, (11.5) says that the subspace of $V$ that is generated by the $G$-orbit $G.v$ of a given $v \in V$ is contained in the subspace spanned by the various $v_{(0)}$, which is finite dimensional. 

### 11.3.2. Rational Actions on Algebras

Now let us start with a comodule algebra $A \in \text{Alg} O^{O(G)}$. Writing the coaction $\delta : A \to A \otimes O(G)$ as $a \mapsto a_{(0)} \otimes a_{(1)}$, we obtain the action rule

$$(11.6) \quad h.a = a_{(0)} \langle h, a_{(1)} \rangle \quad (h \in O(G)^\circ, a \in A),$$

which makes $A$ into a left $O(G)^\circ$-module algebra (Proposition 10.26). Restriction of (11.6) to $G \cong G(O(G)^\circ)$ gives an action $G \C A$ by $\mathbb{k}$-algebra automorphisms (Lemma 10.25); in fact, applying Proposition 11.12(a) to the multiplication and unit maps of $A$, one sees that a right $O(G)$-coaction $\delta$ satisfies condition (10.24) defining comodule algebras if and only if the $G$-action $g.a = a_{(0)} \langle g, a_{(1)} \rangle$ is an action by $\mathbb{k}$-algebra automorphisms. Actions of this form are called rational $G$-actions. Thus, an action $G \C A$ is rational if $G$ acts by $\mathbb{k}$-algebra automorphisms and $A$ becomes a rational representation of $G$ through this action.

The following criterion is often useful.

**Lemma 11.13.** Assume that $G$ acts by automorphisms on $A \in \text{Alg} \mathbb{k}G$ and that $A$ is generated, as $\mathbb{k}$-algebra, by a $G$-stable subspace $V \subseteq A$ that is a rational representation of $G$. Then the action $G \C A$ is rational.

**Proof.** By the foregoing, it suffices to show that the given $G$-action makes $A$ a rational representation of $G$. But our hypotheses imply that $A$ is a homomorphic image of $TV = \bigoplus_{n \geq 0} V \otimes^n$ in $\text{Rep} \mathbb{k}G$ and Proposition 11.12 tells us that $TV$ and all its images in $\text{Rep} \mathbb{k}G$ are rational. \[\square\]
Example 11.14 (G-action on O(G)). The Hopf algebra O(G) carries the regular right O(G)-comodule algebra structure, given by the comultiplication of O(G). Hence O(G) is a left O(G)*-module algebra; the action (11.6) for this setting was stated earlier as (10.26). Restricting this action to G as above, we obtain the following rational action G ⊗ O(G):

\[ g.h = h_{(1)} \langle g, h_{(2)} \rangle \quad (g \in G, h \in O(G)). \]

Example 11.15 (Torus actions and \( \mathbb{Z}^n \)-graded algebras). Consider the group algebra \( kL \) of the lattice \( L = \mathbb{Z}^n \). The category \( \text{Alg}_{kL} \) of right \( H \)-comodule algebras is equivalent to the category \( \text{Alg}_{k} \) of all \( L \)-graded \( k \)-algebras (10.27) and the affine algebraic group associated to \( kL \) is the algebraic \( n \)-torus \( G = (k^+)^n \); so \( O(G) \equiv kL \) (Example 11.10). For any \( L \)-graded algebra \( A = \bigoplus_{\lambda \in L} A_\lambda \), we thus have a rational torus action \( G \otimes A \) and every rational torus action arises in this way. Explicitly, viewing \( \lambda = (z_1, \ldots, z_n) \in L \) as the \( k \)-valued function on \( G \) that is given by \( \lambda(g) = \prod_i \gamma_i^{z_i} \) for \( g = (\gamma_1, \ldots, \gamma_n) \in G \) as in Example 11.10, the action \( G \otimes A \) is given by

\[ g.a = \sum_{\lambda} \lambda(g) a_\lambda \quad (g \in G, a = \sum_{\lambda} a_\lambda \in A). \]

Thus, \( A_\lambda = \{ a \in A \mid g.a = \lambda(g)a \} \) is the \( \lambda \)-weight space of \( A \).

Exercises for Section 11.3

11.3.1 (Finite groups). Let \( G \) be a finite group, viewed as an affine algebraic \( k \)-group (Example 11.9). Show that every \( V \in \text{Rep}_k G \) is rational. Consequently, every action \( G \otimes A \) by automorphism on \( A \in \text{Alg}_k \) is rational.

11.3.2 (Rational representations and homomorphisms). Let \( \varphi: G \to D \) be a map of affine algebraic \( k \)-groups and let \( V \in \text{Rep}_k D \) be rational. Show that \( \varphi^*V \in \text{Rep}_k G \) is also rational.

11.3.3 (Additive and multiplicative group). Show that the action (11.7) of \( G_a(k) = (k,+ \) on \( k[x] \) is given by \( \gamma.x = x + \gamma (\gamma \in k) \). For \( G_m(k) = k^* \subset k[x^{\pm 1}] \), show that (11.7) takes the form \( \gamma.x = \gamma x (\gamma \in k^*) \).

11.4. Linearity

This section gives another description of affine algebraic groups, which does not explicitly mention Hopf algebras or group schemes. It turns out that all affine algebraic \( k \)-groups “come from” the general linear groups \( GL_n(k) \); see Theorem 11.17 below. Therefore, affine algebraic groups are often called linear algebraic groups in the literature. The realization of affine algebraic groups as subgroups of general linear groups allows for a quick and explicit construction, comparable to the Cayley embedding of finite groups into symmetric groups. However, the functorial approach via Hopf algebras taken in the previous sections has formal advantages.
11.4.1. The Zariski Topology of an Affine Algebraic Group

Let $G$ be an affine algebraic $\mathbb{k}$-group. Viewing the Hopf algebra $O(G)$ as an (affine) algebra of functions $G \to \mathbb{k}$ (11.4), we define, for any subset $S \subseteq O(G)$,

$$V(S) \overset{\text{def}}{=} \{ g \in G \mid h(g) = 0 \text{ for all } h \in S \}.$$  

With subsets of this form as the closed sets, one obtains the so-called Zariski topology of $G$; the topology axioms are checked as in Section C.3. All topological statements below refer to this topology. In particular, subgroups of $G$ that are closed for the Zariski topology are simply called closed subgroups of $G$. Note that singletons $\{ g \} \subseteq G$ are closed, being given by $\{ g \} = V(S)$ with $S = \{ h \in O(G) \mid \{ g, h \} = 0 \}$ a maximal ideal of $O(G)$, and that the Zariski topology on direct products is not the usual product topology.

**Proposition 11.16.**

(a) Homomorphism of affine algebraic $\mathbb{k}$-groups are continuous. Similarly for the multiplication $G \times G \to G$ of an affine algebraic group $G$ as well as the maps $G \to G$ that are given by right or left multiplication with a given element of $G$ or inversion.

(b) Let $G$ be an affine algebraic $\mathbb{k}$-group and let $D \leq G$ be an arbitrary subgroup. Then the closure $\overline{D}$ is also a subgroup: it is an affine algebraic $\mathbb{k}$-group and the inclusion $\overline{D} \hookrightarrow G$ is a homomorphism of algebraic groups that is given by a surjective Hopf algebra map $O(G) \twoheadrightarrow O(D)$.

**Proof.** (a) Let $f : G \to E$ be a homomorphism of affine algebraic $\mathbb{k}$-groups and let $f^* = O f : O(E) \to O(G)$, $h \mapsto h \circ f$. For any closed subset $V(S) \subseteq E$, the preimage under $f$ is closed in $G$:

$$f^{-1}(V(S)) = \{ g \in G \mid (h \circ f)(g) = 0 \text{ for all } h \in S \} = V(f^*(S)).$$

This proves continuity of $f$. Identifying $O(G \times G)$ with $O(G) \otimes O(G)$ (Example 11.11), the multiplication map $\mu : G \times G \to G$ comes from the comultiplication $\Delta : O(G) \to O(G) \otimes O(G) = O(G \times G)$, $h \mapsto h \circ \mu = h^{(1)} \otimes h^{(2)}$. As above, this implies continuity of $\mu$. For a given $x \in G$, the map $\lambda_x : G \to G$, $g \mapsto xg$, comes from the map $O(G) \to O(G)$, $h \mapsto h \circ \lambda_x = \langle x, h^{(1)} \rangle \otimes h^{(2)}$; likewise for right multiplication by $x$. Finally, inversion of $G$ comes from the antipode $S$ of $O(G)$ (Exercise 11.2.4). All these maps are therefore continuous.

(b) Being a continuous function $G \to G$ that maps $D$ to itself, inversion also maps $\overline{D}$ to itself. Left multiplication by any $x \in G$ is a homeomorphism $G \to G$, having a continuous inverse. Therefore, $x\overline{x} = \overline{x}x$ for any subset $X \subseteq G$. If $x \in D$, then $xD = D$ and we conclude that $\overline{x} \overline{D} \subseteq \overline{D}$. Hence, $D\overline{D} \subseteq \overline{D}$. Now let $x \in \overline{D}$. Then $Dx \subseteq D\overline{D} \subseteq \overline{D}$ and so $\overline{D}x \subseteq \overline{D}$. A similar argument as above, for right multiplication by $x$, shows that $\overline{D}x = \overline{D}x$. Therefore, $\overline{D} \subseteq \overline{D}$ and so $D\overline{D} \subseteq \overline{D}$. This shows that $\overline{D}$ is a subgroup of $G$. 


Now let $X \subseteq G$ be any closed subset, say $X = V(S)$ with $S \subseteq O(G)$. Then $I := \{ h \in O(G) \mid h(x) = 0 \text{ for all } x \in X \}$ is a semiprime ideal of $O(G)$; in fact, $I = \sqrt{(S)}$ by the Nullstellensatz (Section C.1). Hence $A := O(G)/I$ is a reduced $\mathbb{k}$-algebra. Furthermore, $S \subseteq I$ and $X \subseteq V(I)$, whence $X \subseteq V(I) \subseteq V(S) = X$ and so $X = V(I)$. The canonical epimorphism $O(G) \to A$ may be regarded as the map that restricts any function $h \in O(G)$ from $G$ to $X$ and, as has been remarked in the proof of (a), the comultiplication $\Delta : O(G) \to O(G) \otimes O(G) = O(G \times G)$ coincides with the map $h \mapsto h \circ \mu$, where $\mu$ is the group multiplication of $G$. Therefore, the inclusion $\mu(X \times X) \subseteq X$ is equivalent to $\Delta I \subseteq I \otimes O(G) \otimes O(G) \otimes I$, the right-hand side being the kernel of the map $O(G) \otimes O(G) \to A \otimes A$ that restricts any function $d \in O(G) \otimes O(G)$ to $G \times G$ from $G \times G$ to $X \times X$. Similarly, the inclusion $X^{-1} \subseteq X$ is equivalent to stability of $I$ under the antipode of $O(G)$ and $1 \in X$ is equivalent to $I \subseteq O(G)^+$, the augmentation ideal of $O(G)$. Therefore, $X$ is a subgroup of $G$ if and only if $I$ is a Hopf ideal of $O(G)$. In this case, the canonical epimorphism $O(G) \twoheadrightarrow A$ is a surjective Hopf algebra map and the resulting embedding $G(A^0) \hookrightarrow G(O(G)^0)$ is a homomorphism of affine algebraic groups. Under the canonical isomorphisms $G(A^0) \cong X$ and $G(O(G)^0) \cong G$, this embedding corresponds to the inclusion $X \hookrightarrow G$. The former isomorphism equips $X$ with the structure of an affine algebraic $\mathbb{k}$-group with $O(X) = A$. \hfill $\Box$

### 11.4.2. Linear Algebraic Groups

A **linear algebraic $\mathbb{k}$-group**, by definition, is a closed subgroup of some $GL_n(\mathbb{k})$. To unfold this definition, recall that $O(GL_n(\mathbb{k})) = \mathbb{k}[X_{ij} \mid 1 \leq i, j \leq n][D^{-1}]$; so each $h \in O(GL_n(\mathbb{k}))$ has the form $h = p/D^r$, with $p \in \mathbb{k}[X_{ij} \mid 1 \leq i, j \leq n]$ and $r \in \mathbb{Z}_+$. As functions $GL_n(\mathbb{k}) \to \mathbb{k}$, the variable $X_{ij}$ assigns the $(i, j)$-entry to each matrix $g \in GL_n(\mathbb{k})$ and $h(g) = p(g)/(\det g)^r$. Since $h(g) = 0$ if and only if $p(g) = 0$, the Zariski topology of $GL_n(\mathbb{k})$ is induced by the Zariski topology of $\operatorname{Mat}_n(\mathbb{k})$ (Section C.3) and linear algebraic $\mathbb{k}$-groups are the subgroups $G \leq GL_n(\mathbb{k})$ having following form, for some subset $S \subseteq O(\operatorname{Mat}_n) = \mathbb{k}[X_{ij} \mid 1 \leq i, j \leq n]$:

$$G = \{ g \in GL_n(\mathbb{k}) \mid h(g) = 0 \text{ for all } h \in S \}.$$ 

Each choice of $S$ gives an explicit linear algebraic group and hence an affine algebraic $\mathbb{k}$-group (Proposition 11.16). For example, $GL_n(\mathbb{k})$ results from $S = \emptyset$ and $SL_n(\mathbb{k})$ from $S = \{ D - 1 \}$, giving the groups in Example 11.8. Here are some further examples of linear algebraic $\mathbb{k}$-groups, some of which have been considered earlier (e.g., Exercise 11.2.3):

- the choice $S = \{ X_{ij} \mid i \neq j \}$ yields the group $D_n(\mathbb{k})$ of all diagonal matrices in $GL_n(\mathbb{k})$, which is clearly isomorphic to $(\mathbb{k}^\times)^n$, the algebraic $n$-torus;
- the group $T_n(\mathbb{k})$ consisting of all upper triangular matrices in $GL_n(\mathbb{k})$ comes from $S = \{ X_{ij} \mid i > j \}$; and
• $S = \{X_{ij} \mid i > j\} \cup \{X_{ii} - 1 \mid \text{all } i\}$ produces the group $U_n(\mathbb{k})$ of all unipotent upper triangular matrices in $GL_n(\mathbb{k})$.

Numerous other examples can be constructed with similar ease. Nonetheless, the following theorem may seem surprising at first; in conjunction with Proposition 11.16, it states that affine and linear algebraic groups are the same up to isomorphism.

**Theorem 11.17.** Every affine algebraic $\mathbb{k}$-group is isomorphic to a linear algebraic $\mathbb{k}$-group.

**Proof.** Let $G$ be an affine algebraic $\mathbb{k}$-group. By Example 11.14, there is a rational action $G \subseteq O(G)$ coming from the comultiplication of $O(G)$. Choose a finite set of generators of the algebra $O(G)$. By Proposition 11.12(c), we may assume that the $\mathbb{k}$-subspace $V \subseteq O(G)$ that is spanned by these generators is $G$-stable and, by part (b) of the same proposition, $V$ is a rational subrepresentation of $O(G)$, that is, a right coideal. Thus, fixing a $\mathbb{k}$-basis $h_1, h_2, \ldots, h_n$ of $V$, there are unique $h_{ij} \in O(G)$ such that $\Delta h_j = \sum_i h_i \otimes h_{ij}$.

It follows that the matrix $(h_{ij}) \in \text{Mat}_n(O(G))$ is invertible, and the counit and coassociative laws give $h_j = \sum_i (e, h_i) h_{ij}$, and $\Delta h_{ij} = \sum_k h_{ik} \otimes h_{kj}$. In particular, the elements $h_{ij}$ also generate $O(G)$. Furthermore, formula (11.7) now takes the form $g.h_j = \sum_i h_i (g, h_{ij})$ for $g \in G$, giving a group homomorphism $f : G \to GL_n(\mathbb{k}), \ g \mapsto (\langle g, h_{ij} \rangle)_{i,j}$.

In fact, $f$ is a homomorphism of affine algebraic $\mathbb{k}$-groups: the corresponding Hopf algebra map $f^* : O(GL_n(\mathbb{k})) \to O(G)$ is given by $X_{ij} \mapsto h_{ij}$, $D^{-1} \mapsto \det(h_{ij})^{-1}$.

This map is surjective, because the $h_{ij}$ generate $O(G)$, proving that $G$ is isomorphic to a closed subgroup of $GL_n(\mathbb{k})$ (Proposition 11.16).

**11.4.3. The Closure of a Linear Group**

The machinery of algebraic groups is oftentimes useful when analyzing a finite-dimensional representation $G \to GL(V) \cong GL_n(\mathbb{k})$ of an arbitrary group $G$. Indeed, we may replace $G$ by its image in $GL_n(\mathbb{k})$, thereby reducing to the case where $G \leq GL_n(\mathbb{k})$. For many questions, the following proposition allows us to further replace $G$ by its closure in $GL_n(\mathbb{k})$, which is an affine algebraic group (Proposition 11.16).

**Proposition 11.18.** Let $G \leq GL_n(\mathbb{k})$ be an arbitrary subgroup and let $\overline{G} \leq GL_n(\mathbb{k})$ denote its closure. Then any $G$-stable subspace $V \subseteq \mathbb{k}^n$ is in fact stable under $\overline{G}$ and $V$ is a rational representation of the algebraic group $\overline{G}$.

**Proof.** Put $G = GL_n(\mathbb{k})$ and $G_V = \{g \in G \mid g.V = V\}$. The first assertion amounts to proving that $G_V$ is closed in $G$. To this end, observe that $\mathbb{k}^n$ is a rational
Then we may write $\delta : \mathbb{k}^n \to \mathbb{k}^n \otimes O(G)$, $e_j \mapsto \sum_i e_i \otimes X_{ij}$.

Now choose a basis $(v_j)_j$ of $V$ and a basis $(w_k)_k$ for a complement of $V$ in $\mathbb{k}^n$. Then we may write $\delta v_i = \sum_j v_j \otimes h_{j,i} + \sum_k w_k \otimes f_{k,i}$ with $h_{j,i}, f_{k,i} \in O(G)$. Thus, $g.v_i = \sum_j v_j g(h_{j,i}) + \sum_k w_k g(f_{k,i})$ for $g \in G$. Since $g \in G_V$ is equivalent to $g.v_i \in V$ for all $i$, we obtain

$$g \in G_V \iff \langle g, f_{k,i} \rangle = 0 \text{ for all } k, i \iff g \in V(f_{k,i} \mid \text{ all } k, i).$$

This shows that $G_V$ is closed.

As for rationality, we have already remarked that the action $G \subset \mathbb{k}^n$ is rational. Rationality of the restricted action $\overline{G} \subset \mathbb{k}^n$ follows from Exercise 11.3.2. In brief, this action comes from the coaction $\mathbb{k}^n \to \mathbb{k}^n \otimes O(G) \to \mathbb{k}^n \otimes O(\overline{G})$, where the first map is the above coaction $\delta$ and the second map comes from the restriction homomorphism $O(G) \to O(\overline{G})$ (Proposition 11.16). Thus, $\mathbb{k}^n$ is a rational representation of $\overline{G}$ and so the subrepresentation $V$ is rational as well (Proposition 11.12).

**Exercises for Section 11.4**

11.4.1 (Automorphism groups). Let $\mathcal{A}$ be a finite-dimensional $\mathbb{k}$-vector space that is equipped with a $\mathbb{k}$-bilinear “multiplication” map $\mathcal{A} \times \mathcal{A} \to \mathcal{A}$, $(a, b) \mapsto ab$, as in §5.1.5. Show that $\text{Aut} \mathcal{A} = \{ f \in \text{GL}(\mathcal{A}) \mid f(ab) = f(a)f(b) \text{ for all } a, b \in \mathcal{A} \}$ is an affine algebraic $\mathbb{k}$-group and $\mathcal{A}$ is a rational representation of $\text{Aut} \mathcal{A}$.

11.4.2 (Automorphisms of the polynomial algebra). Show:

(a) Restriction to $\mathbb{k}[x] \subset \mathbb{k}[x]$ gives an isomorphism $\text{Aut}_{\text{Alg}}(\mathbb{k}[x]) \xrightarrow{\sim} \left( \begin{array}{cc} 1 & \mathbb{k} \\ 0 & \mathbb{k}^n \end{array} \right)$ in Groups. In this way, we may regard $\text{Aut}_{\text{Alg}}(\mathbb{k}[x])$ as an affine algebraic $\mathbb{k}$-group.

(b) The standard action $\text{Aut}_{\text{Alg}}(\mathbb{k}[x]) \subset \mathbb{k}[x]$ is rational.

11.4.3 (Stabilizers). Let $G$ be an affine algebraic $\mathbb{k}$-group, let $V \in \text{Rep} \mathbb{k}G$ be rational, and let $U \subseteq V$ be an arbitrary $\mathbb{k}$-subspace. Put $G_U = \{ g \in G \mid g.U = U \}$. Show:

(a) If $g \in G$ satisfies $g.U \subseteq U$, then $g \in G_U$.

(b) $G_U$ is a closed subgroup of $G$ (and so $G_U$ is an affine algebraic $\mathbb{k}$-group) and $U$ is a rational representation of $G_U$.

11.4.4 (Normalizers and centralizers). Let $G$ be an affine algebraic $\mathbb{k}$-group.

(a) For any $x \in G$, the map $c_x : G \to G, g \mapsto g x g^{-1}$, is continuous.
(b) If $X \subseteq G$ is any subset, then $C_G(X) = \{ g \in G \mid gx = xg^{-1} \text{ for all } x \in X \}$ is a closed subgroup of $G$; if $X$ is closed, then $N_G(X) = \{ g \in G \mid gXg^{-1} = X \}$ is a closed subgroup.

**11.4.5 (Orbits).** Show, that, for any $v \in \mathbb{k}^n$, the orbit map $\gamma_v: \text{GL}_n(\mathbb{k}) \to \mathbb{k}^n$, $g \mapsto g.v$, is continuous for the Zariski topologies of $\text{GL}_n(\mathbb{k})$ and $\mathbb{k}^n$.

### 11.5. Irreducibility and Connectedness

A topological space $X$ is said to be **irreducible** if $X$ cannot be written as the union of two proper closed subsets. This property is generally stronger than being **connected**, that is, $X$ cannot be written as the union of two disjoint proper closed subsets. However, for an affine algebraic group $G$ with the Zariski topology ($\S 11.4.1$), irreducibility and connectedness are in fact equivalent; see Proposition 11.20 below. In order to avoid confusion with the representation theoretic meaning of irreducibility, affine algebraic groups $G$ that are topologically irreducible are usually referred to as “connected,” especially when $G$ is realized as a linear algebraic group. It turns out that every affine algebraic group $G$ is “almost” connected (Proposition 11.20). This fact makes it possible to reduce many questions about affine algebraic groups to the connected case.

#### 11.5.1. Primes and Irreducibility

We momentarily suspend our discussion of algebraic groups for some topological considerations in a slightly more general setting. This subsection assumes familiarity with §1.3.4 and Exercise 1.3.1. The reader may also want to have a look at §5.6.3 and Exercises 5.6.3, 5.6.4.

**Jacobson-Zariski Topology.** Let $A \in \text{Alg}_\mathbb{k}$ be arbitrary. Recall that the Jacobson-Zariski topology on the set $\text{Spec} A$ of all prime ideals of $A$ is defined by choosing as closed sets the subsets of the form

$$\mathcal{V}(I) \overset{\text{def}}{=} \{ P \in \text{Spec} A \mid P \supseteq I \}.$$  

Here, $I$ may be any subset of $A$, but without loss of generality, we may assume that $I$ is a semiprime ideal of $A$. Conversely, any subset $X \subseteq \text{Spec} A$ yields a semiprime ideal of $A$ by putting

$$\mathcal{J}(X) \overset{\text{def}}{=} \bigcap_{P \in X} P.$$  

In this way, we obtain inclusion reversing bijections that are inverse to each other:

\[
\begin{array}{c}
\text{closed subsets of Spec } A \overset{\mathcal{V}(\cdot)}{\leftrightarrow} \text{semiprime ideals of } A \\
\text{semiprime ideals of } A \overset{\mathcal{J}(\cdot)}{\leftrightarrow} \text{closed subsets of Spec } A
\end{array}
\]
Under the bijection (11.9), irreducible closed subsets of \( \text{Spec} A \) correspond to prime ideals of \( A \) (Exercise 1.3.1). Since (11.9) reverses inclusions, the maximal ideals of \( A \) correspond to the minimal closed subsets of \( \text{Spec} A \), the 1-point subsets \( \mathcal{V}(P) = \{P\} \) with \( P \in \text{MaxSpec} A \), and the minimal prime ideals of \( A \) correspond to the maximal closed subsets of \( \text{Spec} A \).

**Irreducible Components.** Some facts from general topology (e.g., Bourbaki [22, Chap. II §4, Prop. 2 and 5]): the closure of any irreducible subset of an arbitrary topological space \( X \) is also irreducible; 1-point subsets of \( X \) are evidently irreducible; and every irreducible subset of \( X \) is contained in a maximal one by Zorn’s Lemma. Thus, the maximal irreducible subsets of \( X \) are automatically closed and \( X \) is their union; these subsets are called the **irreducible components** of \( X \). By (11.9), the irreducible components of \( \text{Spec} A \) are the sets \( \mathcal{V}(P) \), where \( P \) is a minimal prime ideal of \( A \). Any right or left noetherian algebra \( A \) has only finitely many minimal primes (Exercise 1.3.2), and hence \( \text{Spec} A \) has finitely many irreducible components.

**Zariski Topology.** Now assume that \( A \in \text{Alg}_k \) is **affine commutative** and consider the subset \( \text{MaxSpec} A \subseteq \text{Spec} A \) consisting of all maximal ideals of \( A \) with the topology that is induced by the Jacobson-Zariski topology of \( \text{Spec} A \). Since the base field \( k \) is assumed algebraically closed, the weak Nullstellensatz (Section C.1) yields a bijection

\[
\Gamma_A(k) \overset{\text{def}}{=} \text{Hom}_{\text{Alg}_k}(A, k) \overset{\sim}{\longrightarrow} \text{MaxSpec} A
\]

\[
g \overset{\sim}{\longmapsto} \ker g
\]

Transporting the topology of \( \text{MaxSpec} A \) to \( \Gamma_A(k) \) by means of this bijection, we obtain a topology whose closed sets are those of the form

\[
\mathcal{V}(I) \overset{\text{def}}{=} \left\{ g \in \Gamma_A(k) \mid \langle g, h \rangle = 0 \text{ for all } h \in I \right\}.
\]

Here, \( I \) is a subset of \( A \), without loss a semiprime ideal. This topology is called the **Zariski topology** of \( \Gamma_A(k) \). For an any affine algebraic group \( G \cong \Gamma_A(k) \) with \( A = O(G) \), this is exactly the topology considered in §11.4.1. In general, for any subset \( X \subseteq \Gamma_A(k) \), we define

\[
I(X) \overset{\text{def}}{=} \bigcap_{g \in X} \ker g.
\]
By the Nullstellensatz (particularly, the Jacobson property) the bijections (11.9) yield the following are inclusion reversing inverse bijections:

\[
\begin{align*}
\{ \text{closed subsets of } \Gamma_A(\mathbb{k}) \} & \xleftrightarrow{\Phi(\cdot)} \{ \text{semiprime ideals of } A \} \\
\end{align*}
\]

Once again, irreducible closed subsets of \( \Gamma_A(\mathbb{k}) \) correspond to prime ideals of \( A \) under this bijection, and the irreducible components of \( \Gamma_A(\mathbb{k}) \) correspond to the minimal prime ideals of \( A \). Thus, \( \Gamma_A(\mathbb{k}) \) has finitely many irreducible components.

### 11.5.2. Products

For any two affine commutative \( \mathbb{k} \)-algebras \( A \) and \( B \), there is a canonical bijection of sets,

\[
\Gamma_A(\mathbb{k}) \times \Gamma_B(\mathbb{k}) \equiv \Gamma_{A \otimes \mathbb{k}}(\mathbb{k}).
\]

This bijection was already considered in the setting of algebraic groups in Example 11.11. As for algebraic groups, the Zariski topology of \( \Gamma_{A \otimes \mathbb{k}}(\mathbb{k}) \) transfers to make the cartesian product \( \Gamma_A(\mathbb{k}) \times \Gamma_B(\mathbb{k}) \) a topological space. If \( X \subseteq \Gamma_A(\mathbb{k}) \) and \( Y \subseteq \Gamma_B(\mathbb{k}) \) are irreducible closed subsets, then \( X \times Y \) is an irreducible closed subset of \( \Gamma_A(\mathbb{k}) \times \Gamma_B(\mathbb{k}) \). To see this, write \( X = V(P) \equiv \Gamma_A(P)(\mathbb{k}) \) and \( Y = V(Q) \equiv \Gamma_B(Q)(\mathbb{k}) \) with \( P \in \text{Spec } A \) and \( Q \in \text{Spec } B \). Then \( X \times Y \equiv \Gamma_{A/P \otimes B/Q}(\mathbb{k}) \) and so irreducibility of \( X \times Y \) amounts to primeness of the algebra \( A/P \otimes B/Q \), which is guaranteed by the following ring theoretic lemma. See Exercise 11.5.1 for an even more general fact.

**Lemma 11.19.** Let \( C \in \text{CommAlg}_\mathbb{k} \) be a domain. If \( A \in \text{Alg}_\mathbb{k} \) is prime (semiprime, a domain) then so is \( A \otimes C \).

**Proof.** Put \( T = A \otimes C \). The three desired conclusions, namely for \( T \) to be prime, semiprime or a domain, are respectively equivalent to \( rTs, rTt \) or \( rs \) being nonzero for any given \( 0 \neq r, s \in T \). Therefore, we may assume without loss that \( C \) is an affine commutative domain. Consider the topological space \( X := \Gamma_C(\mathbb{k}) \) with the Zariski topology. By hypothesis on \( C \), the space \( X \) is irreducible. For each \( x \in X \) consider the map \( \overline{x} = \text{Id}_A \otimes x : A \otimes C \to A \otimes \mathbb{k} = A \) in \( \text{Alg}_\mathbb{k} \).

First assume that \( A \) is a domain and let \( r, s \in T \) be given such that \( rs = 0 \). Then \( \overline{x}(r)\overline{x}(s) = \overline{x}(rs) = 0 \) for all \( x \in X \). Since \( A \) is a domain, we conclude that, for each \( x \in X \), we must have \( \overline{x}(r) = 0 \) or \( \overline{x}(s) = 0 \). Fix a \( \mathbb{k} \)-basis \( (a_i) \) for \( A \) and write \( r = \sum a_i \otimes r_i \) and \( s = \sum a_i \otimes s_i \) with \( r_i, s_i \in C \). Then, for each \( x \in X \), we must have either \( x(r_i) = 0 \) for all \( i \) or \( x(s_i) = 0 \) for all \( i \). In other words, \( X = R \cup S \), where \( R = V(r_i \mid \text{ all } i) \) and \( S = V(s_i \mid \text{ all } i) \) are closed. Since \( X \) is irreducible, it follows that \( R = X \) or \( S = X \). In the former case, all \( r_i = 0 \) and so \( r = 0 \); the latter case similarly yields \( s = 0 \). This shows that \( T \) is a domain.
The case where $A$ is prime can be handled in analogous fashion. Indeed, $rTs = 0$ implies that $\bar{x}(r)A\bar{x}(s) = \bar{x}(rTs) = 0$ for all $x \in X$, and hence $\bar{x}(r) = 0$ or $\bar{x}(s) = 0$ by primeness of $A$. Now proceed as above to conclude that $r = 0$ or $s = 0$. Finally, for the semiprimeness assertion, take $r = s$ in this argument. □

11.5.3. The Identity Component

Every topological space $X$ is the disjoint union of closed connected subsets such that every connected subset of $X$ is contained in one of them; these maximal connected subsets of $X$ are called the connected components of $X$ (e.g., Munkres [154, Sections 3.1, 3.3]). For an affine algebraic group $G$ equipped with the Zariski topology, the proposition below describes the connected component of $G$ containing the identity element; this component is called the identity component of $G$.

**Proposition 11.20.** Let $G$ be an affine algebraic $k$-group and let $G_1$ denote the identity component of $G$. Then $G_1$ is a closed normal subgroup $G$ having finite index in $G$. The cosets of $G_1$ are the connected as well as the irreducible components of $G$. In particular, $G$ is topologically irreducible if and only if $G$ is connected.

**Proof.** Let $C, D$ be irreducible components of $G$ both containing the identity element $1 \in G$. Then $C \times D$ is an irreducible subset of $G \times G$ (Lemma 11.19). Since continuous images of irreducible topological spaces are again irreducible and the multiplication map $\mu: G \times G \to G$ is continuous (Proposition 11.16), it follows that $\mu(C \times D) = CD$ is an irreducible subset of $G$. By maximality of $C$ and $D$, we must have $C = CD = D$. Consequently, there is a unique irreducible component $C \subseteq G$ such that $1 \in C$ and we know that $C$ is closed under multiplication. By Proposition 11.16, inversion is a homeomorphism $G \to G$ fixing 1. Thus, we must have $C^{-1} = C$, which shows that $C$ is a subgroup of $G$.

Left or right multiplication by any $g \in G$ also gives a homeomorphism $G \to G$ (Proposition 11.16), and hence the same holds for conjugation by $g$. Since the latter map fixes 1, we obtain that $gCg^{-1} = C$, proving normality of $C$. Furthermore, the various cosets $gC$ with $g \in G$ are all maximal irreducible subsets of $G$; hence, these cosets are irreducible components of $G$. Since there are only finitely many irreducible components, because $O(G)$ is affine, we obtain that $C$ has finite index in $G$. Thus, $G$ is the finite disjoint union of the distinct cosets $gC$, all of which are closed irreducible subsets of $G$. Any connected subset of $G$ must be contained in some $gC$, and hence these cosets are the connected and the irreducible components of $G$. This proves the proposition. □

**Algebra Structure of $O(G)$.** When $G$ is finite, then $G_1 = 1$ and we have the familiar algebra isomorphism $O(G) = (kG)^{\ast} \cong k^{\times |G|}$ (Example 9.17). A similar direct product decomposition of $O(G)$ occurs in general. In detail, let $p_1 = I(G_1) = \{ h \in O(G) \mid h(g) = 0 \text{ for all } g \in G_1 \}$ denote the ideal of $O(G)$ corresponding in (11.10) to the identity component $G_1$. Thus, $p_1$ is a Hopf ideal of $O(G)$, a minimal
prime ideal, and \(O(G) \cong O(G)/p_1\). By Proposition 11.20, the other irreducible components of \(G\) all have the form \(xG_1\) with \(x \in G\). Each of them corresponds to a minimal prime ideal \(p_x\) (though not a Hopf ideal). Specifically, left translation by \(x\) is a homeomorphism \(\lambda_x : G \to G, g \mapsto xg\), and the map \(\lambda^*_x : O(G) \to O(G)\) is given by \(h \mapsto h \circ \lambda_x = \langle x, h_1(1) \rangle \otimes h_2\); see the proof of Proposition 11.16. This is an algebra automorphism of \(O(G)\) (but not a Hopf map), and

\[
p_x = I(xG_1) = \{ h \in O(G) \mid (h \circ \lambda_x)(g) = 0 \text{ for all } g \in G_1 \} = (\lambda^*_x)^{-1}p_1.
\]

Since distinct components \(xG_1 \neq yG_1\) are disjoint, we must have \(p_x + p_y = O(G)\). Therefore, the Chinese Remainder Theorem gives the isomorphism

\[
O(G) \cong \prod_{x \in G/G_1} O(G)/p_x \cong O(G_1)^{\times |G/G_1|}.
\]

**Orthogonal Groups.** With the exception of finite groups, all affine algebraic groups that we have encountered thus far are connected (Exercise 11.5.3). To exhibit another non-connected example, let us assume that \(\text{char} \mathbb{k} \neq 2\). With \(T\) denoting the matrix transpose and \(1 = 1_{n \times n} \in \text{Mat}_n(\mathbb{k})\), we define

\[
O_n(\mathbb{k}) = \{ g \in \text{GL}_n(\mathbb{k}) \mid g^T g = 1 \} \quad \text{and} \quad \text{SO}_n(\mathbb{k}) = O_n(\mathbb{k}) \cap \text{SL}_n(\mathbb{k}).
\]

**Proposition 11.21.** \(O_n(\mathbb{k})\) and \(\text{SO}_n(\mathbb{k})\) are closed subgroups of \(\text{GL}_n(\mathbb{k})\). Furthermore, \([O_n(\mathbb{k}) : \text{SO}_n(\mathbb{k})] = 2\) and \(\text{SO}_n(\mathbb{k})\) is the identity component of \(O_n(\mathbb{k})\).

**Proof.** Let us drop \(\mathbb{k}\) from our notation. It is clear that \(O_n\) and \(\text{SO}_n\) are subgroups of \(\text{GL}_n\). The condition \(g^T g = 1\) can be expressed by the vanishing of certain polynomials in the coordinate functions \(X_{ij}\) of \(\text{Mat}_n\). Therefore, \(O_n\) is closed in \(\text{GL}_n\). Recall also from Example 11.8 that the determinant \(\det : \text{GL}_n \to \text{GL}_1\) is a homomorphism of affine algebraic groups. Hence, its restriction to \(O_n\) is a continuous map \(\det : O_n \to \text{GL}_1\) (Proposition 11.16). For \(g \in O_n\), we have \(1 = \det(g^T g) = (\det g)^2\) and so \(\det g = \pm 1\). The preimages of \(1\) and \(\pm 1\) in \(O_n(\mathbb{k})\) give a partition into two nonempty closed subsets, the former being \(\text{SO}_n\). Thus, \(\text{SO}_n\) is closed and has index \(2\) in \(O_n\). To prove the assertion about the identity component, it remains to show that \(\text{SO}_n\) is irreducible (connected).

Put \(M = \{ x \in \text{Mat}_n \mid \det(1 + x) \neq 0 \} \) and \(g = \{ y \in \text{Mat}_n \mid y^T = -y \} \); this is a Lie subalgebra of \(s\text{li}_n\). Observe that the map \(x \mapsto r(x) = (1 - x)(1 + x)^{-1}\) sends \(M\) to itself, because \(1 + r(x) = 2(x + 1)^{-1}\) has nonzero determinant. The map \(r\) is inverse to itself and continuous for the Zariski topology on \(M\); so it gives a homeomorphism \(M \xrightarrow{\sim} M\). Furthermore, \(r\) sends \(U := M \cap O_n\) to \(V := M \cap g\). For, if \(y = (1 - x)(1 + x)^{-1}\), then

\[
y^T = (1 - x^T)(1 + x)^{-T}\cdot^{-1} = (1 - x^{-1})(1 + x^{-1})^{-1}
\]

\[= (1 - x^{-1})x x^{-1}(1 + x^{-1})^{-1} = (x - 1)(x - 1)^{-1} = -y.
\]
Similarly, \( r(V) \subseteq U \). Therefore, \( r \) gives a homeomorphism \( V \xrightarrow{\sim} U \). The set \( U \) is contained in \( \text{SO}_n \), because
\[
\det(1 + x) \det x = \det(1 + x)x^T = \det(x^T + 1) = \det(1 + x) \quad (x \in O_n).
\]
In fact, \( U \) is dense in \( \text{SO}_n \) (Exercise 11.5.5). Since \( g \) is evidently irreducible and \( V \) is a nonempty open subset, \( V \) is dense in \( g \) and therefore also irreducible. Consequently, \( U \) is irreducible as well, being homeomorphic to \( V \). Since \( U \) is dense in \( \text{SO}_n \), it follows that \( \text{SO}_n \) is irreducible, as desired. \( \square \)

**Exercises for Section 11.5**

11.5.1 (Primeness of tensor products). Let \( A, B \in \text{Alg}_k \) both be prime. The goal of this exercise is to show that \( A \otimes B \) is prime as well.\(^3\) Familiarity with the symmetric ring of quotients \( QA \) and the extended center \( CA = \mathcal{Z}(QA) \) is assumed; see Appendix E. Consider the central closure \( A' := A(CA) \); this is an algebra with \( A \subseteq A' \subseteq QA \). Moreover, \( A' \) is prime and if \( 0 \neq x, y \in A' \) are such that \( xay = yax \) for all \( a \in A \), then \( x = zy \) for some \( z \in CA \) (Exercise E.3.2). Likewise for \( B \). Show:

- (a) If \( A' \otimes B' \) is prime, then so is \( A \otimes B \). (Use Exercise 1.3.5).
- (b) Every nonzero two-sided ideal of \( A' \otimes B' \) contains a nonzero element of the form \( c(a \otimes b) \) with \( c \in CA \otimes CB \), \( a \in A' \) and \( b \in B' \).
- (c) Conclude from (b) that \( A' \otimes B' \) is prime, and hence so is \( A \otimes B \) by (a).
- (d) Algebraic closedness of \( k \) is necessary for the result of this exercise to hold.

11.5.2 (Connectedness and stability). Let \( G \) be a connected affine algebraic \( k \)-group. Show:

- (a) If \( V \in \text{Rep} \, kG \) be rational and \( U \subseteq V \) is a \( k \)-subspace whose orbit under the action of \( G \) is finite, then \( U \) is in fact \( G \)-stable. (Use Exercise 11.4.3.)
- (b) Every finite normal subgroup of \( G \) is contained in the center \( ZG \). (Use Exercise 11.4.4.)

11.5.3 (Examples). Show that the algebraic groups \( G_m(k), G_n(k), GL_n(k), \text{SL}_n(k), T_n(k), D_n(k), U_n(k) \) are all connected (§11.4.2).

11.5.4 (Monomial matrices). Let \( G \) be the subgroup of \( GL_n(k) \) consisting of all monomial matrices, i.e., matrices with exactly one nonzero entry in each row and column. Show that \( G \) is a closed subgroup of \( GL_n(k) \) having identity component \( D_n(k) \), the group of all diagonal matrices in \( GL_n(k) \), and \( [G : D_n(k)] = n! \).

11.5.5 (For the proof of Proposition 11.21). Show that \( \{ x \in \text{SO}_n(k) \mid \det(1 + x) \neq 0 \} \) is dense in \( \text{SO}_n(k) \).

---

\(^3\)This result is due to G. Bergman [14, Proposition 17.2]. The corresponding statement for semiprime algebras is a consequence, but it fails for domains. In fact, there are even division \( k \)-algebras \( D \) and \( E \), finite over their respective centers, such that \( D \otimes E \) has zero divisors [177].
11.6. The Lie Algebra of an Affine Algebraic Group

In this section, we introduce the Lie algebra that is associated to an affine algebraic group \( G \) and establish some of its most basic properties. As an application, we prove the theorem of Chevalley that gave rise to the terminology “Chevalley property” (§10.2.5). The theorem illustrates how theory of Lie algebra representations, as developed in Part III, can be brought to bear on the analysis of group representations in characteristic 0.

Throughout this section, \( G \) is an affine algebraic \( \k \)-group. We continue to assume that the field \( \k \) is algebraically closed in §11.6.1 while §11.6.5 works over an arbitrary field \( \k \) with \( \text{char} \, \k = 0 \).

11.6.1. Definition and Finiteness

Recall that \( G \) arises as the group of group-like elements of the Hopf \( \k \)-algebra \( O(G) \). The Lie algebra of \( G \) can be defined as the Lie algebra of primitive elements (§9.3.2) of the same Hopf algebra:

\[
\text{Lie } G \defeq \text{L}(O(G)\circ) = \{ x \in O(G)\circ \mid \Delta x = x \otimes 1 + 1 \otimes x \}
\]

This yields a covariant functor,

\[
\text{Lie}: \text{AffineAlgebraicGroups}_\k \longrightarrow \text{Lie}_\k,
\]

the composite of the contravariant functors \( O \) and \( \cdot \circ \) with the covariant primitive element functor \( \text{L} \). If \( \text{char} \, \k = p > 0 \), then \( \text{Lie } G \) is a Lie \( p \)-algebra.

With \( m_1 = O(G)^+ \) denoting the augmentation ideal of \( O(G) \), Proposition 9.19 yields the following description of \( \text{Lie } G \):

\[
(11.12) \quad \text{Lie } G = \text{Der}(O(G), \mathbb{1}) \cong (m_1/m_1^2)^* \quad \text{in } \text{Vect}_\k.
\]

The isomorphism here is in \( \text{Vect}_\k \). Since \( O(G) \) is noetherian, the ideal \( m_1 \) is finitely generated and so \( \text{Lie } G \) is finite dimensional (Proposition 9.19).

11.6.2. Reduction to the Identity Component

It follows from (11.12) that the Lie algebra of both \( G_m(\k) \) and \( G_a(\k) \) is the 1-dimensional Lie algebra over \( \k \), despite the fact that these algebraic groups are not isomorphic (Exercise 11.2.2). Also, \( \text{Lie } G = 0 \) for every finite group \( G \), because \( m_1 = m_1^2 \) in this case. In fact, the following proposition shows that, in general, \( \text{Lie } G \) only “sees” the identity component of \( G \). Therefore, \( \text{Lie } G \) is most useful for connected \( G \).
11.6. The Lie Algebra of an Affine Algebraic Group

Proposition 11.22. If \( H \leq G \) is a closed subgroup, then there is an embedding \( \text{Lie} \, H \hookrightarrow \text{Lie} \, G \) in \( \text{Lie}_k \). For \( H = G_1 \), the identity component of \( G \), this embedding is an isomorphism.

Proof. Since the inclusion \( H \hookrightarrow G \) is a map of affine algebraic groups, it gives rise to a Lie algebra map \( \text{Lie} \, H \to \text{Lie} \, G \) by functoriality. More specifically, the inclusion corresponds to a surjective Hopf algebra map \( O(G) \to O(H) \) (Proposition 11.22), which in turn yields an “inflation” embedding \( O(H)^\circ \hookrightarrow O(G)^\circ \) and therefore an embedding \( \text{Lie} \, H = L(O(G)^\circ) \hookrightarrow L(O(G)^\circ) = \text{Lie} \, G \).

Applying this to \( H = G_1 \), it suffices to show that \( \text{Lie} \, G_1 \) and \( \text{Lie} \, G \) have the same dimension. By (11.12), \( \text{Lie} \, G \cong (m_1/m_1^2)^* \) and \( \text{Lie} \, G_1 \cong (m_1/m_1^2)^* \), where \( m_1 = \mathfrak{I}(\{1\}) \) is the augmentation ideal of \( O(G) \) and \( \gamma : O(G) \to O(G_1) \). The kernel of this map is \( \mathfrak{p}_1 = \mathfrak{I}(G_1) \), a minimal prime ideal of \( O(G) \) that is contained in \( m_1 \). In fact, since \( \mathfrak{p}_1 = \mathfrak{p}_1^2 \) by (11.11), we also have \( \mathfrak{p}_1 \cong m_1^2 \) and so \( m_1/m_1^2 \cong m_1^2/m_1^3 \). This proves the proposition. \( \square \)

11.6.3. Calculation with Dual Numbers

The algebra \( k[t]/(t^2) \) is called the algebra of dual numbers. Let \( \tau \) denote the residue class of the variable \( t \), we may write this algebra as \( k[\tau] = k \oplus k\tau \) with \( \tau^2 = 0 \). Sending \( \tau \mapsto 0 \) gives a surjection \( d : k[\tau] \to k \) of commutative \( k \)-algebras.

Proposition 11.23. Let \( \Gamma \cong \Gamma_H \) be an affine group scheme over \( k \), where \( H \) is a commutative Hopf \( k \)-algebra. Then there is a split exact sequence in Groups,

\[
1 \to 1 + \text{Der}(H, \mathfrak{I}) \to \Gamma(k[\tau]) \xrightarrow{\Gamma(d)} \Gamma(k) \to 1.
\]

The Lie bracket in \( \text{Der}(H, \mathfrak{I}) \) comes from the commutator in \( \Gamma(k[\tau] \otimes k[\tau]) \).

Proof. We may assume that \( \Gamma = \Gamma_H \). Every \( f \in \Gamma(k[\tau]) = \text{Hom}_{\text{Algebras}}(H, k[\tau]) \) can be written as \( f = f_0 + f_1 \tau \) with unique \( f_0, f_1 \in H^\circ \). In fact, \( f_0 \in \Gamma(k) = \text{Hom}_{\text{Algebras}}(H, k) \) and \( \Gamma(d)f = f_0 \). The component \( f_1 \) satisfies the identity

\[
\langle f_1, xy \rangle = \langle f_0, x \rangle \langle f_1, y \rangle + \langle f_1, x \rangle \langle f_0, y \rangle \quad (x, y \in H).
\]

The kernel of \( \Gamma(d) \) consists of those \( f \in \Gamma(k[\tau]) \) with \( f_0 = \varepsilon \), the counit of \( H \) and identity element of \( \Gamma(k) \), and then the above identity reduces to the identity (9.36) defining \( \text{Der}(H, \mathfrak{I}) \). This gives the exact sequence, which is split because \( d \) is split by the unit of \( k[\tau] \).

To explain the assertion about the Lie bracket, let us write \( k[\tau] \otimes k[\tau] = k[\tau, \tau'] \) with \( \tau \tau' - \tau' \tau = \tau^2 = \tau^2 = 0 \). For given elements \( v, w \in \text{Der}(H, \mathfrak{I}) \), consider the pair \( (1 + \tau v, 1 + \tau w) \in \Gamma(k[\tau]) \times \Gamma(k[\tau']) \cong \Gamma(k[\tau, \tau']) \). Since \( (1 + \tau v)^{-1} = 1 - \tau v \) in \( \Gamma(k[\tau]) \) and similarly for \( 1 + \tau w \), the commutator of this pair is is given by

\[
(1 + \tau v)(1 + \tau w')(1 - \tau v)(1 - \tau v') = 1 + (vw - wv)\tau \tau',
\]

where \( vw - wv = [v, w] \in \text{L}(H^\circ) \). \( \square \)
With $H = O(G)$, the above proposition yields a very convenient recipe for calculating Lie $G$.

**Example 11.24** (The general linear group). The general linear group $G = GL_n(\mathbb{k})$ arises from the group scheme $\Gamma = GL_n : \text{CommAlg}_\mathbb{k} \to \text{Groups}$ that is given by $GL_n(R) = \text{Mat}_n(R)^\times$. The kernel of the map $\Gamma(d) : \text{Mat}_n(\mathbb{k}[\tau])^\times \to \text{Mat}_n(\mathbb{k})^\times$ consists of all matrices of the form $1 + \tau v$ with $v \in \text{Mat}_n(\mathbb{k})$ and $1 = \mathbf{1}_{n \times n}$. Thus, Proposition 11.23 yields that Lie $G = \text{Der}(\mathbb{O}(G)), \mathbb{1}) = \text{Mat}_n(\mathbb{k})$ with the usual Lie commutator of matrices as the Lie bracket. So

$$\text{Lie } GL_n(\mathbb{k}) = gl_n(\mathbb{k}).$$

**Example 11.25** (The special linear group). Now the underlying group scheme is given by $\text{SL}_n(R) = \{ M \in \text{Mat}_n(R) \mid \det M = 1 \}$. The kernel of $\Gamma(d)$ consists of all matrices of the form $1 + \tau v$ with $v \in \text{Mat}_n(\mathbb{k})$ such that $\det(1 + \tau v) = 1$. But, with $\lambda_i$ denoting the eigenvalues of $v$, we have $\det(1 + \tau v) = \prod_i (1 + \lambda_i \tau) = 1 + (\sum_i \lambda_i) \tau = 1 + \text{trace}(v) \tau$. Thus, the determinant requirement amounts to the condition $\text{trace}(v) = 0$. Proposition 11.23 therefore gives

$$\text{Lie } \text{SL}_n(\mathbb{k}) = sl_n(\mathbb{k}).$$

**Example 11.26** (The orthogonal group). For the group scheme that is given by $O_n(R) = \{ M \in \text{Mat}_n(R) \mid \det M^T M = 1 \}$, the kernel of $\Gamma(d)$ consists of all matrices of the form $1 + \tau v$ with $v \in \text{Mat}_n(\mathbb{k})$ such that $(1 + \tau v)^T (1 + \tau v) = 1$ or, equivalently, $v^T + v = 0$. Thus, Proposition 11.23 now gives the following Lie algebra, which already played a role in the proof of Proposition 11.21,

$$\text{Lie } O_n(\mathbb{k}) = \{ v \in \text{Mat}_n(\mathbb{k}) \mid v^T = -v \}.$$  

### 11.6.4. Invariant Subspaces

There is much to be said about the manifold connections between $G$ and Lie $G$, a topic that we will have to skirt here, because it would take us too far afield. We merely state, without proof, one particular fact about the relationship between $G$ and Lie $G$ that will be needed in the proof of Chevalley’s Theorem.

Let $V$ be a rational representation of $G$; so $V$ is equipped with a right $O(G)$-comodule structure (§11.3.1). This makes $V$ a locally finite left module over the algebra $O(G)^\ast$, with $O(G)^\ast$-action given by (9.24):

$$f.v = v_{(0)} \langle f, v_{(1)} \rangle \quad (f \in O(G)^\ast, v \in V). \tag{11.13}$$

By restriction along the embedding Lie $G \hookrightarrow O(G)^\circ \hookrightarrow O(G)^\ast$, the space $V$ becomes a (locally finite) representation of the Lie algebra Lie $G$. The proof of the following proposition may be found in Hochschild [100, Corollary IV.3.2].

**Proposition 11.27.** Let $G$ be a connected affine algebraic $\mathbb{k}$-group, with char $\mathbb{k} = 0$, and let $V \in \text{Rep}_{\mathbb{k}G}$ be rational. Then the $G$-stable subspaces of $V$ coincide with the subspaces that are stable under Lie $G$. 

11.6.5. Complete Reducibility of Tensor Products

First, we deal with Lie algebras. The base field \( k \) need not be algebraically closed at the outset, since a field extension argument reduces the problem at hand to the algebraically closed case (Exercise 1.4.7).

**Theorem 11.28.** Let \( g \in \text{Lie}_k \) be arbitrary and assume that \( \text{char} k = 0 \). Then the tensor product of any two finite-dimensional completely reducible representations of \( g \) is again completely reducible.

**Proof.** Let \( V, W \in \text{Rep}_{\text{fin}} g \) be completely reducible. In order to show that \( V \otimes W \) is completely reducible, it suffices to consider the case where \( V \) and \( W \) are both irreducible. As in the proof of Proposition 6.7, we may also assume that \( k \) is algebraically closed. Replacing \( g \) by its image in \( \mathfrak{gl}(V \oplus W) \), we may further assume that \( g \) is reductive (Proposition 6.7); so \( g = Z_g \oplus g' \) with \( g' \) semisimple (6.9). By Schur’s Lemma, \( Z_g \) acts on both \( V \) and \( W \) by scalars, say via \( \lambda, \mu \in (Z_g)\hat{} \), respectively. Thus, \( Z_g \) acts on \( V \otimes W \) via \( \lambda + \mu \), and hence it suffices to show that \( V \otimes W \) is completely reducible for \( g' \). This is guaranteed to be the case by Weyl’s Theorem (Section 6.2). \( \square \)

We are now ready to present the theorem of Chevalley [43, Proposition 2 in Chap. IV §5]. The theorem fails in positive characteristics; see Example 10.19.

**Chevalley’s Theorem.** The tensor product of any two finite-dimensional completely reducible representations of an arbitrary group \( G \) over a field \( k \) with \( \text{char} k = 0 \) is again completely reducible.

**Proof.** Let \( V, W \in \text{Rep}_{\text{fin}} kG \) be completely reducible. In order to show that \( V \otimes W \) is a completely reducible representation of \( kG \), we make a number of reductions. First, we may once again assume that \( k \) is algebraically closed. We may also clearly replace \( G \) by its image in \( \text{GL}(V \oplus W) \), and Proposition 11.18 allows us to further replace \( G \) by its Zariski closure \( \overline{G} = G \) in \( \text{GL}(V \oplus W) \), thereby reducing to the case where \( G \) is an affine algebraic \( k \)-group and \( V, W \) are completely reducible rational representations of \( G \). Since \( \text{char} k = 0 \), the relative Maschke Theorem (Exercise 3.4.3) tells us that representations of \( G \) over \( k \) are completely reducible if and only if this holds for their restriction to the identity component \( G_1 \) (§11.5.3). Therefore, there is no loss in assuming that \( G \) is connected.

In this setting, rational representations of \( G \) are completely reducible for \( G \) if and only if this holds for \( \text{Lie} G \) (Proposition 11.27). Therefore, \( V \) and \( W \) are both completely reducible for \( \text{Lie} G \), and hence so is \( V \otimes W \) (Theorem 11.28). Since \( V \otimes W \) is rational (Proposition 11.12), it follows that \( V \otimes W \) is completely reducible for \( G \) as well, proving the theorem. \( \square \)
Exercises for Section 11.6

11.6.1 (Some Lie algebras). Show that Lie $T_n = t_n$, Lie $U_n = n_n$ and Lie $D_n = d_n$ (omitting reference to $k$).

11.6.2 (Automorphisms and derivations). Using the notation of Exercise 11.4.1 and §5.1.5, show that $\text{Lie}(\text{Aut} A) = \text{Der} A$.

11.6.3 (Adjoint representation). Recall that, for any $H \in \text{HopfAlg}_k$, the adjoint (conjugation) action $GH \triangleleft H$ induces an action $GH \triangleleft LH$ by Lie algebra automorphisms (§10.4.3). In particular, any affine algebraic group $G \cong G(\mathcal{O}(G)\delta)$ acts by Lie algebra automorphisms on $\mathfrak{g} := \text{Lie} G = \mathcal{L}(\mathcal{O}(G)\delta)$ in this way. Let $\text{Ad} x \in \text{Aut} \mathfrak{g}$ denote the automorphism given by the action of $x \in G$. Recall also that $\text{Aut} \mathfrak{g}$ is an affine algebraic $k$-group with $\text{Lie}(\text{Aut} \mathfrak{g}) = \text{Der} \mathfrak{g}$ (Exercise 11.6.2). Show:

(a) $\text{Ad} : G \rightarrow \text{Aut} \mathfrak{g}$ is a map of affine algebraic $k$-groups.

(b) $\text{ad} = \text{Lie}(\text{Ad}) : \mathfrak{g} \rightarrow \text{Der} \mathfrak{g} = \text{Lie}(\text{Aut} \mathfrak{g})$; see (5.6).

11.7. Algebraic Group Actions on Prime Spectra

We conclude the chapter on algebraic groups with an application to the study of prime spectra of algebras. Exercises 10.4.3, 10.4.4 and 11.7.2, 11.7.3 develop some of the material below in the larger context of Hopf algebra actions.

Any action of a group $G$ on an algebra $A$ is potentially a useful instrument for the investigation of the ideal structure of $A$. Indeed, since each $g \in G$ acts by an algebra automorphism, the image $g.I$ of any ideal $I$ of $A$ is an ideal such that $A/g.I \cong A/I$. Consequently, the properties of being maximal, prime, primitive etc. transfer from $I$ to $g.I$. Our focus will be on a rational action $G \triangleleft A$ of an algebraic group $G$ (§11.3.2) and on the resulting action $G \triangleleft \text{Spec} A$ on the set of prime ideals of $A$.

Below, we continue to assume that $G$ is an affine algebraic $k$-group and that $k$ is algebraically closed. Furthermore, we assume throughout that we are given a rational action $G \triangleleft A$ with $A \in \text{Alg}_k$.

11.7.1. $G$-Cores

Ideals of $A$ that are stable under the action of $G$ will be called $G$-ideals. If $I$ is an arbitrary ideal of $A$, then the sum of all $G$-ideals that are contained in $I$ is evidently the largest $G$-ideal that is contained in $I$; it will be called the $G$-core of $I$ and denoted by $I:_G G$. It is easy to see that

$$I:_G G = \bigcap_{g \in G} g.I.$$  

The following lemma gives another expression for $I:_G G$ in terms of the coaction $\delta : A \rightarrow A \otimes \mathcal{O}(G)$ for the given rational action $G \triangleleft A$. 

Lemma 11.29. The $G$-core of an ideal $I$ of $A$ is given by $I:G = \delta^{-1}(I \otimes \mathcal{O}(G))$.

Proof. Writing the coaction $\delta: A \to A \otimes \mathcal{O}(G)$ as $a \mapsto a(0) \otimes a(1)$, the composite of $\delta$ with the embedding $A \otimes \mathcal{O}(G) \to A \otimes \mathbb{K}^G$ that is given by (11.4) takes the form $a \mapsto a(0) \otimes (\langle g, a(1) \rangle)_{g \in G}$ for $a \in A$. Now $a \in I:G$ means that $g.a = a(0)\langle g, a(1) \rangle \in I$ for all $g \in G$ or, equivalently, $\delta(a) \in I \otimes \mathcal{O}(G)$.

Example 11.30 (Torus actions). Recall from Example 11.15 that a rational action of the torus $G = (\mathbb{K}^n)^n$ on $A$ amounts to a grading by the lattice $L = \mathbb{Z}^n$: so $A = \bigoplus_{\lambda \in L} A_\lambda$ and $A_\lambda A_\mu \subseteq A_{\lambda + \mu}$. Here, $\mathcal{O}(G)$ is the group algebra $\mathbb{K}L$ and the coaction $\delta: A \to A \otimes \mathcal{O}(G)$ is given by $\delta a = \sum_\lambda a_\lambda \otimes \lambda$, where $a_\lambda$ is the $\lambda$-homogeneous component of $a \in A$. Therefore, $I:G = \delta^{-1}(I \otimes \mathcal{O}(G)) = \bigoplus_{\lambda \in L}(I \cap A_\lambda)$, the largest graded ideal of $A$ that is contained in $I$.

11.7.2. Primes and $G$-Primes

A $G$-ideal $I$ of $A$ is said to be $G$-prime if $I \neq A$ and $JK \subseteq I$ for $G$-ideals $J, K$ of $A$ implies that $J \subseteq I$ or $K \subseteq I$. An easy argument shows that $G$-cores of prime ideals are $G$-prime. It is less clear that all $G$-primes arise as $G$-cores of primes, but this is in fact the case for rational actions by part (b) of the next proposition. Moreover, while $G$-ideals that are prime are evidently $G$-prime as well, $G$-primes need not be prime in general: the $G$-cores $P:G$ with $G$ finite and $P \in \text{Spec } A$ need not be prime.

Proposition 11.31. (a) For each $P \in \text{Spec } A$, the orbit map $G \to \text{Spec } A, \ g \mapsto g.P$ is continuous.

(b) The $G$-primes of $A$ are exactly the $G$-cores $P:G$ with $P \in \text{Spec } A$.

(c) If $G$ is connected, then all $G$-primes of $A$ are prime. Thus, $G$-primes of $A$ are identical to $G$-ideals that are prime in this case.

Proof. (a) Denoting the orbit map by $\pi$, we need to show that preimages $\pi^{-1}(\mathcal{V}(I))$ of closed subsets of $\mathcal{V}(I) \subseteq \text{Spec } A$ are closed in $G$. But $g \in \pi^{-1}(\mathcal{V}(I))$ if and only if $g^{-1}.a = a(0)\langle g^{-1}, a(1) \rangle \in P$ for all $a \in I$. Fix a basis $(b_j)_{j \in J}$ of $P$ and a basis $(b_k)_{k \in K}$ of a complement of $P$ in $A$ and write $a(0) \otimes a(1) = \sum_{j \in J} b_j \otimes a_{j,1} + \sum_{k \in K} b_k \otimes a_{k,1}$.

\[\text{This result is originally due to Chin [44], with a different proof. See also Exercise 11.7.3 for yet another proof.}\]
Then the condition $a_0(g^{-1}, a_{(1)}) \in P$ becomes $g^{-1}, a_{(1)} = 0$ for all $k \in K$ or, equivalently, $(g, S(a_{(1)})) = 0$ for all $k \in K$. Thus,

$$\pi^{-1}(T(I)) = \bigcap_{k \in K, a \in I} V(S(a_{k,1})).$$

Since the set on the right is closed, (a) is proved.

(b) We have already remarked that $P: G$ is $G$-prime for any prime $P$ of $R$. For the converse, let $I$ be a $G$-prime ideal of $A$. We will show that there is an ideal $P$ of $A$ which is maximal subject to the condition $P: G = I$; the ideal $P$ is then easily seen to be prime. In order to prove the existence of $P$, we use Zorn’s Lemma. So let $\{I_j\}$ be a chain of ideals of $A$ such that $I_j: G = I$ holds for all $j$ and consider the ideal $I_* = \bigcup_j I_j$. It suffices to show that $I_*: G = I$. For this, let $a \in I_*: G$ be given. Since the orbit $G.a$ spans a finite-dimensional subspace of $I_*$ by Proposition 11.12(c), it follows that $G.a \subseteq I_j$ for some $j$. Therefore, $a \in I_j: G = I$, as desired.

(c) The hypothesis that $G$ is connected is equivalent to irreducibility of $G$ for the Zariski topology (Proposition 11.20). Furthermore, continuous images of irreducible topological spaces are again irreducible and the closure of an irreducible subset is also irreducible. Therefore, (a) gives that the orbit $G.P$ of any $P \in \text{Spec } A$ and its closure $\overline{G.P}$ are irreducible subsets of $\text{Spec } A$. By the bijection (11.9), we conclude that the ideal $\mathcal{I}(\overline{G.P})$ is prime. But $\mathcal{I}(\overline{G.P}) = \mathcal{I}(G.P) = \bigcap_{Q \subseteq G.P} Q = P: G$ by (11.14), completing the proof.

**Example 11.32** (Torus actions). Since the algebraic torus $G = (k^\times)^n$ is connected (Exercise 11.5.3) and $G$-ideals of $A$ are the same as graded ideals for the $\mathbb{Z}^n$-grading of $A$ that corresponds to the action $G \subseteq A$ (Example 11.30), it follows from Proposition 11.31 that the graded ideal $\bigoplus_{\lambda \in \mathbb{Z}^n} (P \cap A_{\lambda})$ is prime for any $P \in \text{Spec } A$ and that $G$-primes of $A$ are the same as graded ideals that are prime.

### 11.7.3. A Stratification of the Prime Spectrum

The foregoing suggest the following approach to the analysis of Spec $A$:

**Step 1.** Determine the set $G$-Spec $A$ consisting of all $G$-prime ideals of $A$.

**Step 2.** For each $I \in G$-Spec $A$, describe $\text{Spec}_I A := \{P \in \text{Spec } A \mid P: G = I\}$.

The set $\text{Spec}_I A$ is the fiber over $I$ of the map $\text{Spec } A \to G$-Spec $A$, $P \mapsto P: G$ that is given by Proposition 11.31. The above steps provide us with the pieces of the partition\(^5\)

\[(11.15)\]

\[
\text{Spec } A = \bigsqcup_{I \in G \text{-Spec } A} \text{Spec}_I A
\]

\(^5\)This partition is known as the **Goodearl-Letzter stratification** of Spec $A$, the parts $\text{Spec}_I A$ being the strata; the partition was pioneered by Goodearl and Letzter [86].
We remark that, for many specific algebras $A$ that are of interest, there is a rational action $G \subseteq A$ of a suitable algebraic group $G$, often a torus, such that $G$-Spec $A$ is finite. Furthermore, for any rational action of an algebraic torus, one can identify each $\text{Spec}_I A$ with the prime spectrum of a suitable commutative algebra. The interested reader may consult [34] as well as [138], [139] for further background and details.

11.7.4. An Example: The Prime Spectrum of Quantum Affine Space

In this subsection, we will carry out the steps laid out in §11.7.3 for the following algebra, which is known as quantum affine $n$-space:

$$A = O_q(\mathbb{k}_n) \overset{\text{def}}{=} \mathbb{k}(x_1, \ldots, x_n)/(x_i x_j - q_{ij} x_j x_i \mid i < j).$$

Here, $q = \{q_{ij}\} \subseteq \mathbb{k}$ are given parameters. The special case of the quantum plane ($n = 2$) was looked at in Example 1.24 and in Exercise 1.1.15; the case where all $q_{ij}$ are identical was mentioned in Example 9.22 and in Exercise 9.3.13. Below, we assume that the reader is somewhat familiar with these examples and exercises—many features discussed there generalize in a straightforward manner to the current situation—and we shall also invoke a modicum of noncommutative ring theory.

As in the earlier instances, we retain the notation $x_i$ for the image of $x_i$ in $A$. The standard monomials $x^A = x_1^{a_1} x_2^{a_2} \cdots x_n^{a_n}$ with $A = (a_1, \ldots, a_n) \in \mathbb{Z}_+^n$ span $A$ as a $\mathbb{k}$-vector space, because reordering the factors in any finite product of the generators $x_i$ only changes the product by a scalar factor. In fact, the $x^A$ form a $\mathbb{k}$-basis of $A$; this follows either from the Diamond Lemma (Example D.6) or else by realizing $A$ as an iterated skew polynomial algebra as in Exercise 1.1.15 ($n = 2$): $A \equiv A'[x_n; \sigma]$, where $A'$ is a quantum affine $(n - 1)$-space. The latter method also shows that $A$ is a noetherian domain. Since $x^A x^\mu \in \mathbb{k} x^{\lambda + \mu}$, we obtain an $\mathbb{Z}_+^n$-grading of $A$:

$$A = \bigoplus_{A \in \mathbb{Z}_+^n} A_A \quad \text{with} \quad A_A = \mathbb{k} x^A.$$

This grading corresponds to a rational action of the algebraic torus $G = (\mathbb{k}^\times)^n$ on $A$, which is given by $g \cdot x_i = \gamma_i x_i$ for $g = (\gamma_1, \ldots, \gamma_n) \in G$ (Example 11.15).

For Step 1, we need to determine $G$-Spec $A$, the collection of all graded prime ideals of $A$ (Example 11.32). First note that, for any subset $X \subseteq \{x_1, \ldots, x_n\}$, the quotient $A/(X)$ is a quantum affine $(n - |X|)$-space. Therefore, $(X)$ is a (completely) prime ideal of $A$ which is graded. Conversely, if $I \in G$-Spec $A$, then $I = \bigoplus A(I \cap \mathbb{k} x^L)$ and $I$ is prime. Therefore, $x^L \in I$ if and only if the monomial $x^L$ involves a factor $x_i \in X_L := I \cap \{x_1, \ldots, x_n\}$; this follows from the fact that all generators $x_i$ are normal elements of $A$ in the sense that $A x_i = x_i A$. Consequently, $I = (X_L)$. In summary, $G$-Spec $A$ is in bijection with the subsets of $\{x_1, \ldots, x_n\}$; so

$$\# G\text{-Spec } A = 2^n.$$
Furthermore, all $G$-primes are in fact completely prime.

In Step 2, we describe the stratum $\text{Spec}_I A$, where $I = (X)$ for a given subset $X \subseteq \{x_1, \ldots, x_n\}$. Explicitly, $\text{Spec}_I A$ consists of all prime ideals of $A$ that contain $X$ but none of the “variables” in $\{x_1, \ldots, x_n\} \setminus X$. As we have remarked above, $A/(X)$ is again a quantum affine space. Thus, it suffices to treat the case $X = \emptyset$ and describe the stratum $\text{Spec}_{(0)} A$, consisting of all $P \in \text{Spec} A$ with $x_i \notin P$ for all $i$ or, equivalently, $x := x_1 x_2 \ldots x_n \notin P$. For this, we consider the following algebra, called a quantum affine $n$-torus:

$$B = O_q((\mathbb{K}^x)^n) \overset{\text{def}}{=} \mathbb{K}(x_1^{\pm 1}, \ldots, x_n^{\pm 1})/(x_i x_j - q_{ij} x_j x_i \mid i < j).$$

As in Exercise 1.1.15 ($n = 2$), one sees that $A \subseteq B$ and that $B$ is an iterated skew Laurent polynomial algebra: $B \cong B'[x_n^{\pm 1}, \sigma]$, where $B'$ is a quantum affine $(n-1)$-torus. Now the standard monomials $x^\lambda = x_1^{\lambda_1} x_2^{\lambda_2} \ldots x_n^{\lambda_n}$ with $\lambda = (\lambda_1, \ldots, \lambda_n) \in \mathbb{Z}_+^n$ form a $\mathbb{K}$-basis of $A'$ and $x^\lambda x^\mu \in \mathbb{K} x^{\lambda+\mu}$ as before. Thus, $B$ is $\mathbb{Z}_+^n$-graded, with $B_\lambda = \mathbb{K} x^\lambda$, and

$$A = \bigoplus_{\lambda \in \mathbb{Z}_+^n} \mathbb{K} x^\lambda \subseteq B = \bigoplus_{\lambda \in \mathbb{Z}^n} \mathbb{K} x^\lambda.$$

The algebra $B$ can also be described as the localization $A[x^{-1}]$ of $A$ at the powers of the normal element $x \in A$. See [149, Chapter 2 §1] for background on noncommutative localization. In particular, by [149, Proposition 2.1.16], the prime ideals of $A$ that do not contain $x$ are in bijection with the primes of $B$ via extension and contraction, that is, $P \mapsto PB$ and $Q \mapsto Q \cap A$. Thus, $\text{Spec}_{(0)} A \cong \text{Spec} B$. The following lemma will further show that $\text{Spec} B \cong \text{Spec} \mathcal{Z} B$. Consequently,

$$\text{Spec}_{(0)} A \cong \text{Spec} \mathcal{Z} B.$$

This identifies all strata of $\text{Spec} A$ with spectra of commutative algebras as was our goal in Step 2.

**Lemma 11.33.** Contraction $I \mapsto I \cap \mathcal{Z} B$ and extension $J \mapsto J B$ give inverse bijections between the sets of ideals of $B$ and $\mathcal{Z} B$; they restrict to inverse bijections $\text{Spec} B \xleftrightarrow{*} \text{Spec} \mathcal{Z} B$. Furthermore, $\mathcal{Z} B$ is a Laurent polynomial algebra over $\mathbb{K}$.

**Proof.** The assertion about prime spectra is an easy consequence of the correspondence for all ideals; so we will focus on the latter. Write $Z = \mathcal{Z} B$ and let $M = \{\lambda \in \mathbb{Z}^n \mid x^\lambda \in Z\}$. Then $M$ is a sublattice of $\mathbb{Z}_+^n$ and $Z = \bigoplus_{\lambda \in M} \mathbb{K} x^\lambda \cong \mathbb{K} M$, the group algebra of $M$. Thus, $Z$ is a Laurent polynomial algebra over $\mathbb{K}$ as claimed. Furthermore, $B$ is free as $Z$-module:

$$B = \bigoplus_{\mu \in \mathbb{Z}^n/M} x^\mu Z.$$

The projection onto the summand for $\mu = 0$ in this decomposition is a $Z$-linear epimorphism $\pi : B \rightarrow Z$. Therefore, for any ideal $J$ of $Z$, the extended ideal $JB$ of $B$ satisfies $\pi(JB) = J \pi(B) = J$, proving that the extension map is injective on
the set of ideals of \( Z \). It remains to prove that \( I = (I \cap Z)B \) for any ideal \( I \) of \( B \). Suppose otherwise and pick an element \( a = \sum_{\mu \in \mathbb{Z}^n/M} x^{\mu}z_\mu \in I \setminus (I \cap Z)B \) involving a minimal number of nonzero summands. Since all \( x^{\mu} \) are units in \( B \), we may assume that \( a = z_0 + x^{\mu_1}z_1 + \cdots + x^{\mu_n}z_n \) with \( n \geq 1 \), pairwise distinct \( 0 \neq \mu_i \in \mathbb{Z}^n/M \) and with \( z_i \in Z \setminus (I \cap Z) \). But then, for any \( \lambda \in \mathbb{Z}^n \), the Lie commutator

\[
[a, x^\lambda] = [x^{\mu_1}, x^\lambda]z_1 + \cdots + [x^{\mu_n}, x^\lambda]z_n \in I \cap \sum_{i=1}^n x^{\mu_i+\lambda}Z
\]

involves fewer summands than \( a \). Therefore, \([a, x^\lambda] \in (I \cap Z)B\). Since \( 0 \neq [x^{\mu_n}, x^\lambda] \) for some \( \lambda \in \mathbb{Z}^n \), it follows that \( z_n \in I \cap Z \), a contradiction. \( \square \)

To finish our discussion of quantum affine space, let us redraw the picture of the prime spectrum of the quantum plane \( \mathcal{O}_q(\mathbb{k}^2) = \mathbb{k}[x, y] \) with \( xy = qyx \) for a non-root of unity \( q \in \mathbb{k}^\times \) from Example 1.24. The acting group is the 2-torus \( G = (\mathbb{k}^\times)^2 \) and there are four \( G \)-primes: (0), (x), (y) and (x, y). The strata of (0) and (x, y), depicted in red and black, are both identical to \text{Spec} \( \mathbb{k} \), consisting of one point each; the strata of (x) and (y), in green and blue, are identical to the prime spectrum of a Laurent polynomial algebra in one variable over \( \mathbb{k} \).

11.7.5. Outlook: The Prime Spectrum of Quantum Matrices

The description of \text{Spec} \( A \) is a more challenging task, and more interesting, for the \( \mathbb{k} \)-algebra \( A = \mathcal{O}_q(\text{Mat}_n) \) of quantum \( n \times n \) matrices (Example 9.22). Once again, there is a fortuitous rational action by a large algebraic torus, now the torus \( G = (\mathbb{k}^\times)^{2n} \). Assembling the algebra generators \( X_{ij} \) of \( A \) into the “generic matrix” \( X = (X_{ij})_{n \times n} \), the action of the element \( g = (\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n) \in G \) on \( A \) is given by

\[
g.X = \begin{pmatrix} \alpha_1 & & \cdots & & \alpha_n \\ & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots \\ & \cdots & \ddots & \ddots & \ddots \\ \beta_1 & & \cdots & & \beta_n \end{pmatrix} X \begin{pmatrix} \beta_1 & & \cdots & & \beta_n \\ & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots \\ & \cdots & \ddots & \ddots & \ddots \\ \alpha_1 & & \cdots & & \alpha_n \end{pmatrix}^{-1}.
\]

In this subsection, we shall endeavor to illustrate the statement—but not the proof—of the following theorem, which achieves Step 1 of §11.7.3 with the optimal outcome that \( G \)-\text{Spec} \( A \) is finite. The theorem summarizes work of Cauchon [39], [40], Launois [131] and others. For information on the Stirling numbers of the 2nd kind, we refer to Stanley [191], but the remaining terms in the theorem will be explained.

---

6Part (c), under the stated hypothesis on \( q \), is a special case of a more general result due to Geiger and Yakimov [206], [81].
Theorem 11.34. Let $A = O_q(\text{Mat}_n)$, where $q \in \mathbb{K}^\times$ is not a root of unity.

(a) There is a bijection between $G$-$\text{Spec} A$ and a certain collection of $n \times n$ diagrams, called Cauchon diagrams.

(b) $\# G$-$\text{Spec} A = \sum_{t=0}^n (t)^2 S(n+1, t+1)^2$, where the $S(\cdot, \cdot)$ are Stirling numbers of the 2nd kind.

(c) There is an order isomorphism between $(G$-$\text{Spec} A, \subseteq)$ and the set of permutations $\{ s \in S_{2n} \mid |si - i| \leq n \text{ for all } i = 1, \ldots, 2n \}$ with respect to the Bruhat order on $S_{2n}$.

Cauchon Diagrams. These are $n \times n$ arrays of black and white boxes satisfying the following requirement: if a box is colored black then all boxes on top of it or all boxes to the left must be black as well. For obvious reasons, Cauchon diagrams are also called $\perp$-diagrams (“le”).

For $n = 2$, there are 14 Cauchon diagrams. Only two possible black-and-white colorings of $2 \times 2$ arrays fail the $\perp$-test; the offending boxes are marked in red:

![Cauchon Diagrams](image)

From Diagrams to Permutations: “Pipe Dreams”. To explain the bijection between $n \times n$ Cauchon diagrams and the set of permutations in part (c) of the theorem, place the labels $1, \ldots, 2n$ along the edges of a given Cauchon diagram and arrows in each box as illustrated in the following example for $n = 6$:

![Pipe Dreams](image)
Following the arrows from edge to edge gives a permutation \( s \in S_{2n} \), which clearly satisfies the condition \(|si - i| \leq n\) for all \( i \), because none of the arrows move to the right or down. We will refer to the permutations in part (c) as restricted permutations. The above \( 6 \times 6 \) Cauchon diagram, for example, gives the restricted permutation \( s = (3765)(49)(1112) \in S_{12} \). In general, one can show that this process yields the desired bijection.

**Diagrams and G-Primes.** The bijection between \( G\)-Spec \( A \) and the collection of Cauchon diagrams is harder to explain and we must refer the reader to the cited references for this. Table 11.1 illustrates the situation for \( n = 2 \). We write the generic \( 2 \times 2 \) matrix as \( \begin{pmatrix} a & b \\ c & d \end{pmatrix} \). In the first picture, showing \( G\)-Spec \( A \) along with its inclusions, parentheses also denote the ideal of \( A \) that is generated by the indicated elements, which are arranged in a \( 2 \times 2 \) array so as to mirror the corresponding Cauchon diagram in the second picture. Granting that the displayed ideals are indeed exactly the \( G\)-primes of \( A \), which we shall not prove, the bijection with Cauchon diagrams is quite apparent here, with containment of \( G\)-primes being reflected by containment of diagrams in all cases except for the match \((D_q) \leftrightarrow \square\). For larger values of \( n \), the correspondence is more complex.

**Bruhat Order.** Fulton [70, §10.5] offers several descriptions of the Bruhat order for the symmetric group \( S_m \). The standard definition is as follows: if \( s, w \in S_m \) then \( s \leq w \) iff some substring of some (or every) reduced word for \( s \) is a reduced word for \( w \). Here, a substring need not be consecutive and a reduced word for \( s \) is a product of \( \ell(s) \) many Coxeter generators \( s_i = (i, i+1) \). Recall that the length \( \ell(s) \) is the number of inversions, \( \ell(s) = \# \{ (i, j) \in [m]^2 | i < j \text{ but } si > sj \} \) (Example 7.10). To state this more directly, write \( s[i] = \{s_1, s_2, \ldots, s_i\} \) with \( s_1 < s_2 < \cdots < s_i \) and similarly for \( w \). Define \( s \leq w \) iff \( s_k \leq w_k \) for all \( k \in [i] \). Then \( s \leq w \) if and only if \( s \leq i \) \( w \) for all \( i \in [m] \) (Exercise 11.7.4).

In the present context, we take \( m = 2n \) and we write each \( s \in S_m \) as the string \( s_1 s_2 \cdots s_m \). Define \( t \in S_{2n} \) to be the permutation \( n+1 n+2 \cdots 2n 1 2 \cdots n \); so \( \ell(t) = n^2 \) and \( t \) is restricted: \( |ti - i| = n \) for all \( i \). In fact, one can show that the Bruhat interval \( \{ s \in S_{2n} | (1) \leq s \leq t \} \) coincides with the collection of restricted permutations in \( S_{2n} \) (Exercise 11.7.4). The last diagram in Table 11.1 shows this interval for \( n = 2 \), with levels indicating length and edges giving the Bruhat order.

---

7The \( G\)-prime spectrum for quantum \( 2 \times 2 \) matrices was first determined by Goodearl and Lenagan [84].
8Dixmier [60, 7.7.3] uses the reverse partial order and covers more general Weyl groups.
**Exercises for Section 11.7**

**11.7.1 (G-primes and the extended center).** Let $A \in \text{Alg}_k$, with $k$ an arbitrary field, and let $G$ be an arbitrary group acting by $k$-algebra automorphisms on $A$. Observe that the definition of $G$-Spec $A$ (§11.7.2) extends verbatim to this setting. Show:

(a) The $G$-action on $A$ extends uniquely to an action by $k$-algebra automorphism on the symmetric ring of quotients $QA$, giving an action $G \subseteq CA$ (Appendix E).

(b) For $I \in G$-Spec $A$, the invariant algebra $C(A/I)^G$ is an extension field of $k$. (Adapt the proof of Proposition E.2.)
11.7.2 (Primeness of $H$-cores). In this exercise, the base field $k$ need not be algebraically closed. We assume familiarity with Exercises 10.4.3 and 10.4.5.

(a) Let $A \in \text{Alg}_k$ and consider the power series algebra $P = A[[X_\lambda \mid \lambda \in \Lambda]]$, where $\Lambda$ is any set (Example 3.3). Let $B$ be a subring of $P$ mapping onto $A$ under $P \to A$, $X_\lambda \mapsto 0$. Show that if $A$ is prime (semiprime, a domain), then so is $B$.

(b) Let $g \in \text{Lie}_k$ (not necessarily finite dimensional) and let $A \in g \text{Alg}$ (§5.5.5). Assume that $\text{char} k = 0$ and put $H = U_g$. Use (a) and Example 9.6 to show that if $I$ is a prime, semiprime or completely prime ideal of $A$, then $I:H$ is likewise.

11.7.3 (Primeness of $H$-cores, again). Let $H \in \text{HopfAlg}_k$ and let $A \in \text{Alg}_H$, with $H$ acting locally finitely on $A$. View $A \in \text{Alg}^H$ and let $\delta : A \to A \otimes H^\circ$ denote the $H^\circ$-coaction (Proposition 10.26). Define $H'(A)$ to be the the smallest $k$-subalgebra of $H^\circ$ such that $\delta A \subseteq A \otimes H'(A)$ and let $H(A)$ be the smallest $S^\circ$-stable subalgebra of $H^\circ$ containing $H'(A)$. Show:

(a) $H'(A)$ is a subbialgebra of $H^\circ$ and $H(A)$ is a Hopf subalgebra; they are not necessarily the same. (Use Exercise 9.2.3.)

(b) Assume that $H(A)$ is a commutative domain. If $I$ is a prime, semiprime or completely prime ideal of $A$, then $I:H$ is likewise. (Use Lemma 11.19 and Exercises 1.3.5, 10.4.6.)

11.7.4 (Bruhat order). (a) For $s, w \in S_m$, show that $s \leq w$ for the Bruhat order if and only if $s \leq_i w$ for all $i \in [m] = \{1, \ldots, m\}$. See §11.7.5 for the definition of $\leq$ in terms of reduced words and for $\leq_i$.

(b) Let $t = n + 1 \cdot n + 2 \cdot 2n + 1 \cdot 2 \cdot n \in S_{2n}$. Show that the Bruhat interval \{s \in S_{2n} \mid (1) \leq s \leq t\} coincides with the collection of restricted permutations, \{s \in S_{2n} \mid |si - i| \leq n \text{ for all } i \in [2n]\}. 
In this final chapter, the spotlight is on a Hopf algebra $H$ that is finite dimensional. We have already seen that the antipode $S$ is bijective in this case (Theorem 10.9). This fact will be significantly strengthened in Radford’s formula, which shows that $S$ has finite order, that is, $S^k = S \circ S \circ \cdots \circ S = \text{Id}_H$ for some $k \in \mathbb{N}$ (Theorem 12.10). We have also seen that if $H$ is involutory (i.e., $S^2 = \text{Id}_H$) and $\text{char} \ k \nmid \dim_k H$, then $H$ is semisimple (Corollary 10.17); of course, $H^*$ is semisimple as well in this case, since the hypotheses carry over. Corollary 12.13 improves upon this by giving a necessary and sufficient condition for $H$ and $H^*$ to both be semisimple. Other highlights in this chapter include the celebrated results of Larson and Radford on semisimplicity of Hopf algebras in characteristic 0 (Theorem 12.15) and the Nichols-Zoeller Theorem (§12.4.5), a version of Lagrange’s Theorem from group theory for finite-dimensional Hopf algebras.

Throughout this chapter, $H = (H, m, u, \Delta, \varepsilon, S)$ denotes a finite-dimensional Hopf algebra over an arbitrary field $k$. Further assumptions will be spelled out as needed.

12.1. Frobenius Structure

Our first goal is to show that $H$ is a Frobenius algebra and to exhibit the data that are associated with this structure: Frobenius form, Casimir element and Nakayama automorphism; these will all be instrumental in proving the main results on finite-dimensional Hopf algebras in subsequent sections. We need to assume that the reader is reasonably comfortable with the material in Section 2.2. Our presentation in this section owes much to Schneider’s Cordoba lecture notes [179].
12.1.1. Integrals

The space of $H$-invariants in the (left) regular representation of $H$, which was earlier shown to be 1-dimensional (Theorem 10.9), will now be denoted by $I^l_H$; its elements will be called the left integrals of $H$ and frequently be written as $\Lambda$. Thus,

$$I^l_H \overset{\text{def}}{=} \{ \Lambda \in H \mid h\Lambda = \langle \varepsilon, h \rangle \Lambda \text{ for all } h \in H \}$$

Similarly, right integrals in $H$ are the invariants of the right regular representation:

$$I^r_H \overset{\text{def}}{=} \{ \Lambda \in H \mid \Lambda h = \langle \varepsilon, h \rangle \Lambda \text{ for all } h \in H \}$$

These definitions do of course make sense for any Hopf algebra, but all left and right integrals of infinite-dimensional Hopf algebras are 0 by Proposition 10.6.

Since $S$ is an algebra anti-automorphism of $H$ satisfying $\varepsilon \circ S = \varepsilon$, it is straightforward to see that $S(I^l_H) \subseteq I^r_H$ and $S(I^r_H) \subseteq I^l_H$; in fact, equality must hold throughout, because $I^l_H$ is 1-dimensional. To summarize,

(12.1) \[ S(I^l_H) = I^r_H, \quad S(I^r_H) = I^l_H \quad \text{and} \quad \dim_k I^l_H = \dim_k I^r_H = 1. \]

If $I^r_H = I^l_H$, then the Hopf algebra $H$ is called unimodular; in this case, we will write $I_H$ for the space of integrals.

Example 12.1 (Finite group algebras). Consider the group algebra $\mathbb{k}G$ of a finite group $G$. We had shown in (9.1.4) that $I^l_{\mathbb{k}G} = \mathbb{k}\Lambda$ with $\Lambda = \sum_{x \in G} x$. Since $S\Lambda = \Lambda$, we also have $I^r_{\mathbb{k}G} = \mathbb{k}\Lambda$. Thus, $\mathbb{k}G$ is unimodular.

Example 12.2 (The Taft Hopf algebras). Let $H = H_{n,q}$ be the Taft algebra, with $q \in \mathbb{k}$ a root of unity of order $n \geq 2$ (Example 9.23). Recall that $H$ has $\mathbb{k}$-basis $g^i x^j$ ($0 \leq i, j < n$) and algebra relations $g^n = 1$, $x^n = 0$, $gx = q \cdot xg$. The comultiplication, counit and antipode are given by

$$\Delta g = g \otimes g, \quad \langle \varepsilon, g \rangle = 1, \quad Sg = g^{-1}$$

$$\Delta x = x \otimes 1 + g \otimes x, \quad \langle \varepsilon, x \rangle = 0, \quad Sx = -g^{-1}x$$

Put $\Lambda = \sum_{i=0}^{n-1} g^i x^{n-1}$. One checks easily that $g\Lambda = \Lambda$ and $x\Lambda = 0$. Thus, $\Lambda$ is a nonzero left integral of $H$ and $I^l_H = \mathbb{k}\Lambda$. Similarly, $I^r_H = \mathbb{k}\Lambda'$ with $\Lambda' = \sum_{i=0}^{n-1} x^i g^i = \sum_{i=0}^{n-1} g^i x^{n-1}$. Therefore, $I^r_H \neq I^l_H$ and $H$ is not unimodular.

Integrals in the dual Hopf algebra $H^*$. The definition of left integrals, when spelled out for the Hopf algebra $H^*$, states that a linear form $\lambda \in H^*$ belongs to $I^l_{H^*}$ if and only if $\lambda f = \langle f, 1_H \rangle \lambda$ for all $f \in H^*$ or, more explicitly, $\langle \lambda, h_{(1)} \rangle \langle f, h_{(2)} \rangle = \langle f, 1_H \rangle \langle \lambda, h \rangle$ for all $f \in H^*$ and $h \in H$. This condition can be stated more concisely in the convolution algebra $\text{End}_k(H) \cong H \otimes H^*$ (§9.1.4). We view $H^*$
as a subalgebra of $\text{End}_\kappa(H)$ as in (9.13), identifying $f \in H^*$ with the map $(h \mapsto \langle f, h \rangle_1) \in \text{End}_\kappa(H)$. Then

$$\lambda \in J^*_H. \quad \iff \quad \forall h \in H : \langle \lambda, h \rangle_1 = \langle \lambda, h_1 \rangle h_2$$

(12.2)

Similarly,

$$\lambda' \in J^*_H. \quad \iff \quad \forall h \in H : \langle \lambda', h \rangle_1 = h_1 \langle \lambda', h_2 \rangle$$

(12.3)

**Example 12.3** (Duals of finite group algebras). Let $G$ be a finite group. Recall from Example 9.17 that, as a $\kappa$-algebra, $(\kappa G)^*$ is isomorphic to the direct product of $|G|$ copies of $\kappa$. The elements $\delta_x (x \in G)$ that are defined by $\langle \delta_x, y \rangle = \delta_{x,y} 1_\kappa$ for $x, y \in G$ form a $\kappa$-basis consisting of orthogonal idempotents of $(\kappa G)^*$ and such that $\sum_{x \in G} \delta_x = 1_{(\kappa G)^*}$. The element $\lambda = \delta_1$ satisfies (12.2) and (12.3): $\langle \lambda, x \rangle 1 = \langle \lambda, x \rangle x$ for all $x \in G$. Therefore, $\lambda$ is a left and right integral of $(\kappa G)^*$; it is the trace form that we know from (3.14): $\langle \lambda, \sum_{x \in G} \alpha_x x \rangle = \alpha_1$.

12.1.2. Modular Elements

Fix $0 \neq \Lambda \in J^*_H$. It is easy to see that $h\Lambda$ is also a right integral of $H$ for any $h \in H$. Since $J^*_H$ is 1-dimensional by (12.1), it follows that

$$h\Lambda = \langle \alpha, h \rangle \Lambda$$

(12.4)

for some $\langle \alpha, h \rangle \in \kappa$. The scalar $\langle \alpha, h \rangle$ is independent of the particular choice of $\Lambda$, which is unique up to scalar multiples, and the resulting map $\alpha : H \to \kappa$ is evidently an algebra map; so $\alpha \in G(H^*) = \text{Hom}_{\text{Alg}}(H, \kappa)$. The element $\alpha$ is often called the **distinguished group-like element** of $H^*$ or the **right modular element**. Clearly,

$$H \text{ is unimodular } \iff \alpha = e$$

(12.5)

Employing left integrals instead of left integrals, we can similarly define a **left modular element** $\alpha' \in G(H^*)$; it is determined by the condition

$$\Lambda' h = \langle \alpha', h \rangle \Lambda'$$

(12.6)

for all $h \in H$ and any $0 \neq \Lambda' \in J^*_H$. The computation $\langle \alpha', h \rangle S \Lambda' = S h \Lambda' \rangle = \langle \alpha^{-1}, h \rangle S \Lambda'$ shows that

$$\langle \alpha, Sh \rangle S \Lambda' = \langle S^* \alpha, h \rangle S \Lambda' = \langle \alpha^{-1}, h \rangle S \Lambda'$$

(12.7)

$\alpha' = \alpha^{-1}$

**Example 12.4** (Taft algebras). Let $H = H_{n,q}$ be as in Example 12.2 and recall that $0 \neq \Lambda = \sum_{i=0}^{n-1} x^{n-1} \varphi' \in J^*_H$. One easily checks that $x \Lambda = 0$ and $g \Lambda = q^{-1} \Lambda$. Thus, the right modular element $\alpha \in G(H^*)$ is given by $\langle \alpha, x \rangle = 0$, $\langle \alpha, g \rangle = q^{-1}$. 

Modular elements of $H$. Working with integrals of the dual Hopf algebra $H^*$, we obtain analogous distinguished group-like elements of $H$. In detail, fixing $0 \neq \lambda \in \ell^1_{H^*}$, equation (12.4) turns into the condition $f \lambda = \langle f, a \rangle \lambda$ for all $f \in H^*$ defining the right modular element $a \in G(H)$. Upon evaluation, this takes the following form:

$$h(1)\langle \lambda, h(2) \rangle = a \langle \lambda, h \rangle \quad (h \in H)$$

(12.8)

This condition can again be stated in the convolution algebra $\text{End}_k(H)$. In addition to the embedding $H^* \hookrightarrow \text{End}_k(H)$, $f \mapsto (k \mapsto \langle f, k \rangle)_H$, we also view $H$ as a subalgebra of $\text{End}_k(H)$ via $h \mapsto (k \mapsto \langle e, k \rangle h)$ as in (9.12). Then the endomorphism $f \ast h = h \ast f \in \text{End}_k(H)$ is given by $k \mapsto \langle f, k \rangle h$. Condition (12.8) is equivalent to

$$a \ast \lambda = \lambda \ast a = \text{Id}_H \ast \lambda$$

(12.9)

Starting with $0 \neq \lambda' \in \ell^1_{H^*}$, we obtain the analogous identity for the left modular element of $H$, which is equal to $a^{-1}$ by (12.7):

$$a^{-1} \ast \lambda' = \lambda' \ast a^{-1} = \lambda' \ast \text{Id}_H$$

12.1.3. Frobenius Form and Casimir Element

Recall that $H^*$ is an $(H, H)$-bimodule with actions

$$\langle h \ast f \ast k, l \rangle = \langle f, klh \rangle \quad (h, k, l, f \in H^*)$$

and that $H$ is said to be Frobenius if $H^*$, viewed as a left $H$-module, is isomorphic to the left regular module $H_{\text{reg}} = \text{H-reg} H$. By Lemma 2.11, this is equivalent to the analogous right-sided condition and also to the existence of an element $\lambda \in H^*$ such that

$$H^* = \lambda \ast H = H \ast \lambda$$

Any such $\lambda \in H^*$ is called a Frobenius form for $H$. The following proposition revisits Theorem 10.9 and some observations made in the course of its proof.

**Proposition 12.5.** Let $H$ be a finite-dimensional Hopf algebra. Then $H$ is a Frobenius algebra and any left or right integral $0 \neq \lambda \in H^*$ serves as Frobenius form.

**Proof.** In the proof of Theorem 10.9, it was shown that $M = (H_{\text{reg}})^*$, equipped with the left $H$-action (10.10), is isomorphic to $H \otimes M^{\text{coH}}$, where $H$ acts on the left factor and $M^{\text{coH}} = \ell^1_{H^*}$ in our current notation; see (10.17). Thus, $H^* = H \ast \lambda$ for any $0 \neq \lambda \in \ell^1_{H^*}$. Since the $H$-action (10.10) can be written as $h \ast f = f \ast Sh$ for $f \in H^*$ and $h \in H$, we obtain

$$H^* = H \ast \lambda = \lambda \ast SH = \lambda \ast H = \lambda \ast \lambda$$

Lemma 2.11

This shows that $\lambda$ is a Frobenius form for $H$. The identity $(S^* f) \ast h = S^* (Sh \ast f)$, which is straightforward to verify, gives $(S^* \lambda) \ast H = S^* (H \ast \lambda) = S^* H^* = H^*$. This
proves that $S^*\lambda$ is also a Frobenius form for $H$, and hence so is any nonzero left integral of $H^*$ by (12.1). This completes the proof of the proposition. \hfill \Box

Let us fix a left integral
\[ 0 \neq \lambda \in \mathfrak{l}_H^r, \]
and use it as our Frobenius form of $H$ in what follows. Thus, we have a right $H$-module isomorphism $\lambda \mapsto H \cong H^*$. Our next goal is to identify the Casimir element $c_{\lambda \mapsto} = y_i \otimes x_i \in H \otimes H$ of the Frobenius algebra $(H, \lambda)$. We continue with the convention of Section 2.2 that summation over repeated indices is understood. Recall the definition of $c_{\lambda \mapsto}$; see (2.23) and (2.27):

\[
H \otimes H \xrightarrow{\sim} H \otimes H^* \xrightarrow{\text{can.}} \text{End}_k(H)
\]

\[\begin{align*}
c_{\lambda \mapsto} = y_i \otimes x_i & \quad \mapsto \quad y_i \otimes (\lambda \mapsto x_i) \quad \mapsto \quad \text{Id}_H
\end{align*}\]

(12.11)

Let $\Lambda \in H$ be the unique element satisfying $\lambda - \Lambda = \epsilon$ and note that

(12.12) \[\lambda - \Lambda = \epsilon \iff \Lambda \in \mathfrak{f}_H^r \text{ and } \langle \lambda, \Lambda \rangle = 1\]

Indeed, the left-hand side gives $\lambda - \Lambda h = \epsilon - h = \langle \epsilon, h \rangle \epsilon = \lambda - \langle \epsilon, h \rangle \Lambda$ for all $h \in H$, whence $\Lambda \in \mathfrak{f}_H^r$. Conversely, if $\Lambda \in \mathfrak{f}_H^r$, then $\lambda - \Lambda = \epsilon \langle \lambda, \Lambda \rangle$.

**Proposition 12.6.** Let $H$ be a finite-dimensional Hopf algebra. With $0 \neq \lambda \in \mathfrak{l}_H^r$, as Frobenius form, we have

\[c_{\lambda \mapsto} = S(\Lambda_{(1)}) \otimes \Lambda_{(2)} = S^{-1}(\Lambda_{(2)}) a^{-1} \otimes \Lambda_{(1)}\]

where $\Lambda \in H$ is as in (12.12) and $a \in G(H)$ is the right modular element (12.9).

**Proof.** Recall that the canonical isomorphism $\text{End}_k(H) \cong H \otimes H^*$ is an isomorphism of $k$-algebras for the convolution algebra structure of $\text{End}_k(H)$, with $f * h = h * f \in \text{End}_k(H)$ for $f \in H^*$ and $h \in H$ corresponding to $h \otimes f \in H \otimes H^*$.

The first equality states that $S(\Lambda_{(1)}) * (\lambda - \Lambda_{(2)}) = \text{Id}_H$ in $\text{End}_k(H)$. To prove this, recall from §10.4.2 that $\text{End}_k(H)$ is a right $H$-module algebra for $\mapsto$ and note that $\text{Id}_H \mapsto h = h_H \in \text{End}_k(H)$ for $h \in H$. With this in hand, we compute

\[
S(\Lambda_{(1)}) * (\lambda - \Lambda_{(2)}) = (12.3) \quad S(\Lambda_{(1)}) * (\text{Id}_H * \lambda - \Lambda_{(2)})
\]

\[\begin{align*}
&= S(\Lambda_{(1)}) * (\text{Id}_H - \Lambda_{(2)}) * (\lambda - \Lambda_{(3)}) \\
&= S(\Lambda_{(1)}) * (\Lambda_{(2)})_H * (\lambda - \Lambda_{(3)}) \\
&= \langle \epsilon, \Lambda_{(1)} \rangle \text{Id}_H * (\lambda - \Lambda_{(2)}) \\
&= \text{Id}_H * (\lambda - \Lambda) \quad (12.12)
\end{align*}\]
The verification of the second equality also uses the identity \((h \ast f) - k = h \ast (f - k)\) for \(f \in H^*\) and \(h, k \in H\):
\[
S^{-1}(\Lambda(2)) \ast a^{-1} \ast (\lambda - \Lambda(1)) = S^{-1}(\Lambda(2)) \ast ((a^{-1} \ast \lambda) - \Lambda(1))
\]
\[(12.10)\]
The rest of the calculation, resulting in \(\text{Id}_H\), proceeds very much like above. This proves the proposition. \(\Box\)

Example 12.7 (Finite group algebras and their duals). For a finite group \(G\), we had seen in Examples 12.1 and 12.3 that nonzero (left and right) integrals of \((\mathbb{k}G)^*\) and \(\mathbb{k}G\) are given by \(\lambda = \delta_1 \in (\mathbb{k}G)^*\) and \(\Lambda = \sum_{x \in G} x \in \mathbb{k}G\). Since \(\langle \lambda, \Lambda \rangle = 1\) as stipulated (12.12), Proposition 12.6 gives \(c_{\lambda^{-1}} = \sum_{x \in G} x^{-1} \otimes x\) for \(\mathbb{k}G\), which was already exhibited in (3.15), and \(c_{\Lambda^{-1}} = \sum_{x \in G} \delta_x \otimes \delta_x\) for \((\mathbb{k}G)^*\).

12.1.4. The Nakayama Automorphism

Continuing with our choice of a fixed \(0 \neq \lambda \in \mathcal{I}_H^*\) as Frobenius form for \(H\), we will now determine the Nakayama automorphism \(\nu = \nu_\lambda \in \text{Aut}_{\text{Alg}}(H)\). Recall from (2.24) that \(\nu\) is given by
\[
\lambda - h = \nu(h) - \lambda \quad (h \in H)
\]
Since \(\lambda\) is unique up to scalar multiples, it is clear that \(\nu\) does not depend on the particular chosen left integral \(\lambda\). The formula for \(\nu\) in the following proposition uses the adjoint action of \(g \in G(H)\) on \(H\) as in (10.25):
\[
\text{ad} \, g := g(\cdot)g^{-1} \in \text{Aut}_{\text{Alg}}(H)
\]

Proposition 12.8. Let \(H\) be a finite-dimensional Hopf algebra. Then, choosing a nonzero left integral of \(H^*\) as Frobenius form, the associated Nakayama automorphism of \(H\) is given by
\[
\nu = S^2 \ast \alpha = \alpha \ast (S^{-2} \circ \text{ad} \, a)
\]
where \(\alpha \in G(H^*)\) and \(a \in G(H)\) are the right modular elements (12.4), (12.9).

Proof. We will use formula (2.25): \(\nu(h) = y_i \langle \lambda, h x_i \rangle\) for \(h \in H\). By Proposition 12.6, \(y_i \otimes x_i = S(\Lambda(1)) \otimes \Lambda(2) = S^{-1}(\Lambda(2))a^{-1} \otimes \Lambda(1)\) with \(\Lambda\) as in (12.12). Thus,
\[
\nu(h) = S(\Lambda(1)) \langle \lambda, h \Lambda(2) \rangle = S^{-1}(\Lambda(2))a^{-1} \langle \lambda, h \Lambda(1) \rangle
\]
(12.13) for all \(h \in H\). Furthermore, the right modular element \(\alpha \in H^*\) can be written as (12.14)
\[
\alpha = \Lambda - \lambda
\]
For, \(\langle \Lambda - \lambda, h \rangle = \langle \lambda, h \Lambda \rangle = \langle \alpha, h \rangle \langle \lambda, \Lambda \rangle = \langle \alpha, h \rangle\) for all \(h \in H\).
The first expression in (12.13) gives
\[
(S^{-2} \circ \nu)(h) = \langle \lambda, h \Lambda(2) \rangle S^{-1}(\Lambda(1)) = h(1) \Lambda(2) \langle \lambda, h(2) \Lambda(3) \rangle S^{-1}(\Lambda(1)) \\
= h(1) \langle \lambda, h(2) \Lambda \rangle = h(1) \langle \alpha, h(2) \rangle
\]
Applying $S^2$ to the first and last expressions, we obtain $\nu(h) = S^2(h(1)) \langle \alpha, h(2) \rangle$, proving the first formula for $\nu$ in the proposition.

The second formula follows from the second equality in (12.13). Indeed, applying the automorphisms $S^2$ and $ad a^{-1}$ to this equality and using the fact that $S^2 a = a$ by (9.30), we obtain
\[
a^{-1}(S^2 \circ \nu)(h)a = a^{-1} \langle \lambda, h \Lambda(1) \rangle S(\Lambda(2)) = S(\Lambda(1)) \langle \lambda, h(1) \Lambda(1) \rangle h(2) \Lambda(2) S(\Lambda(3)) \\
= \langle \lambda, h(1) \Lambda \rangle h(2) \langle \alpha, h(1) \rangle h(2)
\]
Finally, applying $S^{-2} \circ ad a$ gives $\nu(h) = \langle \alpha, h(1) \rangle (S^{-2} \circ ad a)(h(2))$, which is the second formula.

12.1.5. Symmetric Hopf Algebras

Having determined the Nakayama automorphism, it is now a simple matter to characterize symmetric Hopf algebras. Recall from §2.2.5 that $H$ is symmetric if and only if the Nakayama automorphism is inner. The theorem below is originally due to Oberst and Schneider [160] as is the material in §§12.1.3 and 12.1.4.

**Theorem 12.9.** A finite-dimensional Hopf algebra $H$ is symmetric if and only if $H$ is unimodular and $S^2$ is an inner automorphism of $H$.

**Proof.** For unimodular $H$, we have $\alpha = \varepsilon$ by (12.5) and so Proposition 12.8 gives $\nu = S^2 * \alpha = S^2$. If, in addition, $S^2$ is inner, then so is $\nu$ and hence $H$ is symmetric.

Conversely, assume that $\nu$ is inner. Then $\varepsilon \circ \nu = \varepsilon$. Since $\varepsilon \circ S = \varepsilon$ and $\varepsilon \circ \alpha = \alpha$, we obtain
\[
\varepsilon = \varepsilon \circ \nu = \varepsilon \circ (S^2 * \alpha) \overset{\text{Prop. 12.8}}{=} (\varepsilon \circ S^2) * (\varepsilon \circ \alpha) = \varepsilon * \alpha = \alpha
\]
Thus, $H$ is unimodular by (12.5) and $S^2 = \nu$ is inner. \qed
Exercises for Section 12.1

12.1.1 (The actions → and ←). Show that the actions \( \langle h \rightarrow f \rightarrow k, l \rangle = \langle f, klh \rangle \) \((h, k, l \in H, f \in H^*)\) can be written as follows: \( h \rightarrow f = f_{(1)} \langle f_{(2)}, h \rangle \) and \( f \rightarrow h = \langle f_{(1)}, h \rangle f_{(2)} \).

12.1.2 (Integrals). Let \( H \) be a finite-dimensional Hopf algebra. Show that \( \langle \lambda, \Lambda \rangle \neq 0 \) for any nonzero left or right integral \( \lambda \in H^* \) and any nonzero left or right integral \( \Lambda \in H \).

\[ \text{[Hint: Use } \lambda \text{ as Frobenius form. Then, after a possible renormalization, we have } \lambda \rightarrow \lambda = \varepsilon \text{ if } \lambda \text{ is a left integral and } \lambda \rightarrow \lambda = \varepsilon \text{ if } \lambda \text{ is a right integral; see (12.12).]} \]

12.1.3 (Integrals of the dual Taft algebra). Continuing with the notation of Example 12.2, let \( H = H_{n,q} \) and consider the dual Hopf algebra \( H^* \). Let \( ((g^i x^j)^*)_{0 \leq i, j < n} \) be the dual \( k \)-basis of \( H^* \) for the basis \( (g^i x^j)_{0 \leq i, j < n} \) of \( H \). Then

\[ x^* \in f_{H^*}^r \quad \text{and} \quad (g^{n-1} x^*)^* \in f_{H^*}^l. \]

Use the quantum binomial formula (Exercise 9.3.14) to verify (12.2) with \( \lambda = x^* \) and (12.3) with \( \lambda' = (g^{n-1} x)^* \). Is \( H \) self-dual?

12.1.4 (Integrals and trace forms). Show that the following are equivalent for a finite-dimensional Hopf algebra \( H \):

(i) \( H \) is unimodular and involutory;

(ii) \( f_{H^*}^l \subseteq H_{\text{trace}}^e \);

(iii) \( f_{H^*}^r \subseteq H_{\text{trace}}^e \).

\[ \text{Can this be fixed?} \]
[Hints: If (i) holds, then the Nakayama automorphism that is associated to a left integral $0 \neq \lambda \in H^*$ is given by $\nu = \text{Id}_H$ (Proposition 12.8), which in turn states exactly that $\lambda \in H^\text{trace}$. Thus, (i) implies (ii). Conversely, assume (ii) and let $0 \neq \lambda \in H^*$. Since $\lambda \in H^\text{trace}$, the Nakayama automorphism for the Frobenius form $\lambda$ is $\nu = \text{Id}_H$. Thus $H$ is symmetric, and Theorem 12.9 tells us that $H$ is unimodular. Proposition 12.8 now gives $\nu = S^2$. Hence $S^2 = \text{Id}_H$, proving (i). Finally, $H^\text{trace}$ is clearly stable under $S^*$, and so (ii) $\iff$ (iii) by (12.1).]

12.1.5 (The Morita context associated with an $H$-module algebra). This exercise assumes familiarity with Exercises 2.1.3 and 2.1.4. Let $H \in \text{HopfAlg}_k$ be finite dimensional and let $A \in H\text{Alg}$. Define $\text{trace}$. Put in stuff from Montgomery §4.5 (uses $\alpha$) and some consequences as in Montgomery §4.3 an 4.4.

12.2. The Antipode

In this section, we prove an important theorem due to Radford [171] showing that the antipode $S$ of any finite-dimensional Hopf algebra $H$ has finite order with respect to composition. We also determine the trace of $S^2 \in \text{Aut}_{\text{Alg}}(H)$; this trace will play a crucial role in our discussion of semisimplicity in the next this section.

12.2.1. Order of the Antipode

Recall that any group-like element $g \in G(H)$ has order at most $|G(H)| \leq \dim_k H$. The adjoint operator $\text{ad} g \in \text{Aut}_{\text{Alg}}(H)$ also has finite order dividing the order of $g$. Similar remarks apply to $\gamma \in G(H^*)$ and to the adjoint operator $\text{ad} \gamma \in \text{Aut}_{\text{Alg}}(H^*)$ and its transpose $(\text{ad} \gamma)^* \in \text{Aut}_{\text{Coalg}}(H)$.

The formula for $S^4$ in following theorem is often referred to as Radford’s formula.

Theorem 12.10. Let $H$ be a finite-dimensional Hopf algebra. Then

$$S^4 = (\text{ad} a) \circ (\text{ad} a)^* = (\text{ad} a)^* \circ (\text{ad} a)$$

where $a \in G(H)$ and $\alpha \in G(H^*)$ are the right modular elements. In particular, the order of $S^4$ divides the least common multiple of the orders of $a$ and $\alpha$.

Proof. We start with some general observations, valid for any $g \in G(H)$ and $\gamma \in G(H^*)$. First, $\langle f, (\text{ad} \gamma)^*(h) \rangle = \langle \gamma f \gamma^{-1}, h \rangle = \langle \gamma, h_{(1)} \rangle \langle f, h_{(2)} \rangle \langle \gamma^{-1}, h_{(3)} \rangle$ for $f \in H^*$ and $h \in H$. Hence,

$$\text{(12.15)} \quad (\text{ad} \gamma)^*(h) = \langle \gamma, h_{(1)} \rangle h_{(2)} \langle \gamma^{-1}, h_{(3)} \rangle \quad (h \in H)$$
In particular, since \( \gamma : H \to \mathbb{k} \) is an algebra homomorphism,
\[
((\text{ad} \gamma)^* \circ (\text{ad} g))(h) = \langle \gamma, gh_{(1)}g^{-1} \rangle gh_{(2)}g^{-1} \langle \gamma^{-1}, gh_{(3)}g^{-1} \rangle
\]
\[
= \langle \gamma, h_{(1)} \rangle gh_{(2)}g^{-1} \langle \gamma^{-1}, h_{(3)} \rangle
\]
\[
= ((\text{ad} g) \circ (\text{ad} \gamma)^*)(h)
\]
The second equality in the formula for \( S^4 \) is a special case of this.

Next, observe that (12.15) amounts to the identity \( (\text{ad} \gamma)^* = \gamma \ast \text{Id}_H \ast \gamma^{-1} \) in the convolution algebra \( \text{End}_k(H) \). Furthermore, for any \( \sigma \in \text{End}_{\text{Alg}}(H) \), the operator \( \sigma_\ast = \sigma \circ \ast \) an algebra endomorphism of \( \text{End}_k(H) \) that is the identity on \( H^* \subseteq \text{End}_k(H) \) (Exercise 9.2.1). In particular, \( \sigma_\ast(\gamma \ast \text{Id}_H \ast \gamma^{-1}) = \gamma \ast \sigma_\ast(\text{Id}_H) \ast \gamma^{-1} = \gamma \ast \sigma \ast \gamma^{-1} \) and so (12.15) gives
\[
\sigma \circ (\text{ad} \gamma)^* = \gamma \ast \sigma \ast \gamma^{-1}
\]
Now Proposition 12.8 provides us with the equality \( S^2 \ast \alpha = \alpha \ast (S^2 \circ \text{ad} a) \) in \( \text{End}_k(H) \) or, equivalently,
\[
S^2 = \alpha \ast (S^2 \circ \text{ad} a) \ast \alpha^{-1}
\]
By (12.16) with \( \sigma = S^2 \circ \text{ad} a \) and \( \gamma = \alpha \), this becomes \( S^2 = S^2 \circ \text{ad} a \circ (\text{ad} \alpha)^* \), which yields the formula for \( S^4 \). It follows that \( S^4^k = (\text{ad} a)^k \circ ((\text{ad} \alpha)^* )^k = (\text{ad} a^k)^\ast \circ (\text{ad} a^k)^\ast \) for all \( k \), where the powers \( a^k \) and \( \alpha^k \) are computed in \( H \) and \( H^* \), respectively, whereas all other powers are formed with composition. The statement about the order of \( S^4 \) follows from this.

12.2.2. The Trace of \( S^2 \)

For an arbitrary Frobenius algebra \((A, \lambda)\) with Casimir element \( c_{\lambda^2} = y_i \otimes x_i \), we have seen earlier that \( \text{trace}(f) = \langle \lambda, f(x_i) y_i \rangle \) holds for any \( f \in \text{End}_k(A) \) (Lemma 2.14). In the present setting, fixing \( 0 \neq \lambda \in f^H \), as Frobenius form, the Casimir element is \( c_{\lambda^2} = S(\Lambda_{(1)}) \otimes \Lambda_{(2)} \) (Proposition 12.6). Thus, the trace formula becomes
\[
\text{trace}(f) = \langle \lambda, f(\Lambda_{(2)})S(\Lambda_{(1)}) \rangle \quad (f \in \text{End}_k(H))
\]
To evaluate this formula for \( f = S^2 \), note that \( S^2(\Lambda_{(2)})S(\Lambda_{(1)}) = S(\Lambda_{(1)})S(\Lambda_{(2)}) = \langle e, \Lambda \rangle 1 \). Thus, (12.17) gives
\[
\text{trace}(S^2) = \langle \lambda, S^2(\Lambda_{(2)})S(\Lambda_{(1)}) \rangle = \langle e, \Lambda \rangle \langle \lambda, 1 \rangle
\]
Our goal is to further analyze this expression. To this end, we identify \( H \) with \( H^{**} \) by means of the canonical isomorphism and we define
\( x \in H \)
to be the character of the left or right regular representation of the algebra \( H^* \) — the characters of the two regular characters agree by (2.30). Self-duality of the
regular representation (Theorem 10.9) and the isomorphism \( H_{\text{reg}} \otimes H_{\text{reg}} \cong H_{\text{reg}}^{\dim H} \) (Corollary 10.5), applied with \( H^* \) in place of \( H \), give the equations

\[(12.19) \quad Sx = x \quad \text{and} \quad x^2 = \dim_k H \cdot x\]

In particular, the automorphism \( S^2 \) of \( H \) stabilizes \( xH \) and so we may consider \( S^2|_{xH} \in \text{Aut}_k(xH) \).

**Proposition 12.11.** Let \( H \) be a finite-dimensional Hopf algebra, let \( 0 \neq \lambda \in \mathcal{f}_H^l \), and let \( \Lambda \in H \) be determined by (12.12). Then

\[\text{trace}(S^2) = \langle \varepsilon, \Lambda \rangle \langle \lambda, 1 \rangle = \dim_k H \cdot \text{trace}(S^2|_{xH})\]

where \( x \in H \) is the character of the regular representation of \( H^* \).

**Proof.** In view of (12.18), it remains to prove the second equality. We first make a general observation that is valid for any finite-dimensional algebra \( A \). Namely, suppose that \( a \in A \) satisfies \( a^2 = \xi a \) with \( \xi \in \mathbb{K} \) and that \( f \in \text{End}_k(A) \) satisfies \( f(aA) \subseteq aA \). Then \( f \circ a_A \in \text{End}_k(A) \) maps \( A \) to \( aA \), and hence \( \text{trace}(f \circ a_A) = \text{trace}(f \circ a_A|_{aA}) \). Since \( f \circ a_A|_{aA} = \xi f|_{aA} \), we obtain that \( \xi \text{trace}(f|_{aA}) = \text{trace}(f \circ a_A) \). By (12.19), all this applies with \( A = H, a = x, f = S^2 \) and \( \xi = \dim_k H \). Therefore,

\[\dim_k H \cdot \text{trace}(S^2|_{xH}) = \text{trace}(S^2 \circ x_H) = \langle \lambda, S^2(x\Lambda(2))S(\Lambda(1)) \rangle = \langle \lambda, xS(\Lambda(1))S(\Lambda(2)) \rangle = \langle \lambda, xS(\Lambda(1))S(\Lambda(2)) \rangle = \langle \varepsilon, \Lambda \rangle \langle \lambda, 1 \rangle\]

Finally, since the right regular action of the left integral \( \lambda \) on \( H^* \) has image \( \mathbb{K}\lambda \) and \( \lambda^2 = \langle \lambda, 1 \rangle \lambda \), it follows from the definition of \( x \) that \( \langle \lambda, x \rangle = \text{trace}(H^* \lambda) = \langle \lambda, 1 \rangle \), proving the second equality in the proposition. \( \square \)

**Exercises for Section 12.2**

to be added

12.3. Semisimplicity

The main goal of this section is to develop criteria for the semisimplicity of a Hopf algebra \( H \). Here, \( H \) is called semisimple if the underlying algebra of \( H \) is semisimple in the sense of §1.4.4. We are also interested in semisimplicity of the dual Hopf algebra \( H^* \) and in the structure of the representation algebra of a semisimple Hopf algebra.
12.3.1. The Generalized Maschke Theorem

The basic result characterizing semisimplicity of Hopf algebras is a direct generalization of Maschke’s Theorem (§3.4.1) for group algebras. Indeed, if \( G \) is a finite group, then \( f_H^1 = \mathbb{k}\lambda \) with \( \lambda = \sum_{g \in G} g \) and \( \langle \varepsilon, \lambda \rangle = |G|/1 \) (Example 12.1). Hence the condition \( \text{char } k \nmid |G| \) in the original version of Maschke’s Theorem is equivalent to \( \langle \varepsilon, f_H^1 \rangle \neq 0 \). The core of the proof in the general case is identical to the proof of Maschke’s Theorem for group algebras, but we give the argument again here.

Maschke’s Theorem for Hopf Algebras. A Hopf algebra \( H \) is semisimple if and only if \( \langle \varepsilon, f_H^1 \rangle \neq 0 \) or, equivalently, \( \langle \varepsilon, f_H^r \rangle \neq 0 \). In particular, \( H \) is finite-dimensional and unimodular in this case.

Proof. First, observe that \( \langle \varepsilon, f_H^1 \rangle = \langle \varepsilon, f_H^r \rangle \) for any Hopf algebra \( H \). For, if \( H \) is infinite dimensional, then \( f_H^1 = f_H^r = 0 \) by Proposition 10.6, and if \( H \) is finite-dimensional, then \( f_H^1 = S(f_H^r) \) by (12.1).

Now assume that \( H \) is semisimple. Then the counit \( \varepsilon : H_{\text{reg}} \to \mathbb{1} \) must be nonzero on the space of invariants \( H_{\text{reg}}^H = f_H^1 \), because \( f_H^1 \) is the \( 1 \)-homogeneous component of the regular representation \( H_{\text{reg}} \). Thus, \( \langle \varepsilon, f_H^1 \rangle \neq 0 \).

Conversely, assume that \( \langle \varepsilon, f_H^1 \rangle \neq 0 \) for some \( t \in f_H^1 \) and put \( \Lambda = \langle \varepsilon, t \rangle^{-1} t \in f_H^1 \). Then \( \langle \varepsilon, \Lambda \rangle = 1 \) and, exactly as in the proof of Proposition ?? for group algebras, \( M^H = \Lambda M \) for every \( M \in \text{Rep } H \); in fact, \( \Lambda M \) is a projection of \( M \) onto \( M^H \). Taking \( M = \text{Hom}_k(V, U) \) for \( V, U \in \text{Rep } H \), we obtain \( \Lambda \cdot \text{Hom}_k(V, U) = \text{Hom}_k(V, U) \) by (10.9). We will show that, if \( U \) is a subrepresentation of \( V \), then there is a map \( \pi \in \text{Hom}_H(V, U) \) with \( \pi|_U = \text{Id}_U \). Then \( \text{Ker } \pi \) will be a complement for \( U \) in \( V \), whence \( V \) is completely reducible as desired. To construct \( \pi \), choose any \( k \)-linear projection \( \pi_0 \in \text{Hom}_k(V, U) \); so \( \pi_0|_U = \text{Id}_U \). Then \( \pi := \Lambda \cdot \pi_0 \in \text{Hom}_H(V, U) \). Furthermore, for \( u \in U \),

\[
\pi(u) = \Lambda(1) \cdot \pi_0(S(\Lambda(2)).u) = \Lambda(1) \cdot S(\Lambda(2)).u = \langle \varepsilon, \Lambda \rangle u = u
\]

because each \( S(\Lambda(2)).u \in U \) and so \( \pi_0(S(\Lambda(2)).u) = S(\Lambda(2)).u \). This proves semisimplicity of \( H \).

Finally, for any \( h \in H \), we have \( \Lambda h = \langle \alpha', h \rangle \Lambda \), where \( \alpha' \in G(H^r) \) is the left modular element, and hence \( \langle \varepsilon, h \rangle = \langle \varepsilon, \Lambda \rangle \langle \varepsilon, h \rangle = \langle \varepsilon, \Lambda h \rangle = \langle \alpha', h \rangle \langle \varepsilon, \Lambda \rangle = \langle \alpha', h \rangle \), proving unimodularity. This completes the proof of the theorem. \( \square \)

Corollary 12.12. If \( H \) is semisimple then so is \( H \otimes K \) for every field extension \( K/k \).

Proof. The condition \( \langle \varepsilon, f_H^1 \rangle \neq 0 \) is evidently preserved under extensions of the base field \( k \). \( \square \)
Finite-dimensional algebras satisfying the conclusion of Corollary 12.12 are called **separable**; see Exercises 1.4.10 and 1.5.6 for more on separable algebras.

We will call a Hopf algebra $H$ **bi-semisimple** if $H$ and the dual algebra $H^*$ are both semisimple algebras. By Maschke’s Theorem for Hopf algebras, we know that $H$ must be finite dimensional in this case; so $H^* = H^*$ is in fact a Hopf algebra. Applying Maschke’s Theorem to $H^*$, we see that $H^*$ is semisimple if and only if $\langle \lambda, 1 \rangle \neq 0$ for some left or right integral $\lambda \in H^*$.

**Corollary 12.13.** A Hopf algebra $H$ is bi-semisimple if and only if $H$ is finite-dimensional and trace$(S^2) \neq 0$. In this case, char $\mathbb{k} \nmid \dim_{\mathbb{k}} H$.

**Proof.** By Proposition 12.11 or (12.18), we have trace$(S^2) = \langle \varepsilon, \Lambda \rangle \langle \lambda, 1 \rangle$; this is nonzero if and only if $\langle \varepsilon, \Lambda \rangle$ and $\langle \lambda, 1 \rangle$ are both nonzero. By Maschke’s Theorem for Hopf algebras, the former says that $H$ is semisimple and the latter that $H^*$ is semisimple. The last assertion, that char $\mathbb{k} \nmid \dim_{\mathbb{k}} H$, is a direct consequence of the second equality in Proposition 12.11. \qed

12.3.2. **Involutory Semisimple Hopf Algebras**

Corollary 12.13 implies the following fact, proved earlier in Corollary 10.17: any finite-dimensional involutory Hopf algebra $H$ such that char $\mathbb{k} \nmid \dim_{\mathbb{k}} H$ is semisimple. Of course, the condition on the characteristic is automatic if char $\mathbb{k} = 0$; so all finite-dimensional involutory Hopf algebras are semisimple in this case. The converse also holds: semisimple Hopf algebras over a field of characteristic 0 are involutory. This important result is due to Larson and Radford [128], [129], which we shall prove next. We remark that, according to a famous conjecture of Kaplansky [116, Appendix 2, Conjecture #5], all semisimple Hopf algebras over any base field are expected to be involutory. This conjecture remains open. However, Etingof and Gelaki [67] have shown that bi-semisimple Hopf algebras are always involutory. Thus, Corollary 12.13 can be improved: a Hopf algebra $H$ is bi-semisimple if and only if $H$ is finite-dimensional involutory and char $\mathbb{k} \nmid \dim_{\mathbb{k}} H$. The proof in [67] uses the theorem of Larson and Radford, to be proved below, and in addition a “lifting” result from positive characteristics to characteristic 0, which is outside the scope of this book.

We start with a lemma on automorphisms of an arbitrary finite-dimensional semisimple algebra $A$ over a field $\mathbb{k}$ with char $\mathbb{k} = 0$. By Wedderburn’s Structure Theorem, $A$ is isomorphic to a direct product of matrix algebras over division $\mathbb{k}$-algebras. The images of these matrix algebras in $A$ are the minimal nonzero ideals of $A$; they are called the simple components of $A$. An element $\alpha \in \mathbb{k}$ is called totally non-negative if $\alpha$ is algebraic over $\mathbb{Q} \subseteq \mathbb{k}$ and the image of $\alpha$ under each embedding $\mathbb{Q}(\alpha) \hookrightarrow \mathbb{C}$ belongs to $\mathbb{R}_+$. 
Lemma 12.14. Let \( A \) be a finite-dimensional semisimple \( k \)-algebra, where \( \mathrm{char}\, k = 0 \), and let \( \phi \in \text{Aut}_{\text{Alg}}(A) \) have finite order. Then \( \text{trace}(\phi) \) is totally non-negative and \( \text{trace}(\phi) \neq 0 \) if and only if there is a simple component \( B \) of \( A \) such that \( \phi(B) = B \) and \( \text{trace}(\phi|_B) \neq 0 \).

Proof. It suffices to show that \( \text{trace}(\phi) \) is totally non-negative; the non-vanishing assertion will then be a simple consequence. Indeed, \( \phi \) permutes the simple components of \( A \). If \( A_1, \ldots, A_n \) are the components that are fixed by \( \phi \), then

\[
\text{(12.20)} \quad \text{trace}(\phi) = \sum_{i=1}^{n} \text{trace}(\phi_i) \quad \text{with} \quad \phi_i = \phi|_{A_i},
\]

Since each \( \phi_i \) is an automorphism of \( A_i \) having finite order, all elements above are totally non-negative by assumption. Consequently, \( \text{trace}(\phi) = 0 \) if and only if all \( \text{trace}(\phi_i) = 0 \).

In order to show that \( \text{trace}(\phi) \) is totally non-negative, we may replace \( A \) by \( A \otimes \overline{k} \) and \( \phi \) by \( \phi \otimes \overline{k} \), where \( \overline{k} \) is an algebraic closure of \( k \), thereby reducing to a case where \( k \) is algebraically closed. (Here, we use the fact that \( A \otimes \overline{k} \) is semisimple, because \( \overline{k} \) is perfect; see Exercises 1.4.10 and 1.5.6. For a Hopf algebra \( A \), we may invoke Corollary 12.12 instead.) Next, decomposing \( A \) into its simple components and using (12.20), we may further reduce to the case where \( A \) is simple. So \( A \) is now a matrix algebra over the algebraically closed field \( \overline{k} \). By the Noether-Skolem Theorem (e.g., [167, 12.6]), \( \phi \) is an inner automorphism of \( A \), that is, \( \phi = g(\cdot)g^{-1} \) for some unit \( g \in A^\times \). Since \( \phi^t = \text{Id}_A \) for some positive integer \( t \), we must have \( g^t \in A(A)^\times \cong \overline{k}^\times \). Writing \( g^t = \lambda^t 1_A \) for some \( \lambda \in \overline{k}^\times \) and replacing \( g \) by \( g^{\lambda^{-1}} \), we may assume that \( g^t = 1 \). We identify \( A \) with \( \text{End}_\overline{k}(S) \cong S \otimes S^* \), where \( S = \overline{k}^{\text{gen}} \).

Then \( g \in \text{GL}(S) \) and \( \phi = g(\cdot)g^{-1} \) becomes \( g \otimes (g^{-1})^* : S \otimes S^* \to S \otimes S^* \). Hence, by (B.25) and (B.26),

\[
\text{trace}(\phi) = \text{trace}(g)\text{trace}(g^{-1})
\]

Let \( \zeta \in \overline{k} \) be a primitive \( t \)-th root of unity. Then \( \text{trace}(\phi) \in \mathbb{Q}(\zeta) \subseteq \overline{k} \) and so \( \text{trace}(\phi) \) is certainly algebraic over \( \mathbb{Q} \). It suffices to show that the image of \( \text{trace}(\phi) \) under each embedding \( \mathbb{Q}(\zeta) \hookrightarrow \mathbb{C} \) belongs to \( \mathbb{R}_+ \). The image of \( \zeta \) under such an embedding is a primitive complex \( t \)-th root of unity and the automorphism \( - \) of \( \mathbb{Q}(\zeta) \) that is given by \( \overline{\zeta} = -\zeta^{-1} \) corresponds to complex conjugation in \( \mathbb{C} \). Therefore, the subfield \( K := \mathbb{Q}(\zeta + \zeta^{-1}) \subseteq \mathbb{Q}(\zeta) \) embeds into \( \mathbb{R} \) and, for any \( \alpha \in \mathbb{Q}(\zeta) \), the element \( \alpha \overline{\alpha} \in K \) is being sent to \( \mathbb{R}_+ \). The above formula shows that \( \text{trace}(\phi) = \alpha \overline{\alpha} \) with \( \alpha = \text{trace}(g) \in \mathbb{Q}(\zeta) \), which has the desired form.

We are now ready to prove the theorem of Larson and Radford.

Theorem 12.15. Let \( H \) be a finite-dimensional Hopf \( k \)-algebra and assume that \( \mathrm{char}\, k = 0 \). Then the following are equivalent:

(i) \( H \) is semisimple;
12.3. Semisimplicity

(ii) $H^*$ is semisimple;
(iii) $H$ is involutory: $S^2 = \text{Id}_H$.

**Proof.** In view of Corollary 12.13, (iii) implies both (i) and (ii), and the implication (i) $\Rightarrow$ (ii) amounts to the assertion that trace($S^2$) $\neq 0$ automatically holds for semisimple $H$. But $S^2$ is an algebra automorphism of $H$ having finite order (Theorem 12.10). The space of (right and left) integrals $f_H$ is a simple component of $H$ that is isomorphic to $\mathbb{k}$ via $\varepsilon$ by Maschke’s Theorem for Hopf algebras. Since $S(f_H) = f_H$ by (12.1) and $\varepsilon \circ S = \varepsilon$, it follows that $S$ is the identity on $f_H$. Therefore, $S^2$ has trace 1 on $f_H$ and Lemma 12.14 gives the desired conclusion, trace($S^2$) $\neq 0$.

The reverse implication (ii) $\Rightarrow$ (i) follows by interchanging $H$ and $H^*$. Thus, (i) and (ii) are equivalent; it remains to show that these conditions imply (iii).

Assume that $H$ and $H^*$ are semisimple; so trace($S^2$) $\neq 0$ by Corollary 12.13. Since $H$ and $H^*$ are both unimodular, we know from Theorem 12.10 that $S^4 = \text{Id}_H$. Therefore, $S^2$ has eigenvalues $\pm 1$ and so trace($S^2$) is a nonzero integer. In fact, trace($S^2$) is positive by Lemma 12.14; so $0 < \text{trace}(S^2) \leq \dim_k H$. On the other hand, the formula trace($S^2$) $= \dim_k H \cdot \text{trace}((S^2)_{xH})$ in Proposition 12.11 shows that trace($S^2$) is an integer multiple of $\dim_k H$. Therefore, we must have trace($S^2$) $= \dim_k H$, whence $S^2 = \text{Id}_H$ as desired. \hfill $\Box$

We close this subsection with some facts concerning involutory Hopf algebras in arbitrary characteristics.

**Proposition 12.16.** Let $H$ be a finite-dimensional Hopf algebra, let $0 \neq \lambda \in f_H^*$, and let $\Lambda \in H$ be determined by (12.12). Then the following are equivalent:

(i) $\chi_{\text{reg}} = \langle \varepsilon, \Lambda \rangle \lambda$
(ii) $\chi_{\text{reg}} \in f_H^*$
(iii) $\chi_{\text{reg}} \in f_H^{*}$

Furthermore, (i) – (iii) are satisfied if $H$ is involutory.

**Proof.** Clearly, (i) implies (ii). For the converse, we use formula (12.17), which gives the expression $\chi_{\text{reg}}(h) = \langle \lambda, h \Lambda(2) S(\Lambda(1)) \rangle$ or, equivalently,

$$\chi_{\text{reg}} = \Lambda(2) S(\Lambda(1)) \lambda$$

Since $f_H^* = \mathbb{k} \lambda$ and $\lambda$ is an isomorphism $H \cong H^*$ by Proposition 12.5, condition (ii) is equivalent to $\Lambda(2) S(\Lambda(1)) \in \mathbb{k} 1_H$, which in turn is equivalent to

$$\Lambda(2) S(\Lambda(1)) \cdot \langle \varepsilon, (\Lambda(2) S(\Lambda(1))) 1_H = \langle \varepsilon, S(\Lambda(1)) \Lambda(2) \rangle 1_H = \langle \varepsilon, \Lambda \rangle 1_H$$

This shows that (ii) implies (i). Since the regular representation of $H$ is self-dual (Theorem 10.9), we have $S^* \chi_{\text{reg}} = \chi_{\text{reg}}$ by (10.21). Hence, in view of (12.1), conditions (ii) and (iii) are equivalent to each other. This proves the equivalence of (i) – (iii). Finally, if $S^2 = \text{Id}_H$, then $\Lambda(2) S(\Lambda(1)) = S(\Lambda(1) S(\Lambda(2))) = \langle \varepsilon, \Lambda \rangle 1_H$ and hence (i) holds. \hfill $\Box$
Corollary 12.17. If $H$ is semisimple and involutory, then $H^*$ is unimodular and $\mathbb{k}_{H^*} = \mathbb{k}_{\chi_{\text{reg}}}$.

Proof. Proposition 12.16 gives $\chi_{\text{reg}} = \langle \varepsilon, \Lambda \rangle \in \mathbb{f}_H^* \cap \mathbb{f}_{H^*}^*$ and Maschke’s Theorem for Hopf algebras further implies that $\langle \varepsilon, \Lambda \rangle \neq 0$. Hence $\chi_{\text{reg}} \neq 0$ and the corollary follows. □

12.3.3. The Representation Algebra

We now focus on a split semisimple Hopf algebra $H$. Thus $H$ is a symmetric algebra (Proposition 2.16) and $H^*$ is at least a Frobenius algebra, with any $0 \neq \Lambda \in \mathbb{f}_H$ serving as Frobenius form (Proposition 12.5). We shall generally choose $\Lambda$ so that $\langle \varepsilon, \Lambda \rangle = 1$. This is possible by Maschke’s Theorem for Hopf algebras (§12.3.1) and it determines $\Lambda$ uniquely. It turns out that the representation algebra $\mathcal{R}_k(H) = \mathcal{R}(H) \otimes_\mathbb{Z} \mathbb{k}$ is also Frobenius, in fact symmetric, and even semisimple when $\text{char} \mathbb{k} = 0$. In positive characteristics, $\mathcal{R}_k(H)$ need not be semisimple as the example $H = (\mathbb{k}G)^*$ for a finite group $G$ shows (Example 10.20).

Frobenius form. For $V$ in $\text{Rep}_\text{fin} H$, define

$$ (12.21) \quad \lambda([V]) \overset{\text{def}}{=} \dim_\mathbb{k} V^H = m(\mathbb{1}, V) $$

where $V^H$ denotes the space of $H$-invariants in $V$ and $m(\mathbb{1}, V)$ is the multiplicity of the trivial representation $\mathbb{1}$ in $V$ (§1.4.2). The functor $(\,,)^H$ is exact: any exact sequence $0 \to U \to V \to W \to 0$ in $\text{Rep} H$ splits, because $H$ is semisimple, and hence $0 \to U^H \to V^H \to W^H \to 0$ is also exact. Consequently, $\lambda$ yields a well-defined group homomorphism

$$ \lambda: \mathcal{R}(H) \to \mathbb{Z} $$

We let $\lambda_\mathbb{k}: \mathcal{R}_k(H) \to \mathbb{k}$ denote the $\mathbb{k}$-linear extension of $\lambda$ and we will also write $[V]_\mathbb{k} = [V] \otimes 1 \in \mathcal{R}_k(H)$. So

$$ \lambda_\mathbb{k}([V]_\mathbb{k}) = (\dim_\mathbb{k} V^H)1_\mathbb{k} $$

Proposition 12.18. Let $H$ be a split semisimple Hopf algebra. Then:

(a) $\lambda_\mathbb{k}$ is a $\ast$-invariant trace form on $\mathcal{R}_k(H)$ and $([S]_\mathbb{k}, [S^\ast]_\mathbb{k})_{S \in \text{Irr} H}$ are dual bases with respect to $\lambda_\mathbb{k}$. Therefore, $(\mathcal{R}_k(H), \lambda_\mathbb{k})$ is a symmetric algebra.

(b) With $\Lambda \in \mathbb{f}_H$ such that $\langle \varepsilon, \Lambda \rangle = 1$, the character map gives a monomorphism of Frobenius algebras $\chi_\mathbb{k}: (\mathcal{R}_k(H), \lambda_\mathbb{k}) \hookrightarrow (H^*, \Lambda)$; so $\Lambda \circ \chi_\mathbb{k} = \lambda_\mathbb{k}$.

(c) If $\text{char} \mathbb{k} = 0$, then $\mathcal{R}_k(H)$ is semisimple.
Proof. (a) Recall from §10.3.4 that $\ast$ yields an involution of $\mathcal{R}(H)$, since $S^2$ is an inner automorphism of $H$ by Theorem 12.9, and that $\ast$ permutes the standard $\mathbb{Z}$-basis $t([S])_{S \in \text{Irr } H}$ of $\mathcal{R}(H)$. For $S, T \in \text{Irr } H$, we have

$$\lambda([S][T^\ast]) = (\dim_k \text{Hom}_H(T, S) = \delta_{S,T}$$

where the last equality uses Schur’s Lemma and the fact that $k$ is a splitting field. Therefore, $([S]_k, [S^\ast]_k)_{S \in \text{Irr } H}$ are dual bases of $\mathcal{R}_k(H)$ with respect to $\lambda_k$ as in (2.21). Furthermore, for any $V \in \text{Rep}_{\text{fin}} H$, we have $m(I, V) = m(I, V^\ast)$ and so

$$\lambda([V]) = \lambda([V^\ast])$$

Thus $\lambda$ and $\lambda_k$ are $\ast$-invariant. Finally, for $S, T \in \text{Irr } H$, we have

$$\lambda([S][T]) = \delta_{S,T^\ast} = \lambda(T)\lambda([S])$$

showing that $\lambda_k$ is a trace form on $\mathcal{R}_k(H)$.

(b) The integral $\Lambda$ is an idempotent of $H$ such that $V^H = \Lambda V$ for every $V \in \text{Rep } H$. Therefore, $\Lambda(\chi_V) = \langle \chi_V, \Lambda \rangle = \text{trace}(\Lambda V) = (\dim_k V^H)1_k$ (Exercise 1.5.1), proving the equality $\Lambda \circ \chi_k = \lambda_k$.

(c) Semisimplicity of $\mathcal{R}_k(H)$ could be deduced from Theorem 2.21, but we prefer to give a more direct argument based on the following observation:

$$0 \neq x \in \mathcal{R}(H) \implies \lambda(x^\ast) > 0$$

To see this, write $x = [V] - [W]$ with $V, W \in \text{Rep}_{\text{fin}} H$ having no common irreducible constituents. Then $\text{Hom}_H(W, V) = 0$ and so $\lambda([V][W^\ast]) = 0 = \lambda([W][V^\ast])$. Therefore, $\lambda(x^\ast) = \dim_k \text{End}_H(V) + \dim_k \text{End}_H(W) > 0$ as claimed. It follows that the $\mathbb{Q}$-algebra $A := \mathcal{R}(H) \otimes_{\mathbb{Q}} \mathbb{Q}$ is semisimple. For, otherwise $N := \mathcal{R}(H) \cap \text{rad } A$ would be a nonzero nilpotent $\ast$-invariant ideal of $\mathcal{R}(H)$ and so $xx^\ast = 0$ for some $0 \neq x \in N$, contradicting our claim above. Thus, $A$ is semisimple – in fact, $A$ is separable (Exercise 1.5.6) – and so $\mathcal{R}_k(H) = A \otimes_k k$ is semisimple as well. □

Casimir element and trace. The Casimir element for the Frobenius trace form $\lambda_k$ in Proposition 12.18 is

$$c_{k} = \sum_{S \in \text{Irr } H} [S]_k \otimes [S^\ast]_k \in \mathcal{R}_k(H) \otimes \mathcal{R}_k(H)$$

We also have the Casimir trace $\gamma_{\lambda_k} : \mathcal{R}_k(H) \to \mathcal{Z} \mathcal{R}_k(H)$; see (2.33). Its value at 1 is given by the class of the adjoint representation:

$$\gamma_{\lambda_k}(1) = \sum_{S \in \text{Irr } H} [S]_k[S^\ast]_k \text{Prop. } 10.30 \text{ } [H_{\text{ad}}]_k \in \mathcal{Z} \mathcal{R}_k(H)$$
Exercises for Section 12.3

Throughout, \( H \) denotes a Hopf algebra.

12.3.1 (Traces of automorphisms). A char \( p \) version of Lemma 12.14 as in Section 10.1 of “HopfMaterials.”

12.3.2 (Symmetry of \( H^* \)). Show that if \( H \) is involutory and semisimple, then \( H^* \) is a symmetric algebra.

12.3.3 (Separability of Hopf algebras). Let \( H \) be a Hopf algebra such that \( \langle \varepsilon, f^*_H \rangle \neq 0 \). Deduce separability of \( H \) from Theorem 2.21 in conjunction with Propositions 12.5 and 12.6.

12.4. Divisibility Theorems

With the exception of the Nichols-Zoeller Theorem (§12.4.5), the results in this section are all modeled on the Frobenius’ Divisibility Theorem for finite group algebras (§3.6.1). Thus, our focus will be on a semisimple Hopf algebra \( H \) over a field \( \mathbb{k} \) of characteristic 0. A celebrated conjecture of Kaplansky, which remains open, states that Frobenius’ Divisibility Theorem holds for \( H \), in the following form: the degree of every absolutely irreducible representation of \( H \) divides \( \dim_\mathbb{k} H \); see [116, Appendix 2, Conjecture #6].

We begin with some reminders. First, by Theorem 12.15, semisimplicity of \( H \) amounts to \( H \) being involutory, which then holds for \( H^* \) as well. Moreover, both \( H \) and \( H^* \) are unimodular by Maschke’s Theorem for Hopf algebras (§12.3.1) and \( \langle \varepsilon, \Lambda \rangle \neq 0 \) for any \( 0 \neq \Lambda \in f_H \). By Proposition 12.16, we further know that

\[
\lambda := \langle \varepsilon, \Lambda \rangle^{-1} \chi_{\text{reg}} \in H^*_\text{trace}
\]

is a nonzero integral of \( H^* \) and that \( \Lambda \) is related to \( \lambda \) as in (12.12): \( \langle \lambda, \Lambda \rangle = 1 \).

Using \( \lambda \) as Frobenius form for \( H \), the associated Casimir element is symmetric by (2.32), that is, \( c_{\lambda\lambda} = c_{\lambda\lambda} \). Denoting this element by \( c_\lambda \), Proposition 12.6 gives the expressions

\[
c_\lambda = S(\Lambda_{(1)}) \otimes \Lambda_{(2)} = \Lambda_{(2)} \otimes S(\Lambda_{(1)})
\]

\[
= S(\Lambda_{(2)}) \otimes \Lambda_{(1)} = \Lambda_{(1)} \otimes S(\Lambda_{(2)})
\]

The Casimir trace \( \gamma_\lambda : H \to \mathbb{Z}^{\otimes H} \) is given by \( \gamma_\lambda(h) = S(\Lambda_{(1)})h\Lambda_{(2)} \) for \( h \in H \); see (2.33). So, identifying \( \mathbb{k} \) with its image in \( H \) under the unit map,

\[
\gamma_{\lambda^*}(1) = \langle \varepsilon, \Lambda \rangle \in \mathbb{k}
\]

Reversing the roles of \( H \) and \( H^* \) in (12.26) and using \( \Lambda \) as the Frobenius form for \( H^* \), we obtain the value of the Casimir trace \( \gamma_\Lambda \) at \( \varepsilon = 1_{H^*} \):

\[
\gamma_\Lambda(\varepsilon) = \langle \lambda, 1 \rangle = \langle \varepsilon, \Lambda \rangle^{-1} \dim_\mathbb{k} H \in \mathbb{k}
\]
The Casimir values at the identity featured prominently in Theorem 2.17 and its corollaries; these results will play a crucial role in this section. Throughout this section, we will assume familiarity with the basic facts about integrality from §2.2.7.

12.4.1. Frobenius Divisibility for Hopf Algebras

We first offer an extension, due to Cuadra and Meir [49, Theorem 3.4], of Frobenius’ Divisibility Theorem to the context of Hopf algebras. The proof is identical to the one in §3.6.1 for finite group algebras, but it is repeated here.

**Theorem 12.19.** Let $H$ be a split semisimple Hopf algebra over a field $\mathbb{k}$ of characteristic 0 and let $\Lambda \in \mathcal{F}_H$ be such that $0 \neq \langle \varepsilon, \Lambda \rangle \in \mathbb{Z}$. Then the following are equivalent:

(i) The degree of every irreducible representation of $H$ divides $\langle \varepsilon, \Lambda \rangle$;

(ii) the Casimir element (12.25) is integral over $\mathbb{Z}$.

**Proof.** Choosing the Frobenius form for $H$ to be $\lambda$ as in (12.24), the Casimir element $c_{\lambda}$ is given by (12.25) and $\gamma(1) = \langle \varepsilon, \Lambda \rangle$ by (12.26). Thus, the theorem is a consequence of Corollary 2.18, which states that the degree of every irreducible representation of $H$ divides $\gamma(1)$ if and only if $c_{\lambda}$ is integral over $\mathbb{Z}$. □

12.4.2. Characters that are Central in $H^*$

As another application of Theorem 2.17, we now present an elegant generalization of Frobenius’ Divisibility Theorem due to S. Zhu [208, Theorem 8]. The hypothesis $\chi_S \in \mathcal{Z}(H^*)$ is automatic whenever $H$ is cocommutative, and hence it certainly holds for finite group algebras $H = kG$.

**Theorem 12.20.** Let $H$ be a semisimple Hopf algebra over a field $\mathbb{k}$ of characteristic 0 and let $S$ be an absolutely irreducible representation of $H$ such that $\chi_S \in \mathcal{Z}(H^*)$. Then $\dim_k S$ divides $\dim_k H$.

**Proof.** Corollary 12.12 allows us to assume that $\mathbb{k}$ is algebraically closed. Thus, $H$ and $H^*$ are both split semisimple. Choose $\Lambda \in \mathcal{F}_H$ such that $\langle \varepsilon, \Lambda \rangle = \dim_k H$; so $\Lambda$ is the character of the regular representation of $H^*$. Then, with $\lambda$ as in (12.24), we have $\gamma(1) = \dim_k H$ by (12.26) and $c_{\lambda}$ is given by (12.25). Thus, Theorem 2.17 gives the following formula for the primitive central idempotent $e(S) \in \mathcal{P}H$:

$$e(S) \frac{\dim_k H}{\dim_k S} = \langle \chi_S, \mathcal{S}(\Lambda(1))\Lambda(2) \rangle \text{ Exercise 12.1.1 } \chi_S - \Lambda$$

In the usual manner, it suffices to show that the element $\chi_S - \Lambda \in \mathcal{Z}(H^*)$ is integral over $\mathbb{Z}$. First, note that $\chi_S$ is integral over $\mathbb{Z}$, because the representation ring $\mathcal{R}(H)$ is a finitely generated $\mathbb{Z}$-module and so all its elements are integral over $\mathbb{Z}$. Furthermore, by hypothesis, $\chi_S \in \mathcal{Z}(H^*)$ and so

$$\chi_S \in \mathcal{Z}(H^*)^\text{int} := \{f \in \mathcal{Z}(H^*) \mid f \text{ is integral over } \mathbb{Z}\}$$
Thus, it suffices to show that all elements of $\mathcal{Z}(H^*)^{\text{int}}$ are integral over $\mathbb{Z}$. But $\mathcal{Z}(H^*) = \prod_{M \in \text{Irr } H} k e(M)$ and $\mathcal{Z}(H^*)^{\text{int}} = \prod_{M \in \text{Irr } H} O e(M)$, where $O$ denotes the integral closure of $\mathbb{Z}$ in $k$. Furthermore, $e(M) = \lambda = (\dim_k M) \chi_M$ by (2.36).

Therefore, $\mathcal{Z}(H^*)^{\text{int}} - \lambda \subseteq \chi(\mathcal{R}(H^*)) O$. Since the ring $\chi(\mathcal{R}(H^*)) O$ is a finitely generated $O$-module, all its elements are integral over $\mathbb{Z}$, completing the proof. □

### 12.4.3. The Class Equation

We now come to the celebrated class equation due to Kac [114, Theorem 2] and Y. Zhu [210, Theorem 1]. The proof given here is based on [136]. To set the stage, recall that the character map gives a monomorphism of Frobenius $k$-algebras $\chi_k : R_k(H) \hookrightarrow H^*$ (Proposition 12.18). Thus, for any $M \in \text{Rep } R_k(H)$, we may consider the induced module $\text{Ind}_{R_k(H)}^H M$. The 1-dimensional representation of $R_k(H)$ that is given by the dimension augmentation $R_k(H) \to k$, $[V] \mapsto (\dim_k V)1$ will be referred to as the trivial representation of the representation algebra and denoted by $1 = 1_{R_k(H)}$. By (10.22) the scalar multiples of the element $[H_{\text{reg}}]\in R_k(H)$ give a copy of $1_{R_k(H)}$ in $R_k(H)$.

**Theorem 12.21** (Class equation). Let $H$ be a semisimple Hopf algebra over an algebraically closed field $k$ of characteristic 0. Then:

(a) $\dim_k \text{Ind}_{R_k(H)}^H M$ divides $\dim_k H$ for every $M \in \text{Irr } R_k(H)$.

(b) $\dim_k H = 1 + \sum_{\lambda \in \text{Irr } R_k(H)} (\dim_k M)(\dim_k \text{Ind}_{R_k(H)}^H M)$.

**Proof.** Part (a) is an application of Corollary 2.19. To check hypotheses, recall from Proposition 12.18 that $R_k(H)$ is split semisimple; a Frobenius form is provided by $\lambda_k \in R_k(H)^{\text{trace}}$; and if $\Lambda \in \mathcal{I}_H$ is such that $\langle e, \Lambda \rangle = 1$, then the map of Frobenius algebras $\chi_k : (R_k(H), \lambda_k) \to (H^*, \Lambda)$ satisfies $\Lambda \circ \chi_k = \lambda_k$. Furthermore, $\gamma_\Lambda(e) = \dim_k H$ by (12.27) and by (12.22), the Casimir element $\epsilon_k \in R_k(H)^{\otimes 2}$ is the image of $\sum_{S \in \text{Irr } H} [S] \otimes [S^*] \in R(H)^{\otimes 2}$. Since all elements of $R(H)^{\otimes 2}$ are integral over $\mathbb{Z}$, it follows that $\epsilon_k$ is likewise. Therefore, Corollary 2.19 applies and yields that the fraction $\frac{\dim_k H}{\dim_k \text{Ind}_{R_k(H)}^H M}$ is integral over $\mathbb{Z}$, proving (a).

(b) Since $\chi_k([H_{\text{reg}}]) = \chi_{\text{reg}}$ is an integral of $H^*$ by Proposition 12.16, we have

$$\text{Ind}_{R_k(H)}^H \mathbb{1}_{R_k(H)} \cong \mathbb{1}_{H^*}.$$ 

Furthermore, $R_k(H)_{\text{reg}} \cong \bigoplus_{M \in \text{Irr } R_k(H)} M^{\otimes \dim_k M}$ by Proposition 12.18 and Wedderburn’s Structure Theorem. Therefore,

$$H^* \cong \mathbb{1}_{H^*} \oplus \bigoplus_{M \in \text{Irr } R_k(H)} \text{Ind}_{R_k(H)}^H M^{\otimes \dim_k M}$$

and (b) follows by taking dimensions. □
12.4. Divisibility Theorems

Applying part (a) above with \( H = (\mathbb{k}G)^* \) for a finite group \( G \), we once again obtain Frobenius’ Divisibility Theorem in its original form for finite group algebras, because \( \chi_k : \mathbb{S}_k(H) \rightarrow H^* = \mathbb{k}G \) is an isomorphism in this case (Example 10.20).

12.4.4. Some Applications of the Class Equation

The first theorem below, due to Y. Zhu [210, Theorem 2], had been conjectured by Kaplansky [116, Appendix 2, Conjecture #8]; it marks the first major contribution to the ongoing classification project of semisimple Hopf algebras in small or otherwise restricted dimensions.

**Theorem 12.22.** Let \( H \) be a Hopf algebra over an algebraically closed field \( \mathbb{k} \) with \( \text{char} \mathbb{k} = 0 \) and assume that \( \dim_k H = p \), a prime. Then \( H \cong \mathbb{k}C_p \), the group algebra of the cyclic group of order \( p \).

**Proof.** We already know that, for any Hopf algebra \( H \), the group algebra \( \mathbb{k}G(H) \) of the group of all group-like elements is a Hopf subalgebra of \( H \) (Lemma 9.1). If \( H \) is finite dimensional, then \( \dim_k H \) is in fact divisible by \( |G(H)| \). This is a consequence of the Nichols-Zoeller Theorem (§12.4.5) which will be proved very shortly. Granting this result for now and assuming the hypotheses of the theorem, it suffices to show that \( G(H) \neq 1 \); for, then we must have \( |G(H)| = \dim_k H = p \) and so \( G(H) \cong C_p \) and \( H = \mathbb{k}G(H) \cong \mathbb{k}C_p \). By the same token, if \( G(H^*) \neq 1 \) then we are able to conclude that \( H^* \cong \mathbb{k}C_p \). Thus, as a \( \mathbb{k} \)-algebra, \( H^* \cong \mathbb{k}[x]/(x^p - 1) \) and there are \( p \) algebra maps \( H^* \rightarrow \mathbb{k} \), given by sending \( x \) to each of the \( p \)th roots of unity in \( \mathbb{k} \). Consequently, \( G(H) = \text{Hom}_{\mathbb{Alg}_{\mathbb{k}}}(H^*, \mathbb{k}) \) has order \( p \) and we are done again.

Observe that \( H \) and \( H^* \) are both unimodular. For, if \( H \) is not unimodular, then the (right or left) modular element is a non-trivial element of \( G(H^*) \) and the foregoing implies that \( H \cong \mathbb{k}C_p \), which is unimodular. This contradiction shows that \( H \) must be unimodular, and the same holds for \( H^* \). Consequently, \( S^4 = \text{Id}_H \) by Radford’s formula (Theorem 12.10). We claim that, in fact, \( S^2 = \text{Id}_H \) or, which amounts to the same by Theorem 12.15, that \( H \) is semisimple. This is of course clear if \( p = 2 \), because \( H \) is then commutative and hence involutory (Exercise 9.3.7). So assume that \( p \neq 2 \) and put \( H_+ = \{ h \in H \mid S^2 h = h \} \). Then \( H = H_+ \oplus H_- \) and we must have \( \dim_k H_+ \neq \dim_k H_- \), because \( p \) is odd. Therefore, \( \text{trace}(S^2) \neq 0 \) and Corollary 12.13 gives the desired conclusion that \( H \) is semisimple.

To finish the proof, we invoke the class equation. Since \( \dim_k H = p \), we must have \( \dim_k \text{Ind}_{\mathbb{k}G(H)}^H M = 1 \) for all \( M \in \text{Irr} \mathbb{S}_k(H) \). If \( M \neq 1_{\mathbb{k}G(H)} \), then \( \text{Ind}_{\mathbb{k}G(H)}^H M \) is not equivalent to \( 1_{H^*} \) and so \( \text{Ind}_{\mathbb{k}G(H)}^H M \equiv \mathbb{k}_g \) for some \( 1 \neq g \in \text{Hom}_{\mathbb{Alg}_{\mathbb{k}}}(H^*, \mathbb{k}) = G(H) \). This gives the desired conclusion \( G(H) \neq 1 \). \( \square \)
The following application of the class equation, due to Masuoka [148], generalizes the familiar group-theoretical fact that finite \( p \)-groups have nontrivial centers. The proof elaborates on some ideas in the proof of Theorem 12.22.

**Theorem 12.23.** Let \( H \) be a semisimple Hopf algebra over an algebraically closed field \( \mathbb{k} \) of characteristic 0 and assume that \( \dim_{\mathbb{k}} H = p^n \) for some prime \( p \) prime and some positive integer \( n \). Then \( \mathbb{G}(H) \cap \mathcal{Z} H \neq \langle 1 \rangle \).

**Proof.** By the class equation, there must exist some \( f \in (12.28) \) write \( \text{Ind} \) field \( k \). Let Theorem 12.23. The proof elaborates on some ideas in the proof of Theorem 12.22. Since \( H \) is simultaneously a left \( \mathcal{Z} \) \( k \)-comodule algebra under \( \Delta \) such that \( \lambda = g^{-1} - \lambda \). But, since \( H^* \) is a left \( H \)-module algebra under \( \cdot \) \( (10.4.2) \), we have \( g^{-1}(g^{-1} - \lambda)f) = \lambda(g - f) = \lambda(g - f, 1) = \lambda(f, g) \) for all \( f \in H^* \) and so

\[
(12.29) \quad (g^{-1} - \lambda)f = (g^{-1} - \lambda)(f, g)
\]

Therefore,

\[
g^{-1} - \lambda = (12.28) \quad (g^{-1} - \lambda)(\langle f', g \rangle) = (12.29) \quad (g^{-1} - \lambda)e' = (12.28) \quad \langle g^{-1} - \lambda, g \rangle e' = e'
\]

as desired. \( \square \)

**12.4.5. Freeness over Hopf Subalgebras**

Recall from §10.1.2 that, for any Hopf algebra \( H \) and any Hopf subalgebra \( K \subseteq H \), we have the category \( \mathcal{H}_K \text{Mod} \) of relative \((H,K)\)-Hopf modules: every \( M \in \mathcal{H}_K \text{Mod} \) is simultaneously a left \( K \)-module and a left \( H \)-comodule, with structure maps \( \mu: K \otimes M \to M \) and \( \delta: M \to H \otimes M \), such that \( \mu \) is an \( H \)-comodule map or, equivalently, \( \delta \) is a \( K \)-module map; both conditions amount to the identity

\[
(12.30) \quad (k.m)_{(-1)} \otimes (k.m)_{(0)} = k_{(1)}m_{(-1)} \otimes k_{(2)}m_{(0)} \quad (k \in K, m \in M)
\]

For example, \( H \in \mathcal{H}_K \text{Mod} \) via the comultiplication \( \Delta_H \) and the left regular action of \( K \) on \( H \).

The following important theorem is due to Nichols and Zoeller [157]. It can be seen as a wide-ranging generalization of the standard fact that any (finite) group algebra is free over any subgroup algebra (Exercise 3.1.3) and, in particular, the
order of any finite group is divisible by the orders of all its subgroups (Lagrange’s Theorem).

**Nichols-Zoeller Theorem.** Let $H$ be a finite-dimensional Hopf algebra and let $K$ be a Hopf subalgebra. Then every object of $\mathcal{H}_K\text{Mod}$ is free as $K$-module; similarly for $\text{Mod}^{H}_K$. In particular, $H$ is free as right and left $K$-module and, consequently, $\dim_k K$ divides $\dim_k H$.

Various generalizations of this result have been proved subsequently. For example, the statement about $\mathcal{H}_K\text{Mod}$ remains true, with the same proof as will be given below, as long as $K$ is a left coideal subalgebra of $H$ that is Frobenius [147]. However, the finite-dimensionality hypothesis on $H$ cannot simply be omitted from the theorem; see [172, Section 9.4] for a counterexample.

We preface the proof of the Nichols-Zoeller Theorem with a short discourse on faithfulness versus freeness. All questions pertaining to these issues are readily answered by means of the Krull-Schmidt Theorem (§1.2.6). A $k$-algebra $A$ is said to be self-injective if the regular representation $A_{\text{reg}}$ is injective, and $A$ called augmented if it is equipped with an algebra map $\varepsilon: A \to k$. Inasmuch as Frobenius algebras are self-injective by Proposition 2.22, any finite-dimensional Hopf algebra is an augmented self-injective algebra.

**Lemma 12.24.** Let $A$ be a finite-dimensional augmented self-injective algebra and let $V \in \text{Rep}_\text{fin} A$.

(a) There exists $r \in \mathbb{N}$ such that $V\oplus^r \cong F \oplus T$ with $F$ free and $T$ non-faithful.

(b) $V$ is free $\iff$ $V\oplus^n$ is free for all $n \in \mathbb{N}$ $\iff$ $V\oplus^n$ is free for some $n \in \mathbb{N}$.

**Proof.** (a) Let $\{P_i\}_{i=1}^t$ be the principal indecomposable representations of $A$ (§2.1.4); so

$$A_{\text{reg}} \cong P_1^\oplus n_1 \oplus \cdots \oplus P_t^\oplus n_t,$$

with $n_i \in \mathbb{N}$. We claim that $V$ is faithful if and only if all $P_i$ are direct summands of $V$. Indeed, if this holds, then $A_{\text{reg}}$ is a summand of $V^\oplus n$ for $n = \max\{n_i\}$, and hence $0 = \text{Ker } V^\oplus n = \text{Ker } V$. To prove the converse, recall that $V$ is faithful if and only if $A_{\text{reg}}$ is a subrepresentation of $V^\oplus n$ for some $n \in \mathbb{N}$; see Exercise 1.2.2 or the proof of Burnside’s Theorem (§1.4.6). In this case, $A_{\text{reg}}$ is in fact a direct summand of $V^\oplus n$ by virtue of our self-injectivity hypothesis, and hence all $P_i$ are direct summands of $V^\oplus n$. Since $V^\oplus n$ and $V$ have the same indecomposable direct summands by the Krull-Schmidt Theorem (disregarding multiplicities), it follows that all $P_i$ are direct summands of $V$ as well. This proves our claim.

To derive (a) from the claim, let $r = 1$ and $F = 0$ in case $V$ is not faithful. Otherwise, write $V \cong P_1^\oplus m_1 \oplus \cdots \oplus P_t^\oplus m_t \oplus Q$ with $m_i \in \mathbb{N}$ and no $P_i$ being a direct summand of $Q$, and take $r = n_i^*$ and $F = A_{\text{reg}}^\oplus m_i^*$, where $n_i^* = \max\{n_i\}$. 


(b) The implications \( \Rightarrow \) all being trivial, we only need to show that freeness of \( V^\otimes n \) for some \( n \) also liberates \( V \). So assume that \( V^\otimes n \cong A_{\text{reg}}^\otimes r \). Inducing along the augmentation \( \varepsilon : A \to \mathbb{k} \), we obtain

\[
(\text{Ind}_A^H V)^{\otimes n} \cong \text{Ind}_A(H_{\text{reg}}^{\otimes n}) \cong \text{Ind}_A(A_{\text{reg}}^\otimes r) \cong \mathbb{k}^\otimes r
\]

and so \( r = nt \) with \( t = \dim_\mathbb{k} \text{Ind}_A^H V \). Therefore, \( V^\otimes n \cong F^\otimes n \) with \( F = A_{\text{reg}}^\otimes m \), and hence \( V \cong F \) by the Krull-Schmidt Theorem, proving that \( V \) is indeed free. \( \square \)

The next preparatory result has some more Hopf specific content.

**Proposition 12.25.** Let \( H \) be a finite-dimensional Hopf algebra and let \( V, W \in \text{Rep}_{\text{fin}} H \), with \( W \) being faithful. If \( W \otimes V \cong V^\otimes \dim_\mathbb{k} W \), then \( V \) is free.

**Proof.** We first observe that \( W \otimes V \) is faithful provided \( V \neq 0 \). To see this, use the fact that \( H_{\text{reg}} \) is a subrepresentation of \( W^\otimes n \) for some \( n \in \mathbb{N} \) (Lemma 12.24), and hence \( H_{\text{reg}} \otimes V \) is a subrepresentation of \( (W \otimes V)^{\otimes n} \). Since \( H_{\text{reg}} \otimes V \cong H_{\text{reg}}^{\otimes \dim_\mathbb{k} V} \) (Corollary 10.5), we conclude that \( H_{\text{reg}} \otimes V \) is faithful for \( V \neq 0 \), and hence so are \( (W \otimes V)^{\otimes n} \) and \( W \otimes V \).

Next, by Lemma 12.24, there exists some \( r \in \mathbb{N} \) such that \( V^\otimes r \cong F \oplus T \), with \( F \) free and \( T \) non-faithful, and it suffices to show that \( T = 0 \). By our observation in the preceding paragraph, this amounts to showing that \( W \otimes T \) is non-faithful. But \( W \otimes F \cong F^\otimes d \) with \( d = \dim_\mathbb{k} W \); see (10.22) and its proof. Thus, our hypothesis \( W \otimes V \cong V^\otimes d \) implies

\[
F^\otimes d \oplus T^\otimes d \cong W \otimes V^\otimes r \cong (W \otimes F) \oplus (W \otimes T)
\]

\[
\cong F^\otimes d \oplus (W \otimes T)
\]

Consequently, \( W \otimes T \cong T^\otimes d \) by the Krull-Schmidt Theorem, whence \( W \otimes T \) is non-faithful and the proof is complete. \( \square \)

**Proof of the Nichols-Zoeller Theorem.** It suffices to prove the assertion for \( H^H \text{Mod} \); passing to \( \text{bi} \text{op} \) will then give the result for \( H^H \text{Mod} \) as well.

First, consider a finite-dimensional \( M \in H^H \text{Mod} \). Fix some \( W \in \text{Rep}_{\text{fin}} H \) such that \( \text{Res}_K^H W \) is faithful; for example, \( W = H_{\text{reg}} \) will do. By Proposition 12.25, it will suffice to show that \( W \otimes M \cong W_{\text{triv}} \otimes M \) as \( K \)-modules, where \( W_{\text{triv}} \) is \( W \) with the trivial \( K \)-action, \( k \cdot w = \langle \varepsilon, k \rangle w \). Define linear maps \( \phi : W_{\text{triv}} \otimes M \to W \otimes M \) by \( \psi(w \otimes m) = m_{(-1)} \cdot w \otimes m_{(0)} \) and \( \psi(w \otimes m) = (S^{-1} m_{(-1)} \cdot w) \otimes m_{(0)} \). A straightforward check shows these maps to be inverse to each other, and the calculation

\[
\phi(k \cdot (w \otimes m)) = \phi(w \otimes k \cdot m) = \phi((k \cdot m)_{(-1)} \cdot w \otimes (k \cdot m)_{(0)})
\]

\[
= (k_1 m_{(-1)} \cdot w) \otimes k_2 m_{(0)} = k \cdot \phi(w \otimes m)
\]

proves the asserted isomorphism in \( \text{Rep}_{\text{fin}} K \).
It remains to deal with an arbitrary $M \in \mathcal{H}_k \text{Mod}$. By Zorn’s Lemma, there is a subset $\mathcal{F} \subseteq M$ that is maximal with respect to being $K$-linearly independent and satisfying $\delta_M \mathcal{F} \subseteq H \otimes K \mathcal{F}$, where $\delta_M$ is the $H$-coaction of $M$. We claim that $M = K \mathcal{F}$. Suppose otherwise and consider the canonical map $M \rightarrow \overline{M} = M / K \mathcal{F}$; this is a map of $(H, K)$-Hopf modules. By the Finiteness Theorem for comodules (§9.2.2), $\overline{M}$ has a finite-dimensional $H$-subcomodule $\overline{N} \neq 0$. Replacing $N$ by $K \overline{N}$, we may assume that $\overline{N}$ is in fact a $(H, K)$-Hopf submodule of $M$. By the preceding paragraph, $\overline{N}$ has a $K$-basis, say $\overline{G}$. Letting $G \subseteq M$ denote the preimage of this basis, the set $G \cup \mathcal{F}$ satisfies the requirements on $\mathcal{F}$ while being strictly larger. This contradiction proves our claim and completes the proof of the theorem. □

Exercises for Section 12.4

Throughout, $H$ denotes a Hopf algebra.

12.4.1 (Relative trace maps). HopfMaterials

12.5. Frobenius-Schur Indicators

In this section, we consider the scalars $\nu_n(V)$, for $V \in \text{Rep}_{\text{fin}} H$, that have already played a role in the proof of the Brauer-Fowler Theorem (§3.6.3) with $n = 2$. In particular, as was promised earlier, we will show here that if $V$ is absolutely irreducible, then $\nu_2(V)$ can only take the values 0 or ±1. We will follow the memoir [118] by Kashina, Sommerhäuser and Zhu rather closely.

12.5.1. Higher Frobenius-Schur Indicators

For an arbitrary Hopf algebra $H$ and $n \in \mathbb{N}$, the Sweedler powers $[n]: H \rightarrow H$ are defined by

$$h^{[n]} := h(1_2 \ldots h(n)) \quad (h \in H)$$

Our focus will be on the case where $H$ is semisimple. Then, by Maschke’s Theorem for Hopf algebras (§12.3.1), there is a unique $\Lambda \in \mathcal{I}_H$ such that $\langle e, \Lambda \rangle = 1$. The $n$th Frobenius-Schur indicator of $V \in \text{Rep}_{\text{fin}} H$ is defined by

$$\nu_n(V) \overset{\text{def}}{=} \chi_V(\Lambda^{[n]}).$$

For example, when $H = \mathbb{k}G$ is the group algebra of the finite group $G$ whose order is invertible in $\mathbb{k}$, then $\Lambda = \frac{1}{|G|} \sum_{g \in G} g$ and so $\nu_2(V) = \frac{1}{|G|} \sum_{g \in G} \chi_V(g^n)$ as in Lemma 3.31.

Our first goal is to give an alternative formula for $\nu_n(V)$ involving the map $t_n \in \text{GL}(V^{\otimes n})$ that is given by that action of the cyclic permutation $(1 2 \ldots n) \in S_n$.
on $V^\otimes n$ as in (3.64):

$$t_n(v_1 \otimes v_2 \otimes \cdots \otimes v_n) := v_n \otimes v_1 \otimes \cdots \otimes v_{n-1} \quad (v_i \in V)$$

**Proposition 12.26.** Let $H$ be a semisimple involutory Hopf algebra. Then, for any $V \in \text{Rep } H$ and $n \in \mathbb{N}$, the space of $H$-invariants $(V^\otimes n)^H$ is stable under $t_n$. If $V$ is finite dimensional, then

$$v_n(V) = \text{trace}(t_n|_{(V^\otimes n)^H})$$

**Proof.** As we have remarked in the proof of Maschke's Theorem for Hopf algebras, $\Lambda_W$ is a projection $W \to W^H$ for any $W \in \text{Rep } H$. We need to show that $\Lambda.V^\otimes n$ is stable under $t_n$. The Hopf algebra $H^*$ is involutory, because $H$ is so, and $H^*$ is unimodular by Corollary 12.17. As in (12.25) (see also Exercise 12.1.4), it follows that $\Lambda(1) \otimes \Lambda(2) = \Lambda(2) \otimes \Lambda(1)$, which further implies that $\Lambda(1) \otimes \cdots \otimes \Lambda(n) = \Lambda(2) \otimes \cdots \otimes \Lambda(n) \otimes \Lambda(1)$ for all $n$. Consequently,

$$t_n(\Lambda.(v_1 \otimes v_2 \otimes \cdots \otimes v_n)) = t_n(\Lambda(1)_v \otimes \cdots \otimes \Lambda(n)_v \otimes v_{n-1} \otimes \Lambda(1)_v)$$

$$= \Lambda(1)_v \otimes \Lambda(2)_v \otimes v_1 \otimes \cdots \otimes \Lambda(1)_v \otimes v_{n-1}$$

Thus $t_n \circ \Lambda_{V^\otimes n} = \Lambda_{V^\otimes n} \circ t_n$, which proves the inclusion $t_n(\Lambda.V^\otimes n) \subseteq \Lambda.V^\otimes n$. Since $\Lambda_{V^\otimes n}$ is a projection of $V^\otimes n$ onto $(V^\otimes n)^H = \Lambda.V^\otimes n$, it also follows that $\text{trace}(t_n|_{(V^\otimes n)^H}) = \text{trace}(t_n \circ \Lambda_{V^\otimes n})$. Finally, $\Lambda_{V^\otimes n} = f_1 \otimes f_2 \otimes \cdots \otimes f_n$ with $f_i = (\Lambda_{i})_{V}$ and so the formula for the Frobenius-Schur indicator $v_n(V)$ will be a consequence of the following trace identity, which holds for any $f_i \in \text{End}_k(V)$,

$$\text{trace}(t_n \circ (f_1 \otimes f_2 \otimes \cdots \otimes f_n)) = \text{trace}(f_1 \circ f_2 \circ \cdots \circ f_n)$$

The verification of this identity is elementary and is left as Exercise 12.5.4. \(\square\)

Since $t_n$ has order $n$, the formula in Proposition 12.26 shows in particular that, if $H$ is semisimple and $\text{char } k = 0$, then $v_n(V) \in \mathbb{Z}[\mu_n]$, where $\mu_n = \{ \zeta \in \overline{k} \mid \zeta^n = 1 \}$ and $\overline{k}$ is an algebraic closure of $k$; in particular, $v_2(V) \in \mathbb{Z}$. If $H$ is cocommutative and $\text{char } k = 0$, then it is not hard to show that all $v_n(V)$ are integers (Exercise 12.5.2). See [118, §7.5] for examples in characteristic 0 with $v_n(V) \not\in \mathbb{R}$.

**12.5.2. The Second Frobenius-Schur Indicator**

We now focus on the original Frobenius-Schur indicator, $v_2$, as introduced by Frobenius and Schur [76] for complex representations of finite groups. The theorem below, due Montgomery and Linchenko [132] and independently to Fuchs, Ganchev, Szlachányi, and Vecsernyés [77], generalizes a result of Frobenius and Schur.

For the statement of the theorem, we need some generalities concerning bilinear forms. A bilinear form $B : V \times V \to k$, for $V \in \text{Vect}_k$, is called **symmetric** if $B(v, v') = B(v', v)$ for all $v, v' \in V$; if $B(v, v) = 0$ for all $v \in V$, then $B$ is said
to be alternating. In case char \( k \neq 2 \), the latter condition is equivalent to skew-symmetry: \( B(v, v') = -B(v', v) \) for all \( v, v' \in V \) (§5.1.1). The collection of all bilinear forms on \( V \) forms a \( k \)-vector space, \( \text{Bil}(V, k) = \text{MultLin}(V^2, k) \), that is canonically isomorphic to \((V \otimes V)^*\) by (B.13) and hence to \( \text{Hom}_k(V, V^*) \) by \( \text{Hom} \otimes \) adjunction (B.15). Explicitly,

\[
\begin{array}{ccc}
\text{Bil}(V, k) & \xrightarrow{\sim} & (V \otimes V)^* \\
\updownarrow & & \updownarrow \\
B & \mapsto & (v \otimes v' \mapsto B(v, v')) \quad \mapsto \quad (v \mapsto B(v, \cdot))
\end{array}
\tag{12.31}
\]

The form \( B \) is said to be non-degenerate on the left if \((v \mapsto B(v, \cdot)) \in \text{Hom}_k(V, V^*)\) is a monomorphism. For finite-dimensional \( V \), this is equivalent to the corresponding right-handed condition being satisfied, in which case \( B \) is simply called non-degenerate (Exercise ??). If \( V \) is a representation of some Hopf algebra \( H \), which may be arbitrary for the time being, then the vector spaces \((V \otimes V)^*\) and \( \text{Hom}_k(V, V^*) \) are equipped with their usual structures of \( H \)-representations and the above isomorphism between them is in fact an isomorphism in \( \text{Rep} H \); see Exercise 10.1.4 for a more general fact. The isomorphism with \((V \otimes V)^*\) allows us to give \( \text{Bil}(V, k) \) an \( H \)-action: \((h.B)(v, v') = B(S(h(2)), v, S(h(1)).v')\). As usual, \( B \) is said to be \( H \)-invariant if \( h.B = \langle e, h \rangle B \) for all \( h \in H \). If the antipode \( S \) is surjective, as it certainly is for semisimple \( H \), then \( H \)-invariance amounts to the condition

\[
B(h(1)).v, h(2), v') = \langle e, h \rangle B(v, v')
\tag{12.32}
\]

for all \( h \in H \) and \( v, v' \in V \). We leave it to the reader to check some of the details of the assertions made in the foregoing (Exercise 12.5.3).

**Theorem 12.27.** Let \( H \) be a semisimple involutory Hopf algebra over a field \( k \) of characteristic \# 2 and let \( S \in \text{Irr} H \) be absolutely irreducible. Then \( \nu_2(S) \in \{0, \pm 1\} \). Furthermore,

- \( \nu_2(S) = 0 \) if and only if \( S \) is not self-dual; in this case, \( S \) admits no nonzero \( H \)-invariant bilinear form.
- \( \nu_2(S) = 1 \) if and only if \( S \) admits an \( H \)-invariant non-degenerate bilinear form that is symmetric.
- \( \nu_2(S) = -1 \) if and only if \( S \) admits an \( H \)-invariant non-degenerate bilinear form that is alternating.

**Proof.** Our remarks above in conjunction with the standard isomorphism \((S \otimes S)^* \equiv S^* \otimes S^*\) in (10.12) yield the following isomorphisms of \( H \)-invariants:

\[
\text{Bil}(S, k)^H \equiv (S^* \otimes S^*)^H \equiv \text{Hom}_k(S, S^*)^H = \text{Hom}_H(S, S^*)
\tag{10.9}
\]

First assume that \( S \) is not self-dual or, equivalently, that \( \text{Hom}_H(S, S^*) = 0 \) (Schur’s Lemma). Then \( S \) admits no nonzero \( H \)-invariant bilinear form and \((S^* \otimes S^*)^H = 0\).
Therefore, \( \nu_2(S') = \text{trace}(t_2|_{(S'^2)^H}) = 0 \) Proposition 12.26. Inasmuch as \( S^* \) is not self-dual either, we also have \( 0 = \nu_2(S'^*) = \nu_2(S) \) in this case. Alternatively, we could invoke Exercise 12.5.1 for the last conclusion.

From now on assume that \( S \cong S^* \). Then the above isomorphisms yield
\[
\text{Bil}(S, \mathbb{k})^H \cong (S^\otimes_2)^H \cong D(S) = \mathbb{k}
\]
where the last equality holds, because \( S \) is absolutely irreducible. Since \( (S^\otimes_2)^H \) is stable under \( t_2 \) by Proposition 12.26 and since \( t_2 \) has order 2, it follows that \( t_2 \) acts on \( (S^\otimes_2)^H \cong \mathbb{k} \) by \( \pm 1 \) and so \( \nu_2(S) = \text{trace}(t_2|_{(S^\otimes_2)^H}) = \pm 1 \). To make the connection with bilinear forms, observe that any nonzero form in \( \text{Bil}(V, \mathbb{k})^H \) is automatically non-degenerate by Schur’s Lemma. Next, recall from (3.67) that
\[
S^\otimes_2 = (S^\otimes_2)(\mathbb{1}) \oplus (S^\otimes_2)({\text{sgn}}) \cong \text{Sym}^2S \oplus \Lambda^2S
\]
as representations of \( \mathbb{k}[t_2] = \mathbb{k}S_2 \). If \( \nu_2(S) = 1 \), then \( (S^\otimes_2)^H \cong (S^\otimes_2)(\mathbb{1}) \cong \text{Sym}^2S \), giving a unique up to scalar multiples \( H \)-invariant non-degenerate bilinear form that is symmetric. Similarly, \( \nu_2(S) = -1 \) leads to \( (S^\otimes_2)^H \cong (S^\otimes_2)({\text{sgn}}) \cong \Lambda^2S \) and an alternating form (Exercise 12.5.3).

**12.5.3. Return to Finite Group Algebras**

We close out this section by elaborating on the classical case \( H = \mathbb{C}G \) for a finite group \( G \). Following Frobenius and Schur [76], the three types of \( S \in \text{Irr} \mathbb{C}G \) that correspond to the three possible values of \( \nu_2(S) \) can also be described as follows. As in §3.1.5, we shall regard characters as a class functions, \( \chi_S : G \rightarrow \mathbb{C} \).

- \( \nu_2(S) = 0 \). As we have seen in (3.40), self-duality of \( S \in \text{Irr} \mathbb{C}G \) is equivalent to \( \chi_S(g) \in \mathbb{R} \) for all \( g \in G \). Thus, by Theorem 12.27, \( \nu_2(S) = 0 \) if and only if \( \chi_S \) is not \( \mathbb{R} \)-valued: \( \chi_S(g) \notin \mathbb{R} \) for some \( g \in G \).

- \( \nu_2(S) = 1 \). In this case, \( \chi_S \) must be \( \mathbb{R} \)-valued, but more is true. In fact, \( \nu_2(S) = 1 \) if and only if the entire group representation \( G \rightarrow \text{GL}(S) \) can be realized over \( \mathbb{R} \) in the sense that, for some basis of \( S \), the matrices of all \( g_S \) with \( g \in G \) have entries in \( \mathbb{R} \). This follows from Lemma 12.28 below.

- \( \nu_2(S) = -1 \). In light of the foregoing, this case must correspond to \( \chi_S \) being \( \mathbb{R} \)-valued but nonetheless \( S \) not being realizable over \( \mathbb{R} \). See Example 12.29 below for a specific \( S \in \text{Irr} \mathbb{C}G \) of this type.

Here is the promised lemma. Note that, for a group algebra \( H = \mathbb{k}G \), the \( H \)-invariance condition (12.32) is equivalent to \( B(g.v, g.v') = B(v, v') \) for all \( g \in G \) and \( v, v' \in V \). We will therefore refer to such bilinear forms \( B \) as \( G \)-invariant.

**Lemma 12.28.** Let \( G \) be a finite group and let \( S \in \text{Irr} \mathbb{C}G \). Then \( S \) can be realized over \( \mathbb{R} \) if and only if there exists a non-degenerate, symmetric, and \( G \)-invariant bilinear form \( B : S \times S \rightarrow \mathbb{C} \).
**Proof.** First assume that $S$ can be realized over $\mathbb{R}$; in other words, $S \cong \mathbb{C} \otimes_{\mathbb{R}} S_0$ for some $S_0$ in $\text{Rep} \mathbb{R} G$ (necessarily irreducible). Fix an arbitrary positive definite symmetric bilinear form $B_0 \in \text{Bil}(S_0, \mathbb{R})$, for example the dot product for some basis of $S_0$, and define $B_1 \in \text{Bil}(S, \mathbb{R})$ by $B_1(s, s') := \sum_{g \in G} B_0(g.s, g.s')$ to obtain a positive definite symmetric bilinear form that is also $G$-invariant. The $\mathbb{C}$-bilinear extension of $B_1$ to $S \times S$ then gives the desired $G$-invariant symmetric bilinear form $B \in \text{Bil}(S, \mathbb{C})$ which is non-degenerate.

Conversely, assume that we are given such a form, say $B \in \text{Bil}(S, \mathbb{C})$. In addition, there is always a $G$-invariant Hermitian inner product $H: S \otimes S \to \mathbb{C}$ that is obtained from an arbitrary Hermitian inner product $H_0: S \otimes S \to \mathbb{C}$ as in the preceding paragraph of the proof; see Exercise 3.4.12. Thus, we have $G$-equivariant additive isomorphisms

$$f_B: S \xrightarrow{\sim} S^*$$

and

$$f_H: S \xrightarrow{\sim} S^*$$

for some $s \in S$.

The isomorphism $f_B$ is $\mathbb{C}$-linear, whereas $f_H$ is $\mathbb{C}$-skew linear: $f_H(\lambda s) = \overline{\lambda} f_H(s)$ for $s \in S$ and $\lambda \in \mathbb{C}$, where $\overline{\lambda}$ denotes complex conjugation. The map $f := f_H^{-1} \circ f_B$ is a $G$-equivariant skew linear automorphism of $S$, and hence $f^2: S \to S$ is an automorphism in $\text{Rep} \mathbb{C} G$. By Schur's Lemma, we must have $f^2 = \lambda \text{Id}_S$ for some $\lambda \in \mathbb{C}^\times$. In fact, we claim that $\lambda \in \mathbb{R}_{>0}$. To see this, we calculate for $s, s' \in S$,

$$H(s, f(s')) = \langle v, (f_H \circ f)(s') \rangle = \langle v, f_B(s') \rangle = B(s, s')$$

By symmetry of $B$, it follows that $H(s, f(s')) = H(s', f(s)) = \overline{H(f(s), s')}$ and hence $\lambda H(s, s) = H(f^2(s), s) = H(f(s), f(s))$. For $s \neq 0$, both $H(s, s)$ and $H(f(s), f(s))$ are positive real, and so $\lambda$ is positive real as well as claimed. Writing $\lambda = \rho^2$ with $\rho \in \mathbb{R}_{>0}$ and replacing $H$ by $\rho H$, we may assume that $\lambda = 1$; so $f^2 = \text{Id}_S$. It follows that $S = S_+ \oplus S_-$ with $S_\pm = \{ s \in S \mid f(s) = \pm s \}$; both are $\mathbb{R} G$-subrepresentations of $S$. Finally, $f(i s) = -i f(s)$ by skew linearity, and so $i S_+ = S_-$. Therefore, $S = \mathbb{C} \otimes_{\mathbb{R}} S_+$, giving the desired realization of $S$ over $\mathbb{R}$. \hfill $\square$

Finally, here is the promised example.

**Example 12.29** (An irreducible complex group representation with $v_2 = -1$). Let $H = \mathbb{R} + i \mathbb{R} + j \mathbb{R} + k \mathbb{R}$ be the real quaternions, with $i^2 = j^2 = k^2 = i j k = -1$, and put $G := \{ i, j \} = \{ \pm 1, \pm i, \pm j, \pm k \} \leq H^\times$. Letting $G$ act on $H$ by left multiplication and writing $S := H = \mathbb{C} \otimes_{\mathbb{R}} j \mathbb{C} \cong \mathbb{C}^2$, we obtain a complex representation $G \hookrightarrow \text{GL}_2(\mathbb{C})$:

$$\begin{align*}
\pm 1 & \mapsto \pm \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & \pm i & \mapsto \pm \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} & \pm j & \mapsto \pm \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} & \pm k & \mapsto \pm \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}
\end{align*}$$

The representation $S$ is irreducible; for, otherwise the image of $G$ in $\text{GL}_2(\mathbb{C})$ would be conjugate to a group of diagonal matrices, contradicting the fact that $G$ is
non-abelian. Since $(\pm 1)^2 = 1$ and $(\pm i)^2 = (\pm j)^2 = (\pm k)^2 = -1$, we obtain

$$\nu_2(S) = \frac{1}{8} \sum_{g \in G} \chi_S(g^2) = \frac{1}{8}(2 \cdot 2 + 6 \cdot (-2)) = -1$$

Thus, while the character $\chi_S$ is certainly $\mathbb{R}$-valued, it is not possible to realize the representation $S$ over $\mathbb{R}$.

**Exercises for Section 12.5**

Throughout, $H$ denotes a Hopf algebra.

12.5.1 (Frobenius-Schur indicators of duals). Show that $S(h^{[n]}) = (Sh)^{[n]}$ for all $h \in H$ and arbitrary $H$. For $H$ semisimple and $V$ in $\text{Rep } H$, conclude that $\nu_n(V) = \nu_n(V^*)$.

12.5.2 (Frobenius-Schur indicators for cocommutative $H$). Let $H$ be a cocommutative Hopf $k$-algebra, with $\text{char } k = 0$, and let $V$ be in $\text{Rep}_{\text{fin }} H$. Show that $\text{trace}(t_n(V^\otimes n)) \in \mathbb{Z}$ for all $n$, where $t_n = (1 \ 2 \ \ldots \ n) \in S_n$. In particular, if $H$ is semisimple, then $\nu_n(V) \in \mathbb{Z}$.

12.5.3 (Details on bilinear forms). Check the details for (12.31). Spell out invariance for groups (done in section) and Lie algebras. Left-right symmetry of “non-deneracy” (via determinant, say) for finite-dimensional $V$. Fill in the details for $\nu_2(S) = \pm 1 \implies$ symmetric/alternating form.

12.5.4 (A trace identity). Let $V$ be a finite-dimensional $k$-vector space and let $t_n \in \text{End}_k(V^\otimes n)$ be given by the place permutation action of $(1 \ 2 \ \ldots \ n) \in S_n$ as in (3.64). Show that $\text{trace}(t_n \circ (f_1 \otimes f_2 \otimes \cdots \otimes f_n)) = \text{trace}(f_1 \circ f_2 \circ \cdots \circ f_n)$ for any $f_i \in \text{End}_k(V)$.
This appendix is a brief compendium of the rudiments pertaining to categories. We shall be rather cavalier about the foundational formalities of category theory—these do require careful attention in a proper treatment of the subject. For an authoritative account, see Mac Lane’s classic *Categories for the Working Mathematician* [140] or the more recent *Abstract and Concrete Categories: the Joy of Cats* by Adámek, Herrlich and Strecker [2].

A.1. Categories

A *category* $\mathcal{C}$ consists of the following data:

(i) To start with, there is a collection $\text{Ob} \mathcal{C}$ of *objects*; this collection need not form a set. In fact, a prime example of a category is the category $\text{Sets}$ having all sets as its objects—the reader will recall that the notion of a set of all sets has paradoxical consequences. Even though $\text{Ob} \mathcal{C}$ may not be a set, we allow ourselves to write $X \in \mathcal{C}$ to mean that $X$ is an object of $\mathcal{C}$.

(ii) For any two objects $X, Y \in \mathcal{C}$, there is a set $\text{Hom}_{\mathcal{C}}(X, Y)$, called the set of all *morphisms* from $X$ to $Y$. Despite the fact that $X$ and $Y$ themselves may not actually be sets, it is common practice to use the familiar set theoretical notations

\[ f : X \to Y \quad \text{or} \quad X \xrightarrow{f} Y \]

in place of $f \in \text{Hom}_{\mathcal{C}}(X, Y)$. Morphisms are sometimes also referred to as “arrows” or “maps” in $\mathcal{C}$.
(iii) For any triple of objects \(X, Y, Z \in C\), there is a \textit{composition} function

\[
\text{Hom}_C(X, Y) \times \text{Hom}_C(Y, Z) \to \text{Hom}_C(X, Z)
\]

Composition is required to be \textit{associative} in the sense that

\[
(h \circ g) \circ f = h \circ (g \circ f)
\]

holds for any \(f: X \to Y\), \(g: Y \to Z\) and \(h: Z \to T\). Composition is often graphically rendered by following arrows in a diagram; a \textit{commutative diagram}

\[
\begin{array}{ccc}
X & \xrightarrow{f} & Y \\
\downarrow{c} & & \downarrow{g} \\
Y & \xrightarrow{g} & Z
\end{array}
\]

means that \(c = g \circ f\).

(iv) Finally, for each object \(X \in C\), there is a distinguished morphism \(\text{Id}_X \in \text{Hom}_C(X, X)\) satisfying the left and right \textit{unit axioms} for composition:

\[
f \circ \text{Id}_X = f \quad \text{and} \quad \text{Id}_Y \circ f = f
\]

for any \(f: X \to Y\). It follows from associativity that \(\text{Id}_X\) is uniquely determined by \(X\).

The entity \(\text{Ob} C\) is often called a “class,” but we will not get bogged down in the fine points of set theory. We shall generally avoid pondering whether two given objects \(X, Y \in C\) are literally the same. Other than for morphisms, strict equality of objects is usually not relevant from a categorical viewpoint, the more crucial notion being that of an \textit{isomorphism}. In detail, \(X \cong Y\) means that there are morphisms \(f: X \to Y\) and \(g: Y \to X\) such that \(g \circ f = \text{Id}_X\) and \(f \circ g = \text{Id}_Y\). In this case, one also writes \(f: X \cong Y\) and \(g: Y \cong X\) and calls \(f\) and \(g\) isomorphisms. An easy argument shows that \(f\) and \(g\) determine each other; one writes \(g = f^{-1}\) and \(f = g^{-1}\). For each \(X \in C\), the set of all isomorphisms \(X \cong X\) forms a group, \(\text{Aut}_C(X)\), called the \textit{automorphism group} of \(X\).

The more economical notation \(C(X, Y)\) is often used instead of \(\text{Hom}_C(X, Y)\). However, most of morphism that we shall encounter are in fact homomorphisms in some familiar setting and so we will stick with \(\text{Hom}_C(X, Y)\). It is also becoming increasingly common in category theory to dispense with the requirement that all \(\text{Hom}_C(X, Y)\) are sets; categories as defined above are then called \textit{locally small}. Thus, all our categories are understood to be locally small. A category \(C\) is said to be \textit{small} if, in addition, the collection of all objects of \(C\) is also a set. In this case, one can consider the set \(\text{Hom}_C = \bigsqcup_{X, Y \in C} \text{Hom}_C(X, Y)\) of all morphisms of \(C\).
A.1. Categories

A.1.1. Some Examples of Categories

We have already mentioned the category $\text{Sets}$; its morphisms are the functions (or maps) between sets, with the ordinary composition, and $\text{Id}_X$ is the usual identity function of $X \in \text{Sets}$. An isomorphism in $\text{Sets}$ is the same as a bijection. Besides $\text{Sets}$, the following categories will play a role in this book; a host of others will be introduced as we go along.

$\text{Vect}_k$, the category of all vector spaces over a given field $k$, with $k$-linear maps as morphisms. This is a subcategory of $\text{Sets}$: all objects of $\text{Vect}_k$ are also objects of $\text{Sets}$; each $\text{Hom}_{\text{Vect}_k}(V, W)$ for $V, W \in \text{Vect}_k$ is a subset of $\text{Hom}_{\text{Sets}}(V, W)$; composition in $\text{Vect}_k$ is obtained by restricting the composition of $\text{Sets}$; and $\text{Id}_V \in \text{Hom}_{\text{Sets}}(V, V)$ does in fact belong to the subset $\text{Hom}_{\text{Vect}_k}(V, V)$ for each $V \in \text{Vect}_k$. In place of $\text{Hom}_{\text{Vect}_k}(V, W)$, we shall use the more common special notation $\text{Hom}_k(V, W)$.

$\text{Groups}$ the category of all groups and group homomorphisms. Again, this is a subcategory of $\text{Sets}$. By considering only abelian groups and group homomorphisms between them, one obtains the subcategory $\text{AbGroups}$ of $\text{Groups}$; this is in fact a full subcategory of $\text{Groups}$ in the sense that $\text{Hom}_{\text{AbGroups}}(A, B) = \text{Hom}_{\text{Groups}}(A, B)$ holds for all $A, B \in \text{AbGroups}$ (not just $\subseteq$).

$\text{R-Mod}$ the category of all left modules and module maps over a given ring $R$ (not necessarily commutative, but with 1). This is a subcategory of $\text{AbGroups}$, but generally not a full one. Again, we will follow common practice in writing $\text{Hom}_R(M, N)$ rather than $\text{Hom}_{\text{R-Mod}}(M, N)$ for $M, N \in \text{R-Mod}$.

Like the above categories, all other categories $C$ considered in this book will be subcategories of $\text{Sets}$ — so the objects will always be sets with additional structure — and the categorical notion of an isomorphism in $C$ will coincide with the familiar one: a morphism in $C$ that is bijective as a set map.

As for more exotic examples, giving a glimpse of the range and flexibility of the concept of a category, we mention that a category $C$ with one object amounts to the datum of a monoid, that is, a set $M$ equipped with an associative law of composition and an identity element: just let $M = \text{Hom}_C(*, *)$, where $*$ denotes the object of $C$. Similarly, a group is essentially the same thing as a category with one object and such that all morphisms are isomorphisms. Any category $C$ having the property that all morphisms of $C$ are isomorphisms is called a groupoid. The morphisms of a category need not be functions, even if the objects are in fact certain sets. For example, one can consider the following category, the only non-locally-small category that we shall mention: $\text{Ob} C$ consists of all rings and, for given rings $A$ and $B$, the collection of morphisms $\text{Hom}_C(A, B)$ consists of all $(B, A)$-bimodules,
with composition given by the tensor product and with \( \text{Id}_A = A A \), the regular \((A, A)\)-bimodule (Example 1.2).

### A.2. Functors

Let \( C \) and \( D \) be categories. A **functor** \( F: C \to D \) is a rule that assigns to every object \( X \in C \) an object \( F(X) = FX \in D \), and to every morphism \( f \in \text{Hom}_C(X, Y) \) a morphism \( F(f) = Ff \in \text{Hom}_D(FX, FY) \). The functor \( F \) is required to respect composition and identities:

\[
F(g \circ f) = Fg \circ Ff \quad \text{and} \quad F(\text{Id}_X) = \text{Id}_{FX}
\]

In particular, \( F \) yields set maps \( \text{Hom}_C(X, Y) \to \text{Hom}_D(FX, FY) \) for each pair of objects \( X, Y \in C \). The functor \( F \) is called **faithful** if all these maps are injective, and **full** if they are surjective. A category \( C \) is called **concrete** if it is equipped with a faithful functor \( F: C \to \text{Sets} \); as was mentioned before, all categories occurring in this book will be concrete.

#### A.2.1. First Examples of Functors

For every category \( C \), there is the **identity functor** \( \text{Id}_C: C \to C \) fixing all objects and all morphisms of \( C \). If \( F: C \to D \) and \( G: D \to E \) are functors, then one can define the composite functor \( G \circ F: C \to E \) in the obvious fashion: \( (G \circ F)X = G(FX) \) for \( X \in C \) and likewise for morphisms. This composition is evidently associative and the identity functors satisfy the unit axioms. With this in hand, the set theoretically ruthless (or expert) may venture to consider the category of all categories, whose objects are the categories and whose morphisms are functors.

Returning to safer territory, we mention another type of functor, called a **forgetful functor**, whose effect is to ignore part of the structure. For example, there is the forgetful functor

\[
\cdot|_{\text{AbGroups}}: \text{Vect}_k \to \text{AbGroups}
\]

that assigns to each \( k \)-vector space its underlying abelian group and regards each \( k \)-linear map as just a homomorphism of abelian groups. Similarly, we have forgetful functors \( \cdot|_{\text{Mod}}: \text{AbGroups} \to \text{Sets} \) etc.

As a slightly more interesting example, we offer the functor

\[
\text{GL}_n: \text{Rings} \to \text{Groups}
\]

Here, \( \text{Rings} \) denotes the category with all rings as objects and with morphism the ring homomorphisms. For any ring \( R \), the group \( \text{GL}_n(R) \) consists of all invertible \( n \times n \)-matrices over \( A \). Any ring homomorphism \( f: R \to S \) gives rise to the group homomorphism \( \text{GL}_n(f): \text{GL}_n(R) \to \text{GL}_n(S), (r_{i,j}) \mapsto (f(r_{i,j})) \). Clearly, \( \text{GL}_n(g \circ f) = \text{GL}_n(g) \circ \text{GL}_n(f) \) and \( \text{GL}_n(\text{Id}_R) = \text{Id}_{\text{GL}_n(R)} \); so \( \text{GL}_n \) is indeed a functor.
A.2.2. Contravariant Functors and Opposite Categories

For emphasis, functors as introduced above are sometimes called covariant, although this is usually understood. In practice, one oftentimes encounters a rule \( F: \mathcal{C} \to \mathcal{D} \) that behaves like a functor inasmuch as it assigns to every object \( X \in \mathcal{C} \) an object \( FX \in \mathcal{D} \) and identity morphisms are preserved as usual, but “arrows are turned around”: to every morphism \( f \in \text{Hom}\_C(X,Y) \) there is an associated morphism \( Ff \in \text{Hom}_D(FY,FX) \). The composition preservation requirement then becomes \( F(g \circ f) = Ff \circ Fg \). In this case, \( F \) is called a contravariant functor.

An example of a contravariant functor is provided by the duality functor \( \cdot^\ast: \text{Vect}_k \to \text{Vect}_k \) that assigns to each \( V \in \text{Vect}_k \) its linear dual \( V^* = \text{Hom}_k(V,k) \) and to each \( k \)-linear map \( f: V \to W \) the transpose map \( f^*: W^* \to V^* \) (§B.3.2).

There is no need for a detailed separate treatment of contravariant functors. Indeed, every contravariant functor \( F: \mathcal{C} \to \mathcal{D} \), can be viewed as a (covariant) functor \( \mathcal{C}^{\text{op}} \to \mathcal{D} \) or else \( \mathcal{C} \to \mathcal{D}^{\text{op}} \). Here \( \mathcal{C}^{\text{op}} \) denotes opposite category of \( \mathcal{C} \); it is defined by giving \( \mathcal{C}^{\text{op}} \) the same objects as \( \mathcal{C} \) but putting \( \text{Hom}_{\mathcal{C}^{\text{op}}}(X,Y) = \text{Hom}_\mathcal{C}(Y,X) \) for any two objects \( X \) and \( Y \). Identity morphisms in \( \mathcal{C}^{\text{op}} \) are the same as those in \( \mathcal{C} \), and the \( \mathcal{C}^{\text{op}} \)-composite \( g \circ_{\text{op}} f \) of morphisms \( f \in \text{Hom}_{\mathcal{C}^{\text{op}}}(X,Y) \) and \( g \in \text{Hom}_{\mathcal{C}^{\text{op}}}(Y,Z) \) is equal to the \( \mathcal{C} \)-composite \( f \circ g \).

A.3. Naturality

In general, it makes little sense to ponder the question as to whether two given functors \( F, G: \mathcal{C} \to \mathcal{D} \) are equal to each other, because this would certainly require the “equality” of objects \( FX = GX \) for every \( X \in \mathcal{C} \), which generally is meaningless for us. The more appropriate concepts is that of natural equivalence or isomorphism of functors, which is indeed one of the central notions of category theory.

A.3.1. Natural Transformations and Functoriality

Let \( F, G: \mathcal{C} \to \mathcal{D} \) be functors. A natural transformation \( \alpha: F \Rightarrow G \) is a collection of morphisms \( \alpha_X: FX \to GX \) in \( \mathcal{D} \), one for each object \( X \in \mathcal{C} \), such that the following diagram commutes for all \( X, Y \in \mathcal{C} \) and all \( f \in \text{Hom}_\mathcal{C}(X,Y) \):

\[
\begin{array}{ccc}
FX & \longrightarrow & GX \\
\downarrow Ff & & \downarrow Gf \\
FY & \longrightarrow & GY \\
\end{array}
\]

(A.1)
Condition (A.1) is called the naturality condition. The notations

\[ \alpha: F \Rightarrow G: C \to D \]

are in use alongside some others.

For each functor \( F: C \to D \), there is the identity transformation \( \text{Id}_F: F \Rightarrow F \) given by \( (\text{Id}_F)_X = \text{Id}_{FX} \) for \( X \in C \). Moreover, natural transformations \( \alpha: F \Rightarrow G \) and \( \beta: G \Rightarrow H \) for functors \( F, G, H: C \to D \) can be composed componentwise:

\[ (\beta \circ \alpha)_X = \beta_X \circ \alpha_X. \]

This composition is manifestly associative and the identity transformations satisfy the unit laws. It is therefore tempting to consider the collection of all functors \( C \to D \) as the objects of a category \( D^C \) with morphisms \( \text{Hom}_{D^C}(F, G) \) given by the natural transformations \( F \Rightarrow G \), but this is generally problematic from a set theoretic perspective. Nonetheless, natural transformations are also called morphisms of functors. For given morphisms of functors \( \alpha, \beta: F \Rightarrow G: C \to D \), the equality \( \alpha = \beta \) is to be interpreted component-wise: \( \alpha_X = \beta_X \) holds in the set \( \text{Hom}_D(FX, GX) \) for all \( X \in C \).

In practice, one often considers a fixed but arbitrary object \( X \in C \) and constructs a morphism (in \( D \))

\[ h: FX \to GX \]

The morphism \( h \) is said to be functorial in \( X \) or natural if \( h = \alpha_X \) for some morphism of functors \( \alpha: F \Rightarrow G \). Thus, the construction of \( h \) is not specific to \( X \) and it is “well-behaved” under change of \( X \).

**Example A.1** (Functoriality of the determinant). It is a well-known fact from linear algebra that, for any commutative ring \( R \), the determinant gives a group homomorphism \( \text{GL}_n(R) \to R^\times \), where \( R^\times \) denotes the group of invertible elements of \( R \) and \( \text{GL}_n(R) \) is the group of invertible \( n \times n \) matrices over \( R \). To see that this group homomorphism is functorial in \( R \), recall that \( \text{GL}_n \) is a functor \( \text{Rings} \to \text{Groups} \). The same is true of \( \cdot \times \), which is just the case \( n = 1 \). Since the ordinary determinant fails to yield a group homomorphism for matrices over arbitrary rings, we consider the restrictions of these functors to the full subcategory \( \text{CommRings} \) of \( \text{Rings} \) consisting of all commutative rings. The determinant then is a morphism of functors

\[ \det: \text{GL}_n \Rightarrow \cdot \times: \text{CommRings} \to \text{Groups} \]

in the above sense: for any morphism \( f: R \to S \) in \( \text{CommRings} \) and any \( n \times n \) matrix \( (r_{i,j}) \) over \( R \), one has

\[ f(\det(r_{i,j})) = \sum_{s \in S_n} \text{sgn}(s) \prod_i f(r_{i,s(i)}) = \det(f(r_{i,j})) \]
A.3. Naturality

giving the requisite commutative diagram specializing (A.1),

\[
\begin{array}{c}
\text{GL}_n(R) \xrightarrow{\text{det}_R} R^X \\
\downarrow \text{GL}_n(f) \quad \downarrow f^X \\
\text{GL}_n(S) \xrightarrow{\text{det}_S} S^X
\end{array}
\]

A.3.2. Natural Equivalences

Let \( \alpha: F \Rightarrow G: \mathcal{C} \to \mathcal{D} \) be a morphism of functors. If each \( \alpha_X \) is an isomorphism, then \( \alpha \) is called a natural equivalence or an isomorphism of functors and one writes

\[ F \cong G \]

In this case, the inverses \( \beta_X = \alpha_X^{-1} \) yield a morphism of functors \( \beta: G \Rightarrow F \) with \( \beta \circ \alpha = \text{Id}_F \) and \( \alpha \circ \beta = \text{Id}_G \); in fact, the existence of such a \( \beta \) is clearly equivalent to \( \alpha \) being an isomorphism.

Oftentimes, one starts with an arbitrary object \( X \) of some category \( \mathcal{C} \) and with given functors \( F, G: \mathcal{C} \to \mathcal{D} \) and constructs an isomorphism

\[ h: FX \xrightarrow{\sim} GX \]

As in A.3.1, the isomorphism \( h \) is called functorial or natural if \( h \) is the \( X \)-component of some isomorphism of functors \( F \cong G \).

**Example A.2** (Naturality of the \( \text{Hom-} \otimes \) isomorphism). Let \( V \) and \( W \) be \( \mathbb{k} \)-vector spaces and assume that \( V \) is finite dimensional. Then there is a \( \mathbb{k} \)-linear isomorphism (Appendix B)

\[ W \otimes V^* \xrightarrow{\sim} \text{Hom}_\mathbb{k}(V, W) \tag{A.2} \]

where \( V^* = \text{Hom}_\mathbb{k}(V, \mathbb{k}) \) and \( \langle \cdot , \cdot \rangle: V^* \times V \to \mathbb{k} \) denotes the evaluation map. This isomorphism is natural (or functorial) in both \( V \) and \( W \). Indeed, \( F = \cdot \otimes V^* \) and \( G = \text{Hom}_\mathbb{k}(V, \cdot) \) both are functors \( \text{Vect}_\mathbb{k} \to \text{Vect}_\mathbb{k} \). Here, for any morphism \( \phi: W \to W' \) in \( \text{Vect}_\mathbb{k} \), the morphism \( F\phi: W \otimes V^* \to W' \otimes V^* \) is given by \((F\phi)(w \otimes f) = \phi(w) \otimes f \) and \( G\phi = \phi \circ \cdot \). The map in (A.2) is an isomorphism \( \alpha_W: FW \xrightarrow{\sim} GW \) in \( \text{Vect}_\mathbb{k} \).

Furthermore, the naturality condition (A.1) for the collection \( \alpha_W \) with \( W \in \text{Vect}_\mathbb{k} \) is readily verified; so (A.2) is in fact the \( W \)-component of an isomorphism of functors

\[ \cdot \otimes V^* \cong \text{Hom}_\mathbb{k}(V, \cdot): \text{Vect}_\mathbb{k} \to \text{Vect}_\mathbb{k} \]

establishing naturality in \( W \). For \( V \), one considers instead the contravariant functors \( F' = W \otimes \cdot^*, G' = \text{Hom}_\mathbb{k}(\cdot, W): \text{Vect}_\mathbb{k} \to \text{Vect}_\mathbb{k} \) and verifies in the same way that (A.2) defines a morphism \( F' \Rightarrow G' \). Thus, the \( \mathbb{k} \)-linear map in (A.2) is always
natural (or functorial) in \( V \), even if \( V \) is infinite-dimensional. Restricting \( F' \) and \( G' \) to the full subcategory \( \text{vect}_k \) of \( \text{Vect}_k \) consisting of all finite-dimensional \( k \)-vector spaces, one obtains the desired isomorphism of (contravariant) functors

\[
W \otimes \cdot \ast \cong \text{Hom}_k(\cdot, W) \colon \text{vect}_k \to \text{Vect}_k
\]

Of course, we could also treat \( V \) and \( W \) simultaneously by considering the functor

\[
\text{Hom}_k(\cdot, \cdot) \colon \text{Vect}_k^{\text{op}} \times \text{Vect}_k \to \text{Vect}_k,
\]

where \( \text{Vect}_k^{\text{op}} \times \text{Vect}_k \) denotes the product of the categories \( \text{Vect}_k^{\text{op}} \) and \( \text{Vect}_k \), defined in the obvious manner. A functor whose domain is a product category, such as \( \text{Hom}_k(\cdot, \cdot) \), is also called a bifunctor.

### A.3.3. Equivalence of Categories

Two categories \( C \) and \( D \) are said to be **equivalent** to each other if there exist functors \( F \colon C \to D \) and \( G \colon D \to C \) such that \( G \circ F \cong \text{Id}_C \) and \( F \circ G \cong \text{Id}_D \). In this case, the functors \( F \) and \( G \) are called **quasi-inverse** to each other and we write

\[
C \equiv D.
\]

In practice, an equivalence \( C \equiv D \) is often established by exhibiting a functor \( F \colon C \to D \) that satisfies the following two conditions:

- \( F \) is **full and faithful**, that is, \( F \) yields a bijection \( \text{Hom}_C(X, X') \cong \text{Hom}_D(FX, FX') \) for each pair of objects \( X, X' \in C \), and
- \( F \) is **essentially surjective**: each object of \( D \) is isomorphic to an object of the form \( FX \) with \( X \in C \).

A quasi-inverse \( G \colon D \to C \) may then be obtained by selecting, with the aid of a sufficiently strong version of the Axiom of Choice, for each \( Y \in D \) an object \( GX \in C \) and an isomorphism \( \epsilon_B : FGY \cong B \) in \( D \). For the proof that this does indeed result in a quasi-inverse \( G \), we refer the reader to [2, 3.36].

For example, the forgetful functor \( F \colon \mathbb{Z}\text{Mod} \to \text{AbGroups} \) of A.2.1 is in fact an equivalence; so \( \mathbb{Z}\text{Mod} \equiv \text{AbGroups} \). For the requisite quasi-inverse \( G \colon \text{AbGroups} \to \mathbb{Z}\text{Mod} \), observe that every abelian group \( X = (A, +) \) can be equipped with a unique \( \mathbb{Z} \)-action \( \alpha : \mathbb{Z} \times A \to A \) making \( (A, \alpha) \) into a \( \mathbb{Z} \)-module with underlying abelian group \( X \); this action corresponds to the unique ring homomorphism \( \mathbb{Z} \to \text{End}(A) \). Thus, we may put \( GX = (A, \alpha) \) and \( Gf = f \) for every homomorphism of abelian groups \( f : X \to Y \), because \( f \) is also a \( \mathbb{Z} \)-module homomorphism \( f : GX \to GY \). One can rightfully say that, in this example, actual equality \( (F \circ G)X = X \) holds for every \( X = (A, +) \in \text{AbGroups} \) and \( (G \circ F)Y = Y \) for every \( Y = (M, \mu) \in \mathbb{Z}\text{Mod} \); so the equivalence just constructed can in fact be called an isomorphism of categories. Similarly, \( \text{AbGroups} \) is isomorphic to the category \( \text{Mod}_\mathbb{Z} \) of right \( \mathbb{Z} \)-modules. However, inasmuch as it involves equality between objects, isomorphism of categories is generally a less useful concept than equivalence. The following examples shed more light on the true reach of the latter notion.
A.3. Skeletons

Example A.3 (Skeleta). A skeleton of a category \( C \) is any full subcategory \( S \) of \( C \) such that every object of \( C \) is isomorphic (in \( C \)) to exactly one object of \( S \). In this case, the inclusion functor \( F : S \to C \) is an equivalence of categories, because it is full, faithful and essentially surjective. Every category \( C \) has a skeleton; this follows from the Axiom of Choice applied to the equivalence relation “is isomorphic to” on the class of objects of \( C \).

For example, a skeleton for the category \( \text{Sets}_{\text{fin}} \) of all finite sets and all functions between them is given by the full subcategory \( S \) with objects \([n] = \{1, 2, \ldots, n\}\) for \( n \in \mathbb{Z}_+ \). Indeed, for every finite set \( X \), there is an isomorphism (bijection) \( X \cong [n] \) with \( n = |X| \), and \( X \) is not isomorphic to any other object of \( S \). Similarly, the category \( \text{vect}_k \) of all finite-dimensional \( k \)-vector spaces has a skeleton with objects \( k^{\otimes n} \) for \( n \in \mathbb{Z}_+ \), an isomorphism \( V \cong k^{\otimes n} \) for \( V \in \text{vect}_k \) being tantamount to choosing a basis of \( V \) (\( n = \dim_k V \)).

Of course, the isomorphisms in the above two examples are generally not unique; so we should not regard the objects in question as “the same.” Equivalence of categories formalizes the basic idea of categories being “the same for most practical purposes.”

Exercises for Section 1.3

A.3.1 (Automorphism groups). Show that, in any category, isomorphic objects have isomorphic automorphism groups. Furthermore, if \( F : C \to D \) is a functor giving an equivalence of categories (so \( F \) has a quasi-inverse), then \( F \) is full and faithful. In particular, for each \( X \in C \), there is an isomorphism \( \text{Aut}_C(X) \cong \text{Aut}_C(FX) \) in Groups.

A.4. Adjointness

The last notion from category theory that we shall mention in this appendix is that of adjointness of functors. Numerous instances of the concept occur throughout the text. In detail, two functors \( F : C \rightleftarrows D : G \) are said to be adjoint to each other if there exists an isomorphism of functors

\[
\alpha : \text{Hom}_D(F \ldots) \cong \text{Hom}_C(\ldots, G \ldots) : C^{\text{op}} \times D \to \text{Sets}
\]

Thus, for any two objects \( X \in C \) and \( Y \in D \), we have a bijection of sets

\[
\alpha_{X,Y} : \text{Hom}_D(FX, Y) \cong \text{Hom}_C(X, GY)
\]

that is natural in both \( X \) and \( Y \) in the sense discussed in Example A.2. More precisely, \( G \) is said to be right adjoint to \( F \) in this case, and \( F \) left adjoint to \( G \). Also, the isomorphism \( \alpha \) is referred as an adjunction.
One can show that any two (left or right) adjoint functors of a given functor are isomorphic (MacLane [140], p. 85), and we also refer the reader to [140], p. 93 for the connection with equivalence of categories. Finally, we remark that, in order to speak of right or left adjoints for a contravariant functor \( F : \mathbb{C} \to \mathbb{D} \), one needs to decide whether to view \( F \) as a (covariant) functor \( \mathbb{C}^{\text{op}} \to \mathbb{D} \) or \( \mathbb{C} \to \mathbb{D}^{\text{op}} \). What will be a left adjoint for one viewpoint, the other calls a right adjoint.

We content ourselves with giving just two examples of adjunctions for now.

**Example A.4** (Abelianization of groups). The abelianization of a group \( G \), by definition, is the abelian group \( G^{ab} = G/[G, G] \), where \([G, G]\) denotes the (normal) subgroup of \( G \) that is generated by the group theoretical commutators \([x, y] = x^{-1}y^{-1}xy \) with \( x, y \in G \). Any group homomorphism \( G \to A \), where \( A \) is some abelian group, contains \([G, G]\) in it kernel and so it factors uniquely through a homomorphism \( G^{ab} \to A \). Conversely, any homomorphism of abelian groups \( G^{ab} \to A \) gives rise to a group homomorphism \( G \to A \) by pulling back along the canonical epimorphism \( G \to G^{ab} \). In this way, we obtain a bijection of sets, for any group \( G \) and any abelian group \( A \),

\[
\text{Hom}_{\text{Groups}}(G, A) \cong \text{Hom}_{\text{AbGroups}}(G^{ab}, A)
\]

(A.3)

In particular, for any group homomorphism \( f : G \to H \), the map \( G \xrightarrow{f} H \to H^{ab} \) corresponds to a unique homomorphism of abelian groups, \( f^{ab} : G^{ab} \to H^{ab} \). It is easy to see that this makes abelianization into a functor,

\[
\xrightarrow{\text{ab}} : \text{Groups} \to \text{AbGroups}
\]

and that the bijection (A.3) is natural (functorial) in both \( G \) and \( A \). In fact, somewhat pedantically, we should really write \( \text{Hom}_{\text{Groups}}(G, IA) \) on the left, where \( I : \text{AbGroups} \to \text{Groups} \) is the inclusion functor. The foregoing then says that the abelianization functor is left adjoint to \( I \). Finally, both Hom-sets in (A.3) are actually abelian groups. Indeed, writing the binary operation of \( A \) as multiplication, the binary operations of the sets in (A.3) are given by “pointwise” multiplication of functions: \((fg)(x) = f(x)g(x)\) for \( x \in G \) or \( G^{ab} \). The bijection (A.3) respects these group structures; so it a natural isomorphism of abelian groups.

**Example A.5** (The free left module on a given set). For any set \( X \) and any ring \( R \), we may consider the set of finitely supported \( R \)-valued functions on \( X \),

\[
R^{(X)} \overset{\text{def}}{=} \{ f : X \to R \mid f(x) = 0 \text{ for almost all } x \in X \}
\]

With the usual pointwise addition of functions and with left \( R \)-action given by \((rf)(x) = rf(x)\) for \( r \in R \), the set \( R^{(X)} \) becomes a left \( R \)-module. Note that \( R^{(X)} \) contains the functions \( \delta_x \) that are defined by \( \delta_x(y) = \delta_{x,y} 1_R \) for \( x, y \in X \), where \( 1_R \) is the identity element of \( R \), and that every \( f \in R^{(X)} \) can be written in the form \( f = \sum_{x \in X} f(x) \delta_x \). Thus, \( R^{(X)} \) is a free \( R \)-module having the family \( (\delta_x)_{x \in X} \) as a basis. It is often notationally more convenient to consider an isomorphic copy
of $R^X$ that has $X$ itself as basis, the free left $R$-module consisting of all \textit{formal $R$-linear combinations} $\sum_{x \in X} r_x x$ with uniquely determined $r_x \in R$ that are almost all 0. Mimicking the form of its elements, we will write this module as

$$RX$$

or occasionally, if $X$ is notationally cumbersome, as $R[X]$. The familiar module theoretic fact that any $R$-module homomorphism $RX \to M$ is determined by its values on the given basis $X$, and that these values may be arbitrarily prescribed, can be concisely stated by saying that, for any $M \in R\text{Mod}$, there is an isomorphism in the category Sets (i.e., a bijection)

$$\Hom_R(RX, M) \xrightarrow{\sim} \Hom_{\text{Sets}}(X, M|_{\text{Sets}})$$

\[ (A.4) \]

where $\cdot|_{\text{Sets}}: R\text{Mod} \to \text{Sets}$ is the forgetful functor as in A.2.1 and $f|_X$ denotes the restriction of $f$ to the basis $X$ of $RX$. In particular, for any map of sets $X \to Y$, we obtain a unique $R$-module homomorphism $RX \to RY$ corresponding to $X \to Y \hookrightarrow RY|_{\text{Sets}}$ in (A.4). As in the previous example, it is straightforward to check that this results in a functor

$$R: \text{Sets} \to R\text{Mod}$$

and that the bijection (A.4) is natural (functorial) in both $X$ and $M$. Thus, the functor $R$ is left adjoint to the forgetful functor $\cdot|_{\text{Sets}}$. Finally, note that both sides of (A.4) are abelian groups under the pointwise addition of functions, using the addition of $M$, and that the bijection (A.4) respects pointwise additions. Moreover, the $R$-action on $M$ can be used to equip $\Hom_{\text{Sets}}(X, M|_{\text{Sets}})$ with the $R$-action $(rf)(x) = rf(x)$, thereby making $\Hom_{\text{Sets}}(X, M|_{\text{Sets}})$ a left $R$-module isomorphic to $M^X$, the direct product of copies of $M$ labelled by $X$. In the same way, $\Hom_R(RX, M)$ can be made into an $R$-module as well provided $R$ is commutative, and (A.4) is an isomorphism of $R$-modules in this case, natural in both $X$ and $M$. 
Background from Linear Algebra

In this appendix, we summarize the basic properties of tensor products and of the trace function. The reader wishing to see the detailed construction of tensor products or proofs for any of the numerous facts stated here without further justification is referred to her or his favorite textbook on algebra; the most encyclopedic reference on this material is Bourbaki [24, Chapters 2 and 3].

We assume that the reader is reasonably familiar with Appendix A.

B.1. Tensor Products

B.1.1. The Basics

Let \( R \) be any ring (with 1) and let \( M \in \text{Mod}_R \) and \( N \in _R \text{Mod} \) be given. Then the tensor product \( M \otimes_R N \) is an abelian group, with binary operation +, that is equipped with a map, called “canonical,”

\[
M \times N \xrightarrow{\text{can}} M \otimes_R N
\]

\[
\begin{align*}
\omega & : (m, n) & \to m \otimes n \\
\psi & : m \otimes n & \to m \\otimes n
\end{align*}
\]

satisfying the following rules, for all \( m, m' \in M, n, n' \in N \) and \( r \in R \),

\[
\begin{align*}
(m + m') \otimes n &= m \otimes n + m' \otimes n \\
m \otimes (n + n') &= m \otimes n + m \otimes n' \\
m \cdot r \otimes n &= m \otimes r \cdot n
\end{align*}
\]
Viewing abelian groups as \( \mathbb{Z} \)-modules (§A.3.3), the first two rules state that the canonical map is \( \mathbb{Z} \)-bilinar: it is a group homomorphism when one of the two arguments is considered variable while the other one is fixed. A \( \mathbb{Z} \)-bilinar map \( \gamma: M \times N \rightarrow G \), where \( G \) is some abelian group, is said to be \( R \)-balanced if \( \gamma \) also has the third basic property of the canonical map: \( \gamma(m, r, n) = \gamma(m, r \cdot n) \) for all \( m \in M, n \in N \) and \( r \in R \). The crucial property of the tensor product is that the canonical map is the “universal” such map: any \( R \)-balanced map \( \gamma: M \times N \rightarrow G \) factors uniquely through the canonical map as in the diagram

\[
\begin{array}{ccc}
M \times N & \xrightarrow{\gamma} & G \\
\text{can.} \downarrow & & \downarrow \exists! \tilde{\gamma} \\
M \otimes_R N & \xrightarrow{\tilde{\gamma}} & \\
\end{array}
\]

So \( \tilde{\gamma}: M \otimes_R N \rightarrow G \) is the unique group homomorphism such that \( \gamma(m, n) = \tilde{\gamma}(m \otimes n) \) for all \( m \in M \) and \( n \in N \). The universal property (B.1) characterizes the tensor product up to isomorphism and it easily implies the following facts.

**Functoriality.** Given morphisms \( f: M \rightarrow M' \) in \( \text{Mod}_R \) and \( g: N \rightarrow N' \) in \( \text{RMod} \), there is a group homomorphism

\[
f \otimes g: \quad M \otimes_R N \xrightarrow{\sim} M' \otimes_R N'
\]

(B.2)

In place of \( \text{Id}_M \otimes g \) one often writes \( M \otimes g \) and similarly \( f \otimes N = f \otimes \text{Id}_N \). In this way, the tensor product gives a functor (or bifunctor)

\[
\cdot \otimes \cdot : \text{Mod}_R \times \text{RMod} \rightarrow \text{AbGroups}
\]

and functors \( M \otimes_R \cdot : \text{RMod} \rightarrow \text{AbGroups} \) and \( \cdot \otimes_R N : \text{Mod}_R \rightarrow \text{AbGroups} \). These functors commutes with direct sums: for any families of modules \( M_i \in \text{Mod}_R \) (\( i \in I \)) and \( N_j \in \text{RMod} \) (\( j \in J \)), there is a natural isomorphism of abelian groups

\[
\left( \bigoplus_{i \in I} M_i \right) \otimes_R \left( \bigoplus_{j \in J} N_j \right) \xrightarrow{\sim} \bigoplus_{(i,j) \in I \times J} (M_i \otimes_R N_j)
\]

(B.3)

Furthermore, viewing \( R \) as a left and right module over itself via multiplication, we have natural isomorphisms of abelian groups, for any \( M \in \text{Mod}_R \) and any
B.1. Tensor Products

\[ N \in \text{RMod}, \]
\[ M \otimes_{\text{R}} \text{R} \xrightarrow{\sim} M \quad \text{and} \quad R \otimes_{\text{R}} N \xrightarrow{\sim} N \]
\[ (B.4) \]
\[ m \otimes r \longmapsto m.r \quad \text{and} \quad r \otimes n \longmapsto r.n \]

In both (B.3) and (B.4), “natural isomorphism” is to be understood as in §A.3.2. For example, the first isomorphism in (B.4) comes from an isomorphism of functors \( \cdot \otimes_{\text{R}} \text{R} \equiv F : \text{Mod}_R \to \text{AbGroups} \) where \( F \) is the forgetful functor of §A.2.1.

**Tensor products with free modules: normal form of elements.** If \( M \in \text{Mod}_R \) is a free \( R \)-module, say with basis \( (b_i)_{i \in I} \), then \( M = \bigoplus_{i \in I} b_i \text{R} \) and \( b_i \text{R} \cong \text{R} \) in \( \text{Mod}_R \) for all \( i \). Therefore, it follows from (B.3) and (B.4) that, for any \( N \in \text{RMod} \), we have isomorphisms
\[ M \otimes_{\text{R}} N \cong \bigoplus_{i \in I} (b_i \text{R} \otimes_{\text{R}} N) \quad \text{and} \quad b_i \text{R} \otimes_{\text{R}} N \cong \text{R} \otimes_{\text{R}} N \cong N \]
for all \( i \). Consequently, every element of \( M \otimes_{\text{R}} N \) has the form
\[ \sum_{i \in I} b_i \otimes n_i \quad \text{with unique} \ n_i \in N \quad \text{that are almost all} \ 0 \]

(B.5)

Similarly, if \( N \in \text{RMod} \) is free with basis \( (c_j)_{j \in J} \), then every element of \( M \otimes_{\text{R}} N \) has the form \( \sum_{j \in J} m_j \otimes c_j \) with unique \( m_j \in M \) that are almost all \( 0 \).

**Exactness properties.** Let \( 0 \to U \xrightarrow{f} V \xrightarrow{g} W \to 0 \) be a short exact sequence in \( \text{RMod} \). Here, “exactness” of the sequence means that, for each module with an incoming and an outgoing arrow, the image of the incoming map equals the kernel of the outgoing map: \( f \) is injective, \( g \) is surjective, and \( \text{Im} f = \text{Ker} g \). Applying the functor \( M \otimes_{\text{R}} \cdot : \text{RMod} \to \text{AbGroups} \) for a given \( M \in \text{Mod}_R \) to the above short exact sequence, the resulting sequence in \( \text{AbGroups} \) will generally not be exact. Specifically, the map \( \text{Id}_M \otimes f \) need not be injective, but at least the sequence
\[ M \otimes_{\text{R}} U \xrightarrow{\text{Id}_M \otimes f} M \otimes_{\text{R}} V \xrightarrow{\text{Id}_M \otimes g} M \otimes_{\text{R}} W \xrightarrow{} 0 \]
is always exact. This property of the functor \( M \otimes_{\text{R}} \cdot \) is called right exactness. The module \( M \) is said to be flat if the functor \( M \otimes_{\text{R}} \cdot \) is in fact exact in the sense that it turns short exact sequences into short exact sequences. All free modules and, more generally, all projective modules (Section 2.1) are flat. In particular, all modules over division rings are flat. All this also applies mutatis mutandis to the functors \( \cdot \otimes_{\text{R}} N: \text{Mod}_R \to \text{AbGroups} \) with \( N \in \text{RMod} \).

B.1.2. Additional Structure from Bimodules

Let \( (R, S) \) be a pair of rings. An \((R, S)\)-bimodule is an abelian group \( M \) that is both a left \( R \)-module and a right \( S \)-module such that the two actions commute in
the sense that, for all \( r \in R, s \in S \) and \( m \in M \),

\[(r.m).s = r.(m.s)\]

Morphisms (or homomorphisms) between \((R, S)\)-bimodules are maps that are both left \(R\)-module and right \(S\)-module homomorphisms. Thus, we have a category, \(R \text{Mod}_S\).

Identifying abelian groups with (right or left) \(Z\)-modules as in §A.3.3, we may identify \(R \text{Mod}_S\) with \(R \text{Mod}_Z\) and \(S \text{Mod}_R\) with \(Z \text{Mod}_R\). In particular, the category \(\text{AbGroups}\) is isomorphic to \(Z \text{Mod}_Z\). We will use the notation \(R M_S\) to indicate that \(M \in R \text{Mod}_S\).

Given bimodules \(R M_S\) and \(S N_T\), the tensor product \(M \otimes_S N\) becomes an \((R, T)\)-bimodule via the actions

\[r.(m \otimes n) := r.m \otimes n \quad \text{and} \quad (m \otimes n).t := m \otimes n.t\]

For any morphisms \(f : M \rightarrow M'\) in \(R \text{Mod}_S\) and \(g : N \rightarrow N'\) in \(S \text{Mod}_T\), the group homomorphism \(f \otimes g : M \otimes_S N \rightarrow M' \otimes_S N'\) in (B.2) is evidently a morphism in \(R \text{Mod}_T\). Thus, the tensor product gives a (bi-)functor

\[\cdot \otimes_S \cdot : R \text{Mod}_S \times S \text{Mod}_T \rightarrow R \text{Mod}_T\]

The earlier functor in §B.1.1 is the special case \(R = T = Z\) and \(S = R\) and its basic properties extend to the current setting: (B.3) gives, in the same way, a natural isomorphism in \(R \text{Mod}_T\), for any \(M_i \in R \text{Mod}_S\) and \(N_j \in S \text{Mod}_T\).

\[(\bigoplus_{i \in I} M_i) \otimes_S \left( \bigoplus_{j \in J} N_j \right) \cong \bigoplus_{(i,j) \in I \times J} (M_i \otimes_R N_j)\]

Furthermore, viewing \(R\) as the regular \((R, R)\)-bimodule in (B.4), with right and left \(R\)-action both given by multiplication, we obtain isomorphisms of functors

\[(B.7) \quad \cdot \otimes_R \cdot \cong \text{Id}_{R \text{Mod}_R} \quad \text{and} \quad R \otimes_R \cdot \cong \text{Id}_{R \text{Mod}_S}\]

In addition, the tensor product of bimodules is associative: for any bimodules \(R L_S, S M_T\) and \(T N_U\), there is an isomorphism of \((R, U)\)-bimodules

\[(B.8) \quad (L \otimes_S M) \otimes_T N \cong L \otimes_S (M \otimes_T N)\]

which is natural in the sense that it comes from an isomorphism of functors (§A.3.2)

\[(\cdot \otimes_S \cdot) \otimes_T \cdot \cong \cdot \otimes_S (\cdot \otimes_T \cdot) : R \text{Mod}_S \times S \text{Mod}_T \times T \text{Mod}_U \rightarrow R \text{Mod}_U\]
B.1.3. Tensor Powers and Multilinear Maps

In this subsection, \( R \) denotes a commutative ring. Any right \( R \)-module \( M \in \text{Mod}_R \) can be made into a \((R,R)\)-bimodule by defining the left \( R \)-action to be the same as the given right action: \( r \cdot m := m \cdot r \). This yields a functor \( F : \text{Mod}_R \rightarrow \text{Mod}_R \). In the other direction, we have the functor \( G : \mathcal{R} \text{Mod}_R \rightarrow \text{Mod}_R \) that forgets the left \( R \)-action; so \((G \circ F)M = M\). In this way, we can identify \( \text{Mod}_R \) with the full subcategory of \( \mathcal{R} \text{Mod}_R \) consisting of all \((R,R)\)-bimodules having identical left and right \( R \)-operations. Similar things do of course apply to \( \mathcal{R} \text{Mod} \); so we have equivalences of categories

\[
\text{Mod}_R \cong \mathcal{R} \text{Mod}
\]

We will freely view one-sided \( R \)-modules as \((R,R)\)-bimodules and just speak of \( R \)-modules. Then the tensor product \( M \otimes_R N \) of \( R \)-modules \( M \) and \( N \) is an \( R \)-module again: \( r.(m \otimes n) = r.m \otimes n = m \otimes r.n \). Furthermore, (B.6), (B.7) and (B.8) with \( S = T = R \) all give isomorphisms of \( R \)-modules, and \( M \otimes_R N \cong N \otimes_R M \) as \( R \)-modules via the switch map, \( m \otimes n \leftrightarrow n \otimes m \).

**Free Modules.** Let \( V \) be a free \( R \)-module with basis \((b_i)_{i \in I}\); so \( V = \bigoplus_{i \in I} Rb_i \) and \( Rb_i \cong R \) as \( R \)-modules. Choosing a ring homomorphism \( \phi : R \rightarrow F \), where \( F \) is some field, say the factor of \( R \) modulo some maximal ideal, we can view \( F \) as an \((F,R)\)-bimodule: \( f'.f.r := f'.f(\phi(r)) \). Then \( F \otimes_R V \) becomes an \( F \)-vector space and it follows from (B.6) and (B.7) that \( F \otimes_R V \cong \bigoplus_{i \in I} (F \otimes_R Rb_i) \) and \( F \otimes_R Rb_i \cong F \) as \( F \)-vector spaces. Consequently, \( \dim_F F \otimes_R V = |I| \), proving that the cardinality of a basis of \( V \) is an invariant of \( V \); it is called the rank of \( V \).

It follows from our remarks about normal form of elements in §B.1.1 that if \( V \) and \( W \) are free \( R \)-modules with respective \( R \)-bases \((b_i)_{i \in I}\) and \((c_j)_{j \in J}\), then \( V \otimes_R W \) is a free \( R \)-module with \( R \)-basis \((b_i \otimes c_j)_{(i,j) \in I \times J}\). In particular,\n
\[
\text{rank}(V \otimes_R W) = (\text{rank } V)(\text{rank } W) \tag{B.9}
\]

**Iterated Tensor Products.** For any given \( R \)-modules \( V_1, \ldots, V_k \) \((k \geq 2)\), we define the iterated tensor product by

\[
V_1 \otimes_R V_2 \otimes_R \cdots \otimes_R V_k \overset{\text{def}}{=} (\cdots (V_1 \otimes_R V_2) \otimes_R \cdots) \otimes_R V_k.
\]

This is an \( R \)-module, and a different choice of bracketing on the right would result in a naturally isomorphic \( R \)-module by (B.8). Elements of the iterated tensor product will be also written without brackets:

\[
v_1 \otimes v_2 \otimes \cdots \otimes v_k = (\cdots (v_1 \otimes v_2) \otimes \cdots) \otimes v_k.
\]

The associativity isomorphism (B.8) is oftentimes implicit in the construction of certain maps. For example, if \( f : V_1 \rightarrow W_1 \) and \( g : V_2 \otimes_R V_3 \rightarrow W_2 \) are \( R \)-module

\[^{1}\text{For noncommutative rings, there is generally no meaningful notion of rank for free modules; see Cohn [46]. However, the argument above works for all rings having a homomorphism into some division ring.}\]

maps, then \( f \otimes g : V_1 \otimes_R V_2 \otimes_R V_3 \to W_1 \otimes W_2 \) is understood to be the map 
\[
(f \otimes g)(v_1 \otimes v_2 \otimes v_3) = f(v_1) \otimes g(v_2 \otimes v_3),
\]
which is in fact the composite of the associativity isomorphism 
\[
V_1 \otimes_R (V_2 \otimes_R V_3) = (V_1 \otimes_R V_2) \otimes_R V_3 \cong V_1 \otimes_R (V_2 \otimes_R V_3)
\]
and the actual map \( f \otimes g : V_1 \otimes_R (V_2 \otimes_R V_3) \to W_1 \otimes_R W_2 \).

**Tensor Powers.** We will often be concerned with the case where all \( V_i \) are equal
(or isomorphic) to some \( R \)-module \( V \). The tensor powers \( V^{\otimes k} \) are then given by
\[
V^{\otimes 0} \overset{\text{def}}{=} R, \quad V^{\otimes 1} \overset{\text{def}}{=} V \quad \text{and} \quad V^{\otimes k} \overset{\text{def}}{=} \underbrace{V \otimes_R \cdots \otimes_R V}_{k \geq 2}
\]
By virtue of (B.7) and (B.8) we have \( V^{\otimes k} \otimes V^{\otimes l} \cong V^{\otimes (k+l)} \) as \( R \)-modules; for
\( k, l \geq 1 \), this associativity isomorphism is given by
\[
\begin{array}{ccc}
V^{\otimes k} \otimes V^{\otimes l} & \sim & V^{\otimes (k+l)} \\
(v_1 \otimes \cdots \otimes v_k) \otimes (v'_1 \otimes \cdots \otimes v'_l) & \mapsto & v_1 \otimes \cdots \otimes v_k \otimes v'_1 \otimes \cdots \otimes v'_l
\end{array}
\]
The tensor powers give functors \( \cdot^{\otimes k} : \text{Mod}_R \to \text{Mod}_R \): for any morphism \( f : V \to W \) in \( \text{Mod}_R \) (i.e., any \( R \)-module homomorphism), \( f^{\otimes k} : V^{\otimes k} \to W^{\otimes k} \) is given by
\[
f^{\otimes k}(v_1 \otimes \cdots \otimes v_k) = f(v_1) \otimes \cdots \otimes f(v_k).
\]

**Multilinear Maps.** Let \( U, V \) and \( W \) be \( R \)-modules. Recall that a map \( \beta : U \times V \to W \) is called \( R \)-balanced if \( \beta \) is additive in both arguments and \( \beta(r.u, v) = \beta(u, r.v) \)
for all \( u \in U, v \in V \) and \( r \in R \); the map \( \beta \) is said to be **\( R \)-bilinear** if, in
addition, \( \beta(r.u, v) = \beta(u, r.v) = r.\beta(u, v) \), that is, \( \beta \) is an \( R \)-module map for both
inputs. By the universal property (B.1) of tensor products, the set of all \( R \)-balanced
maps \( \beta : U \times V \to W \) is in bijection with the set of homomorphisms of abelian
groups \( \beta : U \otimes_R V \to W \). This bijection restricts to a bijection between the subsets
consisting of the \( R \)-bilinear maps on the one side and the \( R \)-module homomorphisms
on the other:
\[
\text{BiLin}(U, V; W) \cong \text{Hom}_R(U \otimes_R V, W)
\]
The \( R \)-module structure of \( W \) imparts standard \( R \)-module structures to both sets in
(B.12), given by “pointwise” addition and \( R \)-multiplication of functions, and (B.12)
is in fact an isomorphism of \( R \)-modules.

More generally, a map \( \mu : V_1 \times \cdots \times V_k \to W \) for \( R \)-modules \( V_1, \ldots, V_k \) and \( W \)
is called **\( R \)-multilinear** if all \( \mu(v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_k) \) are \( R \)-module maps. As
above, the set consisting of all these maps forms an \( R \)-module that is isomorphic to
the \( R \)-module \( \text{Hom}_R(V_1 \otimes_R \cdots \otimes_R V_k, W) \). In particular, if \( V_1 = \cdots = V_k =: V \),
then we obtain an isomorphism of $R$-modules

\[
\text{MultLin}(V^n, W) \cong \text{Hom}_R(V^\otimes n, W)
\]

\[\mu \mapsto \overline{\mu} = (v_1 \otimes \cdots \otimes v_k \mapsto \mu(v_1, \ldots, v_k))\]

\[\text{(B.13)}\]

**B.2. Hom-$\otimes$ Relations**

**B.2.1. The Hom-Functor**

Let $R, S$ and $T$ be arbitrary rings and let $_R M_S$ and $_R N_T$ be bimodules. Then the group $\text{Hom}_R(M, N)$ of all left $R$-module maps $\phi: M \to N$ becomes an $(S, T)$-bimodule by defining

\[(s, \phi, t)(m) := \phi(m.s) \cdot t\]

For any morphisms $f: M' \to M$ in $_R \text{Mod}_S$ and $g: N \to N'$ in $_R \text{Mod}_T$, the map $\text{Hom}_R(f, g): \text{Hom}_R(M, N) \to \text{Hom}_R(M', N')$, $\phi \mapsto g \circ \phi \circ f$ is a morphism in $_S \text{Mod}_T$. This yields a (bi-)functor

\[\text{Hom}_R(\cdot, \cdot): \_R \text{Mod}_S \times _R \text{Mod}_T \to _S \text{Mod}_T\]

This functor is covariant in the second variable but contravariant in the first. In an analogous fashion, one obtains a functor

\[\text{Hom}_R(\cdot, \cdot): _S \text{Mod}_R \times _T \text{Mod}_R \to _T \text{Mod}_S\]

In either case, the morphism $\text{Hom}_R(\text{Id}_M, g)$ is often denoted by $\text{Hom}_R(M, g)$ or $g_*$ and one also writes $\text{Hom}_R(f, \text{Id}_N) = \text{Hom}_R(f, N) = f^*$.  

**Left Exactness.** Let $0 \to N' \xrightarrow{f} N \xrightarrow{g} N''$ be an exact sequence in $_R \text{Mod}_T$ and let $_R M_S$ be given. Then the following is an exact sequence in $_S \text{Mod}_T$:

\[0 \to \text{Hom}_R(M, N') \xrightarrow{f^*} \text{Hom}_R(M, N) \xrightarrow{g_*} \text{Hom}_R(M, N'')\]

This property of the functor $\text{Hom}_R(M, \cdot)$ is called **left exactness**.

Similarly, for any $_R N_T$, the contravariant functor $\text{Hom}_R(\cdot, N)$, is also left exact when viewed as a covariant functor $\text{Hom}_R(\cdot, N): _R \text{Mod}_S^{\text{op}} \to _T \text{Mod}_S$. An exact sequence $0 \to M' \to M \to M''$ in $_R \text{Mod}_S^{\text{op}}$ corresponds to an exact sequence $M'' \xrightarrow{f^*} M \xrightarrow{g_*} M' \to 0$ in $_R \text{Mod}_S$ and the resulting $_T \text{Mod}_S$-sequence

\[0 \to \text{Hom}_R(M', N) \xrightarrow{g_*} \text{Hom}_R(M, N) \xrightarrow{f^*} \text{Hom}_R(M'', N)\]

is again exact. All this also holds for $\text{Hom}_R(\cdot, \cdot): _S \text{Mod}_R \times _T \text{Mod}_R \to _T \text{Mod}_S$. 

Direct Sums and Products. Let $M_i \in R\text{Mod}_S \ (i \in I)$ and $N_j \in R\text{Mod}_T \ (j \in J)$ be families of bimodules. Then the following is an isomorphism in $S\text{Mod}_T$:

$$\Hom_R(\bigoplus_{i \in I} M_i, \prod_{j \in J} N_j) \xrightarrow{\sim} \prod_{j \in J} \Hom_R(M_i, N_j)$$  \hspace{1cm} (B.14)

Here, $\mu_i : M_i \to \bigoplus_{i' \in I} M_{i'}$, $m \mapsto (\delta_{i,i'} m)_{i'}$ and $\pi_j : \prod_{j' \in J} N_{j'} \to N_j$, $(n_{j'})_{j'} \mapsto n_j$ are the canonical embedding and projection maps. Since finite direct sums and direct products are isomorphic, (B.14) says in particular that the functors $\Hom_R(M, \cdot)$ and $\Hom_R(\cdot, N)$ commute with finite direct sums.

B.2.2. Hom-⊗ Adjunction

Let $R$, $S$, $T$ and $U$ be rings. For given bimodules $R M_S$, $S L_T$ and $U N_T$, we obtain bimodules $R(M \otimes_S L)_T$ and $U(\Hom_T(L, N))_S$ and then $\Hom_T(M \otimes_S L, N)$ and $\Hom_S(M, \Hom_T(L, N))$ in $U\text{Mod}_R$. The latter two bimodules are naturally isomorphic via

$$\Hom_T(M \otimes_S L, N) \xrightarrow{\sim} \Hom_S(M, \Hom_T(L, N))$$  \hspace{1cm} (B.15)

Similarly, for bimodules $R M_S$, $S L_T$ and $R N_U$, the following map is a natural isomorphism in $T\text{Mod}_U$,

$$\Hom_R(M \otimes_S L, N) \xrightarrow{\sim} \Hom_S(L, \Hom_R(M, N))$$  \hspace{1cm} (B.16)

B.3. Vector Spaces

We will now leave the general ring theoretic setting behind and concentrate on the category $\text{Vect}_k$ of vector spaces over a given base field $k$ with $k$-linear maps as morphisms. Thus, the material and the conventions of §B.1.3 will apply; in particular, $k$-vector spaces are $(k,k)$-bimodules with identical left and right $k$-actions.
B.3.1. The Bifunctor \( \otimes \)

We will simply write \( \otimes = \otimes_k \) throughout; this is a bifunctor

\[
\cdot \otimes \cdot : \text{Vect}_k \times \text{Vect}_k \longrightarrow \text{Vect}_k
\]

which is exact in either argument (§B.1.1). Furthermore, \( V \otimes W \equiv W \otimes V \) for any \( V, W \in \text{Vect}_k \) and the rank formula (B.9) becomes

\[
\dim_k(V \otimes W) = (\dim_k V)(\dim_k W).
\]

By (B.2), any pair of morphisms \((\alpha, \alpha') \in \text{Hom}_k(V, W) \times \text{Hom}_k(V', W')\) gives rise to the morphism \(\alpha \otimes \alpha' \in \text{Hom}_k(V \otimes V', W \otimes W')\) that is given by

\[
(\alpha \otimes \alpha')(v \otimes v') = \alpha(v) \otimes \alpha'(v').
\]

This yields a natural \(k\)-linear map

\[
\text{(B.17)} \quad \text{Hom}_k(V, W) \otimes \text{Hom}_k(V', W') \longrightarrow \text{Hom}_k(V \otimes V', W \otimes W').
\]

The map (B.17) is always injective and it is an isomorphism if at least one of the pairs \((V, V')\) or \((V, W)\) or \((V', W')\) consists of finite-dimensional \(k\)-vector spaces; see [24, Chap. II §7 Prop. 16].

B.3.2. The Bifunctor \( \text{Hom}_k \) and the Linear Dual

Similarly, the bifunctor

\[
\text{Hom}_k(\cdot, \cdot) : \text{Vect}^{\text{op}}_k \times \text{Vect}_k \longrightarrow \text{Vect}_k
\]

is exact in either argument, not merely left exact, due to the existence of complements for subspaces. In many cases, this bifunctor can be described in terms of \( \otimes \) and the so-called linear dual; the latter is the (contravariant and exact) functor

\[
\cdot^* \quad \text{def} \quad \text{Hom}_k(\cdot, k) : \text{Vect}_k \to \text{Vect}_k
\]

The elements of \( V^* \) are called linear forms on \( V \). We will use the notation

\[
\langle \cdot, \cdot \rangle : V^* \times V \to k
\]

for the evaluation map. If \( V \) is finite dimensional, then \( \dim_k V = \dim_k V^* \); if \((v_i)_{i=1}^n\) is a \(k\)-basis of \( V \), then \( V^* \) has the dual basis \((f_i)_{i=1}^n\) that is defined by \(\langle f_i, v_j \rangle = \delta_{i,j}\).

For infinite-dimensional \( V \), one has \( \dim_k V < \dim_k V^* \) by the Erdős-Kaplansky Theorem [109, p. 247].

Recall that, for a \(k\)-linear map \( \alpha : V \to W \), the morphism \( \alpha^* : W^* \to V^* \) is given by \(\alpha^*(f) = f \circ \alpha\). The map \(\alpha^*\) is also called the transpose of \( \alpha \), because the matrices of \( \alpha \) and \( \alpha^* \) with respect to dual bases for \( V, V^* \) and \( W, W^* \) are transposes of each other.
**Homomorphisms and Tensor Products.** For any \( V, W \in \text{Vect}_k \), there is a natural \( k \)-linear monomorphism

\[
W \otimes V^* \hookrightarrow \text{Hom}_k(V, W)
\]

(B.18)

\[
w \otimes f \mapsto (v \mapsto \langle f, v \rangle w)
\]

To check injectivity, write a given \( 0 \neq x \in W \otimes V^* \) as a finite sum \( x = \sum_i w_i \otimes f_i \) with \( \{w_i\} \) linearly independent and nonzero \( f_i \in V^* \) and observe that the image of \( x \) in \( \text{Hom}_k(V, W) \) is plainly nonzero again. The image of the map (B.18) consists exactly of all \( k \)-linear maps \( V \rightarrow W \) that have finite rank (i.e., the image is a finite-dimensional subspace of \( W \)). In particular, if at least one of \( V \) or \( W \) is finite-dimensional, then (B.18) gives an isomorphism in \( \text{Vect}_k \), which will be referred to as “canonical,”

\[
W \otimes V^* \cong \text{Hom}_k(V, W)
\]

(B.19)

For \( U, V, W \in \text{Vect}_k \), (B.15) gives \( \text{Hom}_k(U \otimes V, W) \cong \text{Hom}_k(U, \text{Hom}_k(V, W)) \) in \( \text{Vect}_k \). If \( \dim_k V < \infty \) or \( \dim_k W < \infty \), then (B.19) further implies that

\[
\text{Hom}_k(U \otimes V, W) \cong \text{Hom}_k(U, W \otimes V^*)
\]

(B.20)

as \( k \)-vector spaces. In particular, \( \text{Hom}_k(V \otimes W, k) \cong \text{Hom}_k(V, k \otimes W^*) \cong \text{Hom}_k(V, W^*) \cong W^* \otimes V^* \). Explicitly,

\[
W^* \otimes V^* \xrightarrow{\sim} (V \otimes W)^*
\]

(B.21)

\[
g \otimes f \mapsto (v \otimes w \mapsto \langle f, v \rangle \langle g, w \rangle)
\]

This isomorphism is also a consequence of (B.17).

**The Double Dual.** For any \( V \in \text{Vect}_k \), there is a natural \( k \)-linear monomorphism

\[
V \hookrightarrow V^{**}
\]

(B.22)

\[
v \mapsto (f \mapsto \langle f, v \rangle)
\]

Injectivity follows from the fact that any \( 0 \neq v \in V \) is part of some \( k \)-basis of \( V \), and hence there is a linear form \( f \in V^* \) with \( \langle f, v \rangle \neq 0 \). The map (B.22) is an isomorphism for finite-dimensional \( V \), because \( \dim_k V = \dim_k V^* = \dim_k V^{**} \) in this case.
B.3.3. The Trace Map

Let $V \in \text{Vect}_k$ be finite dimensional. Then (B.18), (B.19) give the canonical isomorphism $V \otimes V^* \cong \text{End}_k(V)$. Viewing this isomorphism as an identification, the trace map is the linear form on $\text{End}_k(V)$ that is defined by evaluation:

$$\text{trace}: \text{End}_k(V) \xrightarrow{\sim} V \otimes V^* \xrightarrow{\text{can.}} k$$

(B.23)

To see that this is indeed the usual trace map, let $(v_i)$ be a basis of $V$ and let $(f_i)$ be the dual basis of $V^*$ (§B.3.2). Then the inverse of the isomorphism (B.18) is easily seen to be the map $\text{End}_k(V) \to V \otimes V^*$, $\alpha \mapsto \sum_i \alpha(v_i) \otimes f_i$. Therefore, (B.23) is explicitly given by $\text{trace}(\alpha) = \sum_i \langle f_i, \alpha(v_i) \rangle$, which is the sum of the diagonal entries of the matrix of $\alpha$ with respect to the chosen basis of $V$.

Composites. The most crucial trace property is expressed by the formula

$$\text{trace}(\alpha \circ \beta) = \text{trace}(\beta \circ \alpha)$$

(B.24)

for any morphisms $\alpha \in \text{Hom}_k(V, W)$ and $\beta \in \text{Hom}_k(W, V)$ with $V, W \in \text{Vect}_k$ finite dimensional. To derive this formula from (B.23), we may assume by linearity that both $\alpha$ and $\beta$ have rank at most 1. Thus, viewing the isomorphisms $\text{Hom}_k(V, W) \cong W \otimes V^*$ and $\text{Hom}_k(W, V) \cong V \otimes W^*$ from (B.19) as identifications, we may write $\alpha = w \otimes f$ with $w \in W$, $f \in V^*$ and similarly $\beta = v \otimes g$. Then the composites are given by “contraction in the middle,”

$$\alpha \circ \beta = (w \otimes f) \circ (v \otimes g) = w \otimes \langle f, v \rangle g$$

and similarly $\beta \circ \alpha = v \otimes \langle g, w \rangle f$. Therefore, evaluation gives $\text{trace}(\alpha \circ \beta) = \langle f, v \rangle \langle g, w \rangle = \text{trace}(\beta \circ \alpha)$.

Transposes. It follows from our remarks in §B.3.2 about the matrix of a transpose that, for any $\alpha \in \text{End}_k(V)$,

$$\text{trace}(\alpha) = \text{trace}(\alpha^*)$$

(B.25)

Alternatively, with the usual identifications $\text{End}_k(V) = V \otimes V^*$ and $\text{End}_k(V^*) = V^* \otimes V^{**}$ and identifying $V^{**}$ with $V$ as in (B.22), the map $(\cdot)^* : \text{End}_k(V) \to \text{End}_k(V^*)$ simply becomes the switch map $V \otimes V^* \to V^* \otimes V$, $v \otimes f \mapsto f \otimes v$. Thus, (B.25) is also evident from (B.23).

Tensor products. The following trace formula, for $\alpha \in \text{End}_k(V)$ and $\beta \in \text{End}_k(W)$ and finite-dimensional $V, W \in \text{Vect}_k$, will be useful:

$$\text{trace}(\alpha \otimes \beta) = \text{trace}(\alpha) \text{trace}(\beta)$$

(B.26)
This can be proved by considering matrices: the matrix of \( \alpha \otimes \beta \in \text{End}_k(V \otimes W) \), with respect to a certain basis of \( V \otimes W \) that is assembled from bases of \( V \) and \( W \), is the so-called “Kronecker product” of the matrices of \( \alpha \) and \( \beta \). Instead, let us use the description (B.23) of the trace map. By linearity, we may again assume that both \( \alpha \) and \( \beta \) have rank at most 1, as in the preceding paragraph; so \( \alpha(x) = \langle f, x \rangle_v \) with \( v \in V \), \( f \in V^* \) and similarly \( \beta(y) = \langle g, y \rangle_w \). Then \( \langle f, v \rangle = \text{trace}(\alpha) \) and \( \langle g, w \rangle = \text{trace}(\beta) \) and, using the isomorphism (B.21), we have

\[
\text{trace}: \quad \text{End}_k(V \otimes W) \rightarrow (V \otimes W) \otimes (W^* \otimes V^*) \rightarrow k
\]

proving (B.26).

**B.3.4. Extensions of the Base Field**

Let \( K/k \) be a field extension. We may view \( K \) as a \((K,k)\)-bimodule via multiplication: \( \mu \cdot \nu \cdot \lambda = \mu \nu \lambda \) for \( \mu, \nu \in K \) and \( \lambda \in k \). Hence, for any \( V \in \text{Vect}_k \), the tensor product \( K \otimes V \) becomes a \( K \)-module, that is, a \( K \)-vector space, as in §B.1.2. This yields the field extension functor

\[
K \otimes \cdot : \text{Vect}_k \rightarrow \text{Vect}_K
\]

It follows from the remarks about normal form of elements in §B.1.1 that if \( V \) has \( k \)-basis \( \{v_i\} \), then \( (1 \otimes v_i) \) is a \( K \)-basis for \( K \otimes V \). In particular,

\[
\dim_K(K \otimes V) = \dim_k V
\]

As for morphisms, note that each \( \mu \in K \) gives a \((K,k)\)-bimodule endomorphism \( \mu_K \) of \( K \) via multiplication: \( \mu_K(\nu) = \mu \nu \) for \( \nu \in K \). Therefore, for any \( f \in \text{Hom}_k(V, W) \), the map \( \mu_K \otimes f : K \otimes V \rightarrow K \otimes W \) is \( K \)-linear. This yields a \( K \)-linear monomorphism, again often called “canonical,”

\[
K \otimes \text{Hom}_k(V, W) \otimes \text{Hom}_K(K \otimes V, K \otimes W) \rightarrow \mu_K \otimes f
\]

For injectivity, write \( 0 \neq x \in K \otimes \text{Hom}_k(V, W) \) as a finite sum \( x = \sum_i \mu_i \otimes f_i \) with \( k \)-linearly independent \( \{\mu_i\} \) and all \( f_i \neq 0 \) and pick \( v \in V \) with \( f_i(v) \neq 0 \) for some \( i \) to obtain \( (\sum_i (\mu_i)_K \otimes f_i)(1 \otimes v) = \sum_i \mu_i \otimes f_i(v) \neq 0 \). One can also show that (B.27) is bijective if \( V \) is finite dimensional or the field extension \( K/k \) is finite; see [24, Chap. II §7 Prop. 22]. The map \( 1_K \otimes f = \text{Id}_K \otimes f \) is generally written as \( K \otimes f \).
Appendix C

Some Commutative Algebra

C.1. The Nullstellensatz

Let \( k[t_1, \ldots, t_n] \) denote the (commutative) polynomial algebra over the field \( k \) and let \( \overline{k} \) denote an algebraic closure of \( k \). Each polynomial \( f \in k[t_1, \ldots, t_n] \) gives rise to a \( \overline{k} \)-valued function on affine \( n \)-space \( \overline{k}^n \) by viewing the variables \( t_i \) as the coordinate functions: \( t_i(x) = A_i \) for \( x = (A_1, \ldots, A_n) \in \overline{k}^n \). Thus, for any subset \( I \subseteq k[t_1, \ldots, t_n] \), we may consider the zero set or vanishing set

\[ V(I) \overset{\text{def}}{=} \{ x \in \overline{k}^n \mid f(x) = 0 \text{ for all } f \in I \} \]

Apparently, the ideal of \( k[t_1, \ldots, t_n] \) that is generated by \( I \) has the same vanishing set as \( I \) itself, and so we may as well assume at the outset that \( I \) is an ideal when considering \( V(I) \). In the opposite direction, for each subset \( X \subseteq \overline{k}^n \), we may define an ideal of \( k[t_1, \ldots, t_n] \) by

\[ I(X) \overset{\text{def}}{=} \{ f \in k[t_1, \ldots, t_n] \mid f(x) = 0 \text{ for all } x \in X \} \]

Evidently, \( I \subseteq I(V(I)) \) and if \( f^n \in I(X) \) for some \( n \in \mathbb{N} \), then \( f \in I(X) \). Therefore, letting \( \sqrt{I} = \{ f \in k[t_1, \ldots, t_n] \mid f^n \in I \text{ for some } n \in \mathbb{N} \} \) denote the radical of the ideal \( I \), which may also be described as the intersection of all prime ideals of \( k[t_1, \ldots, t_n] \) that contain \( I \) [8, Proposition 1.8], we certainly have the inclusion \( \sqrt{I} \subseteq I(V(I)) \). In fact, equality holds here:

**Hilbert’s Nullstellensatz.** \( I(V(I)) = \sqrt{I} \) for every ideal \( I \) of \( k[t_1, \ldots, t_n] \).
Observe that if \( V(I) = \emptyset \), then the Nullstellensatz gives \( \sqrt{I} = I(\emptyset) = \mathbb{k}[t_1, \ldots, t_n] \).
Hence \( 1 \in \sqrt{I} \) and so \( 1 \in I \). Thus, part (a) of the following statement is a direct
consequence of the Nullstellensatz; part (b) is also known as Zariski’s Lemma.

**Weak Nullstellensatz.**  
(a) **Classical formulation.** \( V(I) \neq \emptyset \) for every proper
ideal \( I \) of \( \mathbb{k}[t_1, \ldots, t_n] \).

(b) **Ring theoretic formulation.** For every maximal ideal \( I \) of \( \mathbb{k}[t_1, \ldots, t_n] \),
the field \( \mathbb{k}[t_1, \ldots, t_n]/I \) is finite dimensional over \( \mathbb{k} \).

The two formulations of the weak Nullstellensatz are in fact equivalent. To see
that (a) implies (b), consider a maximal ideal \( I \) of \( \mathbb{k}[t_1, \ldots, t_n] \) and pick \( x \in V(I) \).
Then the evaluation map \( \mathbb{k}[t_1, \ldots, t_n] \to \mathbb{k}, \ f \mapsto f(x) \), is a ring homomorphism
with kernel \( I \). Therefore, \( \mathbb{k}[t_1, \ldots, t_n]/I \) is a finitely generated algebraic field
extension of \( \mathbb{k} \), and so \( \mathbb{k}[t_1, \ldots, t_n]/I \) is finite dimensional over \( \mathbb{k} \), proving (b).
Conversely, assume (b) and let \( I \) be a proper ideal of \( \mathbb{k}[t_1, \ldots, t_n] \). In order to
show that \( V(I) \neq \emptyset \), we may assume that \( I \) is maximal, because \( V(\cdot) \) is inclusion reversing. But then \( K = \mathbb{k}[t_1, \ldots, t_n]/I \) is a finite field extension of \( \mathbb{k} \) by (b), and
so there is an embedding \( K \hookrightarrow \mathbb{k} \) that is the identity on \( \mathbb{k} \). Letting \( \lambda_i \in \mathbb{k} \) denote
the image of \( t_i + I \in K \), we obtain a point \( x = (\lambda_1, \ldots, \lambda_n) \in V(I) \). This proves the
equivalence of (a) and (b). The foregoing also shows that the weak Nullstellensatz
gives the following description of the maximal ideals of \( \mathbb{k}[t_1, \ldots, t_n] \): they are exactly the ideals of the form \( I(\{x\}) \) for some point \( x \in \mathbb{k}^n \).

In fact, the full Nullstellensatz can also be derived from (b) via the following
argument, which is known as the “Rabinowitsch trick” [170]. Given an ideal \( I \) of
\( A = \mathbb{k}[t_1, \ldots, t_n] \), we must show that \( f \in P \) holds for all \( f \in I(V(I)) \) and all prime
ideals \( P \) of \( A \) such that \( I \subseteq P \). Suppose otherwise and fix \( f \) and \( P \) with \( f \notin P \).
Form the polynomial algebra \( A[x] = \mathbb{k}[t_1, \ldots, t_n, x] \) and consider the ideal of \( A[x] \)
that is generated by \( P \) and the element \( 1 - xf \); this is a proper ideal, because it is
contained in the kernel of the ring homomorphism \( A[x] \to \text{Fract}(A/P) \) that is given
by \( t_i \mapsto t_i + P \) and \( x \mapsto (f + P)^{-1} \). Therefore, we may choose a maximal ideal \( J \)
of \( A[x] \) such that \( P \subseteq J \) and \( 1 - xf \in J \). Then, exactly as above, \( K = A[x]/J \) is a
finite field extension of \( \mathbb{k} \) by (b), giving a point \( x = (\lambda_1, \ldots, \lambda_n, \mu) \in V(J) \).
Thus, \( 1 - \mu f(\lambda_1, \ldots, \lambda_n) = 0 \) and \( g(\lambda_1, \ldots, \lambda_n) = 0 \) for all \( g \in P \); so \( (\lambda_1, \ldots, \lambda_n) \in V(P) \).
Since \( V(P) \subseteq V(I) \) and \( f \in I(V(I)) \), we must have \( f(\lambda_1, \ldots, \lambda_n) = 0 \) while, on the other hand, \( \mu f(\lambda_1, \ldots, \lambda_n) = 1 \). This contradiction completes the argument.

Ideals \( I \) satisfying \( \sqrt{I} = I \) are called **semi-prime**. In this case, the Nullstellensatz
gives \( I = I(V(I)) = \cap_{x \in V(I)} I(\{x\}) \). Thus, every semi-prime ideal of \( \mathbb{k}[t_1, \ldots, t_n] \)
is an intersection of maximal ideals. This is called the **Jacobson property** of
\( \mathbb{k}[t_1, \ldots, t_n] \). Finally, for a proof of version (b) of the weak Nullstellensatz, we refer
the reader to Eisenbud [65, p. 142-144] or to Atiyah-Macdonald [8, p. 69-70].
C.2. The Generic Flatness Lemma

The following lemma is due to Grothendieck [95, Exp. IV, Lemme 6.7] or [94, Lemme 6.9.2]. For any commutative domain $R$ and any $0 \neq f \in R$, we let $R_f$ denote the localization of $R$ at the powers of $f$, that is, the subring of the field of fractions of $R$ that is generated by $R$ and $f^{-1}$.

**Generic Flatness Lemma.** Let $R$ be a commutative domain and let $S$ be a finitely generated commutative $R$-algebra. Then, for any finitely generated $S$-module $M$, there exists $0 \neq f \in R$ such that $M_f = R_f \otimes_R M$ is a free as $R_f$-module.

**Proof.** Since $R_f$ is flat as $R$-module, we can argue by induction on the number of generators of $M$ to reduce to the case where $M$ is cyclic. Thus, $M \cong S/I$ for some ideal $I$ of $S$. Replacing $S$ by $S/I$, we may further assume that $M = S$.

Fix a set of generators $x_1, \ldots, x_n$ for the $R$-algebra $S$. We will write monomials in these generators as $x^m = x_1^{m_1} \cdots x_n^{m_n}$ with $m = (m_1, \ldots, m_n) \in \mathbb{Z}_{+}^n$ and we also put $|m| = \sum_i m_i$. Define a total order $\preceq$ on $\mathbb{Z}_{+}^n$ by ordering all subsets $\{m \in \mathbb{Z}_{+}^n \mid |m| = d\}$ for fixed $d$ lexicographically and by declaring $m < n$ if $|m| < |n|$. Note that that $\preceq$ respects the additive structure of $\mathbb{Z}_{+}^n$ in the sense that $m \leq n$ implies $m + k \leq n + k$ for all $m, n, k \in \mathbb{Z}_{+}^n$. Furthermore, $\mathbb{Z}_{+}^n$ is isomorphic to $(\mathbb{Z}_{+}, \preceq)$ as an ordered set. Now put

$$S_n^\sim := \sum_{m<n} R x^m \subseteq S_n := Rx^n + S_n^\sim$$

Thus, $S_n/S_n^\sim \cong R/I_n$ with $I_n = \{ r \in R \mid rx^n \in S_n^\sim \}$. For $r \in I_n$ and $m \in \mathbb{Z}_{+}^n$, one has $r x^m + n = x^m r x^n \in x^m S_n \subseteq S_{m+n}$. Therefore, $I_n \subseteq I_{m+n}$; in terms of the partial order $\preceq$ on $\mathbb{Z}_{+}^n$ that is given by $m \preceq n$ if and only if $n - m \in \mathbb{Z}_{+}^n$, this says that $m \preceq n$ implies $I_m \subseteq I_n$. A noteworthy feature of $\preceq$ is that any subset $N \subseteq \mathbb{Z}_{+}^n$ is a finite (possibly empty) union of sets of the form $N_{n_i} = \{ n \in N \mid n \geq n_i \}$ with $n_i \in N$ (Exercise C.3.1). Taking $N = \{ n \in \mathbb{Z}_{+}^n \mid I_n \neq 0 \}$, we conclude that there are finitely many nonzero $I_n$, so that each nonzero $I_n$ contains one of them. Since $R$ is a domain, we may pick an element $0 \neq f \in \bigcap_{n \in N} I_n$. Then $R_f \otimes_R S_n/S_n^\sim \cong R_f \otimes_R R/I_n = 0$ for $n \in N$, while $R_f \otimes_R S_n/S_n^\sim \cong R_f$ for $n \notin N$. Consequently, $S_f = R_f \otimes_R S$ is the union of a chain of $R_f$-submodules whose successive nonzero factors are all isomorphic to $R_f$, and hence $S_f$ is free as $R_f$-module (Proposition 2.1).

C.3. The Zariski Topology on a Vector Space

Let $V$ be a finite-dimensional vector space over the field $\mathbb{k}$. The **algebra of polynomial functions** on $V$, by definition, is the symmetric algebra of the dual space:

$$\mathcal{O}(V) \overset{\text{def}}{=} \text{Sym} V^*$$
The algebra $O(V)$ is related to the algebra $\mathbb{k}^V$ of all $\mathbb{k}$-valued functions on $V$, with the usual “pointwise” addition and multiplication of functions, as follows. Evaluation $\langle \cdot , \cdot \rangle : V^\ast \times V \to \mathbb{k}$ gives a $\mathbb{k}$-linear map $V^\ast \to \mathbb{k}^V$, $f \mapsto (v \mapsto \langle f, v \rangle)$. By the universal property of symmetric algebras (1.8), this map lifts uniquely to an algebra map $\Phi : O(V) \to \mathbb{k}^V$. If $(x_i)_d^j$ is a $\mathbb{k}$-basis of $V$ and $(x_i^j)_d$ the dual basis of $V^\ast$, then $O(V) \cong \mathbb{k}[x^1, \ldots, x^d]$ and the map $\Phi$ takes the form

$$\Phi : \quad O(V) \cong \mathbb{k}[x^1, \ldots, x^d] \longrightarrow \mathbb{k}^V = \{\text{functions } V \to \mathbb{k}\}$$

(C.1)

$$f \longmapsto \left( \sum_i \lambda_i x_i \mapsto f(\lambda_1, \ldots, \lambda_d) \right)$$

This map is an embedding if — and for $d \neq 0$ only if — $\mathbb{k}$ is infinite. If $|\mathbb{k}| \geq n$, then at least the restriction of $\Phi$ to the space $O^n(V)$ of homogeneous polynomial functions of total degree $n$ is injective. See Exercise C.3.2 for all this. Even when the map $\Phi$ is not necessarily injective, each $f \in O(V)$ yields a $\mathbb{k}$-valued function on $V$ via $\Phi$, and we will write $f(v) = \Phi(f)(v) \in \mathbb{k}$ for $v \in V$.

For any subset $S \subseteq O(V)$, we may now define

$$V(S) \overset{\text{def}}{=} \{v \in V \mid f(v) = 0 \text{ for all } f \in S\}.$$  

If $S = \{f\}$, then we simply write $V(f)$. The reader will readily verify the following assertions: $V(1) = \emptyset$ and $V(0) = V$; $\bigcap_{\sigma} V(S_{\sigma}) = V(\bigcup_{\sigma} S_{\sigma})$ for any collection $\{S_{\sigma}\}$ of subsets of $O(V)$; and $V(S) \cup V(S') = V(SS')$ for any two subsets $S, S' \subseteq O(V)$, where $SS' \subseteq O(V)$ consists of the products $ff'$ with $f \in S$ and $f' \in S'$. Thus, declaring the subsets of the form $V(S)$ to be closed, we obtain a topology on $V$, the so-called Zariski topology. The open subsets of $V$ are the sets $V(S)^C = \bigcup_{f \in S} V_f$, where we have put

$$V_f = V(f)^C = \{v \in V \mid f(v) \neq 0\}.$$  

The sets $V_f$ are called the principal open subsets of $V$. Evidently, $V_f \neq \emptyset$ is equivalent to $\Phi(f) \neq 0$. Thus, a subset $D \subseteq V$ is dense for the Zariski topology, or Zariski dense for short, if and only if $D \cap V_f \neq \emptyset$ for all $f \in O(V)$ with $\Phi(f) \neq 0$. It is not hard to show that, if $\mathbb{k}$ is infinite, then all nonempty open subsets of $V$ are Zariski dense; if $\mathbb{k}$ is finite, then the Zariski topology is discrete: all subsets of $V$ are open and closed (Exercise C.3.3).

**Exercises for Section C.3**

**C.3.1** (A detail for the proof of the Generic Flatness Lemma). Let $\preceq$ denote the partial order on $\mathbb{Z}_n^m$ that is given by $\mathbf{m} \preceq \mathbf{n}$ if and only if $\mathbf{n} - \mathbf{m} \in \mathbb{Z}_n^m$. Show that any subset $N \subseteq \mathbb{Z}_n^m$ is a finite (possibly empty) union of sets of the form $N_{\mathbf{n}_f} = \{\mathbf{n} \in N \mid \mathbf{n} \geq \mathbf{n}_f\}$ with $\mathbf{n}_f \in N$.  

C.3.2 (Polynomial functions). For a finite-dimensional \( \mathbb{k} \)-vector space \( V \), consider the algebra map \( \Phi: \mathcal{O}(V) \to \mathbb{k}^V \) in (C.1). Prove:

(a) If \( |\mathbb{k}| > n \) then the restriction of \( \Phi \) to the space \( \mathcal{O}_n(V) \) of polynomial functions of total degree \( \leq n \) is mono. The converse also holds when \( V \neq 0 \).

(b) If \( |\mathbb{k}| \geq n \) then the restriction of \( \Phi \) to the space \( \mathcal{O}^n(V) \) of homogeneous polynomial functions of total degree \( n \) is mono. The converse holds for \( \dim_\mathbb{k} V \geq 2 \).

C.3.3 (Zariski topology). Let \( V \) be a finite-dimensional \( \mathbb{k} \)-vector space. Prove:

(a) All “points” \( v \in V \) and all subspaces \( U \subseteq V \) are closed in the Zariski topology.

(b) If \( \mathbb{k} \) is finite, then the Zariski topology on \( V \) is the discrete topology: all subsets of \( V \) are open and closed.

(c) If \( V \neq 0 \), then all nonempty open subsets of \( V \) are Zariski dense if and only if \( \mathbb{k} \) is infinite.

(d) \( V \) is quasi-compact: if \( V = \bigcup_{i \in I} U_i \) for some collection of open subsets of \( U_i \subseteq V \), then \( V = \bigcup_{i \in I'} U_i \) with \( I' \subseteq I \) finite.
The Diamond Lemma

D.1. The Goal

We are interested in exhibiting a \( k \)-basis of an associative \( k \)-algebra \( A \) that is given to us by a presentation of the form

\[
A = \mathcal{K}(X)/ (w_\sigma - f_\sigma \mid \sigma \in S)
\]

Here \( \mathcal{K}(X) \) is the free \( k \)-algebra generated by a given set \( X \) as in §1.1.2(B) all \( f_\sigma \in \mathcal{K}(X) \) and the \( w_\sigma \) are finite products with factors from \( X \), also called monomials or words. The collection of all words forms a submonoid \( W = \langle X \rangle \leq (\mathcal{K}(X), \cdot) \) that is a \( k \)-basis of \( \mathcal{K}(X) \). Therefore, the image of \( W \) in \( A \) spans \( A \) as a \( k \)-vector space. The issue is to identify a subset of \( W \) that yields a \( k \)-basis of \( A \). In this section, we will describe the method of [15].

Here are two simple examples of the situation described in (D.1); they will serve to illustrate the discussion that follows.

Example D.1 (The Weyl algebra). Recall from (1.15) that the first Weyl algebra over \( k \) is defined by

\[
A = A_1(k) = \mathcal{K}(x, y)/(yx - xy - 1)
\]

So \( X = \{x, y\} \) and there is a single relation \( w - f \) with \( (w, f) = (yx, xy + 1) \).

Example D.2 (The enveloping algebra of \( sl_2 \)). Consider the algebra

\[
A = \mathcal{K}(e, f, h)/(ef - fe - h, hf - fh + 2f, eh - he + 2e)
\]

We will see later that \( A \cong U(sl_2) \), the enveloping algebra of the Lie algebra \( sl_2(k) \). Here, \( X = \{e, f, h\} \) and we have three relations \( w_i - f_i \) given by \( (w_1, f_1) = (ef, fe + h), (w_2, f_2) = (hf, fh - 2f) \) and \( (w_3, f_3) = (eh, he - 2e) \).
D. The Method

As in Examples D.1 and D.2 above, we will encode the relations in (D.1) in a set of pairs,

\[ S = \{ \sigma = (w_\sigma, f_\sigma) \} \]

We will refer to \( S \) as a reduction system. Indeed, each member of \( S \) gives rise to a substitution \( (aw_\sigma b \mapsto af_\sigma b) \) that can be applied to all words \( aw_\sigma b \in W \) containing \( w_\sigma \) as a subword. The \( \mathbb{k} \)-linear endomorphism of \( \mathbb{k} \langle X \rangle \) that sends the specific word \( aw_\sigma b \) to \( af_\sigma b \) and fixes all other words will be denoted by

\[ r_{a,\sigma,b} : \mathbb{k} \langle X \rangle \to \mathbb{k} \langle X \rangle \]

Any finite composite of such endomorphisms will be called a reduction. We have the following requests to ensure that \( S \) leads to an effective reduction process:

1. For each \( \sigma \in S \), the expression \( af_\sigma b \) should be an improvement over \( aw_\sigma b \) in some way;
2. Starting with any word \( w \in W \), we should reach, after finitely many reductions, a final expression \( f \in \mathbb{k} \langle X \rangle \) that can no longer be improved by further reduction, that is, \( f \) involves no monomial of the form \( aw_\sigma b \); and
3. The fully reduced expression \( f \) in (2) should be unambiguously determined by the starting word \( w \) and not depend on the particular sequence of reductions chosen.

We denote the collection of all fully reduced words by \( W_{irr} \); it will turn out that \( W_{irr} \) gives us the desired \( \mathbb{k} \)-basis of the algebra \( A \). Here are the technicalities that will ensure that our above wish list is fulfilled.

For (1), we assume that we are given a partial order \( \leq \) on \( W \). If \( w' < w \) then \( w' \) will be considered an improvement over \( w \). Similarly, any \( f \in \mathbb{k} \langle X \rangle \) that is a \( \mathbb{k} \)-linear combination of words \( < w \) will be considered better than \( w \). To make sure that the substitutions \( (aw_\sigma b \mapsto af_\sigma b) \) for \( \sigma \in S \) result in improvements, we assume that

- \( \leq \) is compatible with \( S \) in the following sense: for all \( \sigma \in S \), the expression \( f_\sigma \) is a \( \mathbb{k} \)-linear combination of words \( < w_\sigma \), and
- \( \leq \) is a semigroup partial order on \( W \), that is, \( \leq \) is preserved by multiplication in \( W \): for all \( a, b, w, w' \in W \), \( w' < w \Rightarrow aw'b < awb \).

For (2), we shall require that the semigroup partial order \( \leq \) on \( W \) satisfies the descending chain condition (DCC):

- Every chain \( w_1 \geq w_2 \geq w_3 \geq \ldots \) with \( w_i \in W \) stabilizes after finitely many steps.

In light of the fact that all reductions lead to improvements, this will ensure that our reduction system \( S \) does not allow for an infinite sequence of reductions.
The most stringent requirement is (3). For this, we shall use the diamond conditions. Observe that an ambiguity in the reduction process of a given word \( w \in W \) arises when several \( \sigma \in S \) are available to rewrite \( w \). Specifically, the following types of ambiguities need to be considered:

**Overlap ambiguities.** These are quintuples \( (a, b, c, \sigma, \tau) \in (W \setminus \{1\})^3 \times S^2 \) such that \( ab = w_\sigma \) and \( bc = w_\tau \). The ambiguity lies in the fact that the word \( w = abc \in W \) has two immediate reductions: \( r_{a,\sigma,c}(abc) = f_\sigma c \) and \( r_{a,\tau,1}(abc) = af_\tau \). We shall say that the overlap ambiguity \( (a, b, c, \sigma, \tau) \) is resolvable if there are reductions \( r \) and \( r' \) such that \( r(f_\sigma c) = r'(af_\tau) \):

\[
(D.4)
\]

**Inclusion ambiguities.** These are quintuples \( (a, b, c, \sigma, \tau) \in W^3 \times S^2 \) such that \( w_\sigma = b \) and \( w_\tau = abc \). The ambiguity is called resolvable if the results of the two immediate reductions, \( af_\sigma c \) and \( f_\tau \), can be reduced to a common expression \( \in \mathbb{k}(X) \):

\[
(D.5)
\]

We are now ready to state the main result of this section. For the proof, we refer to the original source [15].

**Diamond Lemma (G. Bergman).** Let \( S = \{ \sigma = (w_\sigma, f_\sigma) \} \subseteq W \times \mathbb{k}(X) \) be given, and let \( \leq \) be a semigroup partial ordering on \( W \) that is compatible with \( S \) and satisfies DCC. Let \( W_{\text{irr}} \) denote the set of all words \( w \in W \) that do not contain any \( w_\sigma \) (\( \sigma \in S \)) as a subword. Then the following are equivalent:

(i) all overlap and inclusion ambiguities are resolvable, that is, the diamond conditions (D.4) and (D.5) are satisfied;

(ii) the residue classes of the words \( w \in W_{\text{irr}} \) form a \( \mathbb{k} \)-basis for the algebra \( \mathbb{k}(X)/(w_\sigma - f_\sigma \mid \sigma \in S) \).

The main interest for our purposes lies in the implication (i) \( \Rightarrow \) (ii). In §5.4.3 below, we will add a third equivalent condition, also from [15], that is often easier to verify than (i). For many applications, however, the formulation of the Diamond Lemma above is adequate.
D.3. First Applications

In order to apply the Diamond Lemma, we need a suitable partial order on the monoid \( W \) of words in the given alphabet \( X \). Here is one way to choose such an order. Fix a total order \( \leq \) on \( X \). Then we may define the \textit{degree-lexicographic order} \( \leq_{\text{deglex}} \) on \( W \) as follows. First, define the degree of a word \( w = x_1 x_2 \ldots x_n \in W \) (\( x_i \in X \)) by \( \deg w = n \) – this is the same as the degree of \( w \) in the free algebra \( \mathbb{k}(X) \) viewed as the tensor algebra with its usual grading. Next, define the lexicographic order between words of a given degree by declaring \( w = x_1 x_2 \ldots x_n \leq_{\text{lex}} w' = x'_1 x'_2 \ldots x'_n \) (\( x_i, x'_i \in X \)) if there is some \( t \) such that \( x_i = x'_i \) for \( i < t \) but \( x_i < x'_i \). Finally, define

\[ w \leq_{\text{deglex}} w' \iff \deg w < \deg w' \text{ or else } \deg w = \deg w' \text{ and } w \leq_{\text{lex}} w' \]

The degree-lexicographic order is easily seen to be a semigroup (total) order on \( W \). Moreover, if the order \( \leq \) on \( X \) satisfies DCC then so does \( \leq_{\text{deglex}} \). Indeed, given a descending chain \( w_1 \geq w_2 \geq w_3 \geq \ldots \) in \( W \), we obtain a descending chain \( \deg w_1 \geq \deg w_2 \geq \deg w_3 \geq \ldots \) in \( \mathbb{Z}_+ \), which must must stabilize. From this point on, the first letters in all words form a descending chain in \( X \), which must also stabilize by hypothesis on \( (X, \leq) \). Moving on to the letters in the second position and so on, the letters in all positions will have stabilized eventually, and hence the given chain in \( W \) has become stable. We remark that, while it always possible to choose order \( \leq \) on \( X \) which satisfies DCC by the well-ordering theorem, \( X \) is often finite in practice and DCC not an issue. However, we must take care to choose \( \leq \) on \( X \) so that \( \leq_{\text{deglex}} \) is compatible with the given reduction system \( S \). As for ambiguities, we will focus on overlap ambiguities below. In fact, it turns out that inclusion ambiguities are, in a sense, always avoidable; see [15, Section 5].

Example D.3 (The Weyl algebra, continued). Recall from Example D.1 that the Weyl algebra \( A = A_1(\mathbb{k}) \) is defined from the free generators \( X = \{ x, y \} \) by means of the singleton reduction system \( S = \{ (yx, xy + 1) \} \). Ordering \( X \) by \( x < y \), the degree-lexicographic order on the monoid \( W \) of words in \( x \) and \( y \) satisfies all requirements specified in the Diamond Lemma. Indeed, it only remains to remark that \( \leq_{\text{deglex}} \) is compatible with \( S \), because \( xy \leq_{\text{deglex}} yx \) and \( 1 \leq_{\text{deglex}} yx \). Since a singleton reduction system does not lead to any ambiguities, the Diamond Lemma tells us that the images of the words in \( W_{\text{irr}} \) form a \( \mathbb{k} \)-basis of \( A_1(\mathbb{k}) \). Finally, \( W_{\text{irr}} = \{ x^r y^s \mid r, s \in \mathbb{Z}_+ \} \). We conclude that the ordered or “standard” monomials in the two generators of \( A_1(\mathbb{k}) \) form a \( \mathbb{k} \)-basis.

There are of course other ways to see this. For example, we may consider the standard representation \( A_1(\mathbb{k}) \to \text{End}_\mathbb{k}(\mathbb{k}[t]) \) that is given by \( x \mapsto t \) and \( y \mapsto \frac{dt}{dr} \); see Example 1.8. To reach the same conclusion as above, it suffices to check that the operators \( \frac{d^r}{dt^r} \in \text{End}_\mathbb{k}(\mathbb{k}[t]) \) are \( \mathbb{k} \)-linearly independent, which is elementary if char \( \mathbb{k} = 0 \) (and false otherwise).
Example D.4 (The enveloping algebra of $\mathfrak{sl}_2$, continued). Let us now return to the algebra $A \cong U(\mathfrak{sl}_2)$ from Example D.2. Recall that $X = \{e, f, h\}$ and $S = \{\rho = (ef, fe + h), \sigma = (hf, fh - 2f), \tau = (eh, he - 2e)\}$. The fully reduced words for this reduction system are all words $w \in W$ that contain none of $ef, hf$ or $eh$ as a subword; so $W_{\text{tr}} = \{f^r h^s e^t \mid r, s, t \in \mathbb{Z}_+\}$. Choosing the degree-lexicographic order on $W$ that is based on the order $f < h < e$ of $X$, the order hypotheses of the Diamond Lemma are easily seen to be satisfied. Moreover, since the words $w_\rho = ef, w_\sigma = hf$ and $w_\tau = eh$ are all distinct but have the same degree, there are no inclusion ambiguities. However, there is one overlap ambiguity that needs to be resolved: $(e, h, f, \sigma, \tau)$. This is done in Figure D.1 below. Now we conclude from the Diamond Lemma that the desired $k$-basis of $A$ comes from the set $W_{\text{tr}}$. Thus, once again, the “standard” monomials in the three ordered generators of $A$ form a $k$-basis.

Example D.5 (Quantum $2 \times 2$-matrices). The quantum matrix algebra $O_q(\text{Mat}_2(k))$ is defined to be the $k$-algebra with four generators, $a, b, c$ and $d$, subject to the relations

\begin{align*}
ab &= q \ ba \quad & ac &= qa \quad & bc &= cb \\
bd &= q \ db \quad & cd &= q dc \quad & ad - da &= \tilde{q} bc
\end{align*}

(D.6)

Here, $q \in k^*$ is a nonzero parameter and $\tilde{q} = q - q^{-1}$. Thinking of the given generators as arranged in the $2 \times 2$-matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$, the four relations on the left are sometimes expressed by saying that generators in the same row or column skew commute. Formally, let us think of $O_q(\text{Mat}_2(k))$ as given by a presentation of the form (D.1), with $X = \{x_1, x_2, x_3, x_4\}$ and

$$S = \{(x_j x_i, q^{-1} x_i x_j) \mid (i, j) = (1, 2), (1, 3), (2, 4) \text{ or } (3, 4)\}$$

$$\cup \{(x_3 x_2, x_2 x_3), (x_4 x_1, x_1 x_4 - \tilde{q} x_2 x_3)\}.$$
The Diamond Lemma implies that the monomials slightly less detail than in Figure D.1. Therefore, as in the previous example, the compatible with the given reduction system of resolvability of an ambiguity, called resolvability relative to following its statement. This addendum concerns a formal weakening of the notion of resolvability for $O_q(\text{Mat}_2(\mathbb{k}))$

\begin{align*}
x_3 x_2 x_1 : & \quad \{ (x_2 x_3) x_1 \mapsto x_2(q^{-1}x_1 x_3) \mapsto q^{-1}(q^{-1}x_1 x_2) x_3 \\
& \quad x_3(q^{-1}x_1 x_2) \mapsto q^{-1}(q^{-1}x_1 x_3) x_2 \mapsto q^{-2}x_1(x_2 x_3) \\
x_4 x_2 x_1 : & \quad \{ (q^{-1}x_2 x_4) x_1 \mapsto q^{-1}x_2(x_1 x_4 - \hat{q} x_2 x_3) \\
& \quad \mapsto q^{-1}(q^{-1}x_1 x_2) x_4 - q^{-1}\hat{q} x_2^2 x_3 \\
x_4 x_3 x_1 : & \quad \{ q^{-1}x_3 x_4 x_1 \mapsto q^{-1}x_3(x_1 x_4 - \hat{q} x_2 x_3) \\
& \quad \mapsto q^{-1}(q^{-1}x_1 x_3) x_4 - q^{-1}\hat{q}(x_2 x_3) x_3 \\
x_4 x_3 x_2 : & \quad \{ q^{-1}x_3 x_4 x_2 \mapsto q^{-1}x_3(q^{-1}x_2 x_4) \mapsto q^{-2}(x_2 x_3) x_4 \\
& \quad x_4(x_2 x_3) \mapsto (q^{-1}x_2 x_4) x_3 \mapsto q^{-1}x_2(q^{-1}x_3 x_4) \\
\end{align*}

Figure D.2. Resolving ambiguities for $O_q(\text{Mat}_2(\mathbb{k}))$

So $a$ is the image of $x_1$, $b$ the image of $x_2$ and so on. Ordering $X$ by $x_1 < x_2 < x_3 < x_4$ and the monoid $W$ of words in the $x_i$'s by the degree-lexicographic order, all order requirements of the Diamond Lemma are satisfied. There are only four overlap ambiguities which are easily resolved; this is shown Figure D.2 below in slightly less detail than in Figure D.1. Therefore, as in the previous example, the Diamond Lemma implies that the monomials $a^k b^l c^m d^n$ ($k, l, m, n \in \mathbb{Z}_{+}$) form a $\mathbb{k}$-basis of $O_q(\text{Mat}_2(\mathbb{k}))$.

Example D.6 (Quantum affine $n$-space).

D.4. A Simplification

We now come to the addendum to the Diamond Lemma that was announced following its statement. This addendum concerns a formal weakening of the notion of resolvability of an ambiguity, called resolvability relative to $\leq$. Here, $\leq$ is the given semigroup partial ordering of the monoid of words $W$, assumed to be compatible with the given reduction system $S$ (and to satisfy DCC).

In detail, let us denote the relation ideal in (D.1) by $I$; so

$$I = (w_\sigma - f_\sigma \mid \sigma \in S)$$

Every element of $I$ is a $\mathbb{k}$-linear combination of expressions of the form $a(w_\sigma - f_\sigma)b = aw_\sigma b - af_\sigma b$ with $a, b \in W$. Since $\leq$ is a semigroup partial order that is compatible with $S$, all words occurring with nonzero coefficient in $af_\sigma b$ are $< aw_\sigma b$.
so $aw_{σ}b$ may be thought of as the leading term of $a(w_{σ} - f_{σ})b$. For any given $w ∈ W$, put

$$I_{w} = \langle a(w_{σ} - f_{σ})b \mid a, b ∈ W, aw_{σ}b < w \rangle_{k}$$

where $\langle \ldots \rangle_{k}$ denotes the $k$-linear span. For later use, we note the following simple lemma.

**Lemma D.7.** Let $w ∈ W$ and suppose that $f ∈ k(X)_{w} := \langle a ∈ W \mid a < w \rangle_{k}$. Then, for any reduction $r$, we have $r(f) ∈ k(X)_{w}$ and $f ≡ r(f) \mod I_{w}$.

**Proof.** It suffices to prove this for a simple reduction of the form $r = r_{a, σ, b}$ with $a, b ∈ W$. By linearity, we may also assume that $f$ is a word; in fact, we may assume that $f = aw_{σ}b$, because $r$ fixes all other words. But then $r(f) = a f_{σ}b ∈ k(X)_{w}$, because $f_{σ} ∈ k(X)_{w, r}$ by compatibility of $≤$ with $S$, and $f − r(f) = a(w_{σ} − f_{σ})b$ with $aw_{σ}b < w$, because $f ∈ k(X)_{w}$. Thus, $f − r(f) ∈ I_{w}$.

Recall from the diamond conditions (D.4) and (D.5) that, for any ambiguity $(a, b, c, σ, τ)$, there are two immediate reductions, say $r_{0}$ and $r_{0}'$, of the word $abc ∈ W$. For instance, if $(a, b, c, σ, τ)$ is an overlap ambiguity, then $r_{0}(abc) = f_{σ}c$ and $r_{0}'(abc) = a f_{σ}$. Similarly for an inclusion ambiguity. Resolving the ambiguity amounts to finding reductions $r$ and $r'$ such that $rr_{0}(abc) = r'r_{0}'(abc)$. Instead, the (overlap or inclusion) ambiguity $(a, b, c, σ, τ)$ is said to be **resolvable relative to** $≤$ if the following condition is satisfied:

$$r_{0}(abc) ≡ r_{0}'(abc) \mod I_{abc}$$

It turns out that this is equivalent to conditions (i) and (ii) in the Diamond Lemma; so we state for the record:

**Addendum to the Diamond Lemma** (G. Bergman). In the situation of the Diamond Lemma, conditions (i) and (ii) are also equivalent to

(i') all overlap and inclusion ambiguities are resolvable relative to $≤$.

Again, we refer the reader to [15] for complete details, but let us at least point out that $r_{0}(abc) ≡ rr_{0}(abc) \mod I_{abc}$ holds for any reduction $r$ and similarly for $r_{0}'(abc)$; this follows from Lemma D.7. Therefore, the ambiguity $(a, b, c, σ, τ)$ is resolvable relative to $≤$ if and only if the following condition is satisfied:

(D.7) there exist reductions $r$ and $r'$ such that $rr_{0}(abc) ≡ r'r_{0}'(abc) \mod I_{abc}$

In particular, resolvable ambiguities are clearly resolvable relative to $≤$, which is the implication (i) $⇒$ (i') in the Diamond Lemma.

**D.5. The Poincaré-Birkhoff-Witt Theorem**

Bring notation in line. ——— As an application of the Addendum to the Diamond Lemma, we shall now prove the Poincaré-Birkhoff-Witt Theorem which was already stated, without proof, in §5.4.3. Recall that if $g$ is a Lie algebra over $k$ and $X$ is
any \( k \)-basis of \( g \), then we may identify the tensor algebra \( T_g \) with the free algebra \( k\langle X \rangle \) and the enveloping algebra \( U_g \) with the factor \( k\langle X \rangle / I \), where \( I \) is the ideal of \( k\langle X \rangle \) that is generated by the elements \( xy - yx - [x, y] \) with \( x, y \in X \). Here, the Lie algebra \( g \subseteq T_g \) is identified with the \( k \)-span \( \langle X \rangle_k \subseteq k\langle X \rangle \); so we have a bracket \( \{ \ldots \} : \langle X \rangle_k \times \langle X \rangle_k \rightarrow k\langle X \rangle \). In order to distinguish this bracket from the commutator bracket of \( k\langle X \rangle_{\text{Lie}} \), we will denote the latter by \( [\ldots]_{k\langle X \rangle} \); so \( [a, b]_{k\langle X \rangle} = ab - ba \) for \( a, b \in k\langle X \rangle \). With these notations, we may write the generators of \( I \) as follows:

\[
\text{(D.8)} \quad xy - yx - [x, y] = [x, y]_{k\langle X \rangle} - [x, y]
\]

with \( x, y \in X \). If \( X \) is totally ordered by \( \leq \) then it suffices to take only \( x, y \in X \) with \( x < y \) here.

**Poincaré-Birkhoff-Witt Theorem.** Let \( g \) be a Lie \( k \)-algebra and let \( X \) be a \( k \)-basis of \( g \) that is totally ordered by \( \leq \). Then a \( k \)-basis of \( U_g \) is given by the residue classes of the ordered words \( x_1x_2\ldots x_n \in W \subseteq k\langle X \rangle \), with \( n \in \mathbb{Z}_+, \ x_i \in X \) and \( x_1 \leq x_2 \leq \cdots \leq x_n \).

**Proof.** Since \( I = (yx - (xy - [x, y])) \mid x, y \in X, x < y \), we may take

\[
S = \{ \sigma_{yx} = (yx, xy - [x, y]) \mid x, y \in X, x < y \}
\]

for our reduction system. This leads to \( W_{\text{irr}} = \{ x_1x_2\ldots x_n \mid n \in \mathbb{Z}_+, \ x_i \in X \) and \( x_1 \leq x_2 \leq \cdots \leq x_n \} \) as our set of fully reduced words, and hence to the desired \( k \)-basis of \( U_g \), provided we can satisfy the order requirements of the Diamond Lemma and can take care of ambiguities.

Since our chosen order \( \leq \) of \( X \) need not satisfy DCC, the associated degree-lexicographic order of \( W \) will not work this time. Instead, we settle for a partial order on \( W \) that uses the **misordering index** of words as a tie-breaker for the degree. Here, the misordering index of \( x_1x_2\ldots x_n \in W \) \( (x_i \in X) \) is defined to be the number of pairs \((i, j)\) with \( i < j \) but \( x_j > x_i \). Thus, \( W_{\text{irr}} \) consists exactly of the words having misordering index 0, while any word \( x_1x_2\ldots x_n \in W \) with \( x_1 > x_2 > \cdots > x_n \) has the largest possible misordering index, \( \binom{n}{2} \). Define a partial order \( \preceq \) on \( W \) by declaring \( w' < w \) if either \( \deg w' < \deg w \) or if \( w' \) arises from \( w \) by a permutation of the factors but has smaller misordering index than \( w \). While \( \preceq \) is only a partial order of \( W \), it is straightforward to check that it respects the semigroup structure, is compatible with our reduction system \( S \), and satisfies DCC. This is all we need for the Diamond Lemma.

For ambiguities, we only have the overlap ambiguities \( (z, x, y, \sigma_{zy}, \sigma_{yx}) \) with \( z > y > x \) in \( X \). We will resolve these ambiguities relative to \( \leq \). Exactly as in
Figure D.1, we compute
\[
\begin{align*}
    (yz - [y, z])x &= yzx - [y, z]x \\
    \rightarrow (y(xz - [x, z]) - [y, z]x) &= yxz - y[x, z] - [y, z]x \\
    \rightarrow (xy - [x, y])z - y[x, z] - [y, z]x &= xyz - f \\
    z(xy - [x, y]) &= zxy - z[x, y] \\
    \rightarrow (xz - [x, z])y - z[x, y] &= xzy - [x, z]y - z[x, y] \\
    \rightarrow x(yz - [y, z]) - [x, z]y - z[x, y] &= xyz - g
\end{align*}
\]
with
\[
f = [x, y]z + y[x, z] + [y, z]x \quad \text{and} \quad g = x[y, z] + [x, z]y + z[x, y]
\]
Our goal is to show that the expression \((xyz - f) - (xyz - g) = g - f\) belongs to \(I_{zyx}\); see (D.7). To see that this is the case, we write \(g - f\) in the form
\[
g - f = [x, [y, z]]_{\mathbb{K}(X)} + [y, [z, x]]_{\mathbb{K}(X)} + [z, [x, y]]_{\mathbb{K}(X)}
\]
Recall that \([a, b]_{\mathbb{K}(X)} \equiv [a, b] \mod I\) for all \(a, b \in \langle X \rangle_{\mathbb{K}}\) by (D.8). In fact, since \(ab - ba - [a, b]\) is a \(\mathbb{K}\)-linear combination of the elements \(yx - (xy - [x, y])\) with \(x, y \in X\) and \(x < y\) and each \(yx - (xy - [x, y])\) belongs to \(I_w\) for any \(w \in W\) with \(\deg w > 2\), we have \([a, b]_{\mathbb{K}(X)} \equiv [a, b] \mod I_w\) for any \(w \in W\) with \(\deg w > 2\). Applying this with \(w = zyx\), we conclude that
\[
g - f \equiv [x, [y, z]] + [y, [z, x]] + [z, [x, y]] \mod I_{zyx}
\]
Finally, the Jacobi identity for \(g\) tells us that \([x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0\), whence \(g - f \in I_{zyx}\) as desired. This completes the proof of the Poincaré-Birkhoff-Witt Theorem.

\[\square\]

Exercises for Section D.5

D.5.1. Recall from class that the algebra \(O_q(\mathbb{K}(\mathbb{K}))\) of quantum 2 \(\times\) 2-matrices is given by the generators \(a, b, c\) and \(d\) satisfying the relations
\[
\begin{align*}
    ab &= q ba & ac &= qa & bc &= cb \\
    bd &= q db & cd &= q dc & ad - da &= \hat{q} bc
\end{align*}
\]
with \(\hat{q} = q - q^{-1}\). Quantum SL2 is defined as the \(\mathbb{K}\)-algebra
\[
O_q(\text{SL}_2(\mathbb{K})) = O_q(\mathbb{K}(\mathbb{K}))/(D_q - 1) \quad \text{with} \quad D_q = ad - qbc
\]
The element \(D_q\) is called the quantum determinant.

(a) Denoting the images of the generators of \(O_q(\mathbb{K}(\mathbb{K}))\) in \(O_q(\text{SL}_2(\mathbb{K}))\) by \(\bar{a}, \ldots, \bar{d}\), show that a defining set of relations for \(O_q(\text{SL}_2(\mathbb{K}))\) is given by
\[
\begin{align*}
    \bar{a} \bar{b} &= q \bar{b} \bar{a} & \bar{a} \bar{c} &= q \bar{c} \bar{a} & \bar{b} \bar{c} &= \bar{c} \bar{b} & \bar{a} \bar{d} &= q \bar{b} \bar{c} + 1 \\
    \bar{b} \bar{d} &= q \bar{d} \bar{b} & \bar{c} \bar{d} &= q \bar{d} \bar{c} & \bar{a} \bar{d} &= q^{-1} \bar{b} \bar{c} + 1
\end{align*}
\]
(b) Use the Diamond Lemma to find a $\mathbb{k}$-basis of $O_q(SL_2(\mathbb{k}))$.

**Hint.** The degree-lexicographic ordering of words that is based on $b < a < d < c$ together with the reduction system coming from the relations in (a) works for the Diamond Lemma.
The Symmetric Ring of Quotients

This appendix provides the ring theoretic background for the treatment of the Dixmier-Mœglin equivalence for enveloping algebras in §5.6.6. Throughout, $A$ will denote an arbitrary ring (associative, with 1) and ideals are understood to be two-sided ideals.

E.1. Definition and Basic Properties

Let $\mathcal{E} = \mathcal{E}(A)$ denote the collection of all ideals $I$ of $A$ having zero left and right annihilator in $A$:

\[
\text{ann}_A(I) = \{ a \in A \mid aI = 0 \} = 0 .
\]

and similarly $\text{r. ann}_A(I) = 0$. Note that $\mathcal{E}$ is closed under intersections and products: if $I, J \in \mathcal{E}$ then $I \cap J \in \mathcal{E}$ and $IJ \in \mathcal{E}$. The symmetric ring of quotients $Q_A$ was originally introduced by Kharchenko [121], [122]; it can be constructed as a direct limit:

\[
Q_A \overset{\text{def}}{=} \lim_{\longrightarrow \atop {I \in \mathcal{E}}} \mathcal{H}_I
\]

with

\[
\mathcal{H}_I := \{ (f, g) \in \text{Hom}(A_I, A_A) \times \text{Hom}(I_A, A_A) \mid (af)b = a(gb) \forall a, b \in I \}
\]

Here, $\text{Hom}(A_I, A_A)$ and $\text{Hom}(I_A, A_A)$ denote the collections of all left and right $A$-module maps $I \to A$, respectively. Furthermore, we have written $af = f(a)$ for $f \in \text{Hom}(A_I, A_A)$ and $gb = g(b)$ for $g \in \text{Hom}(I_A, A_A)$. We shall refer to the condition $(af)b = a(gb)$ for all $a, b \in I$ as “associativity.”
Proof. We first check the asserted properties (i) – (iv) for QA. Given two pairs, \((f, g) \in \mathcal{H}_I\) and \((f', g') \in \mathcal{H}_{I'}\) with \(I, I' \in \mathcal{E}\), we write 
\((f, g) \sim (f', g')\) if \(f|_J = f'|_J\) for some \(J \in \mathcal{E}\) with \(J \subseteq I \cap I'\). By associativity, the latter condition holds if and only if \(g|_J = g'|_J\); in fact, either condition, \(f|_J = f'|_J\) or \(g|_J = g'|_J\), implies that \(f|_{I \cap I'} = f'|_{I \cap I'}\) and \(g|_{I \cap I'} = g'|_{I \cap I'}\). It is easy to see that \(\sim\) defines an equivalence relation on the disjoint union \(\bigsqcup_{I \in \mathcal{E}} \mathcal{H}_I\). By definition, QA is the set of equivalence classes for \(\sim\). Denoting the class of \((f, g)\) by \([f, g]\), addition in QA is defined by pointwise addition of functions:

\[ [f, g] + [f', g'] := [f + f', g + g'] \]

with \(f + f', g + g' : I \cap I' \to A\). Multiplication in QA is defined by composition:

\[ [f, g] \cdot [f', g'] := [ff', gg'] \]

Here, \(ff'\) is again thought of as being written to the right of its input. Thus, \(ff' = f' \circ f : I'I \to I' \to R\) and also \(gg' = g \circ g' : I'I \to I \to R\). This makes QA into a ring with 1 = [Id\(_A\), Id\(_A\)].

The following proposition extracts the operative ring theoretic facts from this construction and gives an abstract characterization of QA.

**Proposition E.1.** The symmetric ring of quotients QA has the following properties:

(i) There is a ring embedding \(A \hookrightarrow QA\).

(ii) For each \(q \in QA\), there exists \(I \in \mathcal{E}\) with \(qI \subseteq A\) and \(Iq \subseteq A\).

(iii) If \(I \in \mathcal{E}\), then \(\text{ann}_{\text{QA}}(I) = 1.\text{ann}_{\text{QA}}(I) = 0\).

(iv) Given \((f, g) \in \mathcal{H}_I\) with \(I \in \mathcal{E}\), there exists \(q \in QA\) with \(aq = af\) and \(qb = gb\) for all \(a, b \in I\).

Furthermore, QA is determined by these properties: If \(Q\) is any ring satisfying (i) – (iv), then there is a unique isomorphism \(QA \to Q\) that is the identity on \(A\).

**Proof.** We first check the asserted properties (i) – (iv) for QA.

(i) Put \(\mu(a) = [\rho_a, \lambda_a]\), where \(\rho_a, \lambda_a : A \to A\) denote right and left multiplication by \(a \in A\), respectively. This yields the desired ring embedding \(\mu : A \hookrightarrow QA\). We will treat this embedding as an inclusion below, that is, we will identify each \(a \in A\) with \(\mu(a) \in QA\).

(ii) Let \(q = [f, g]\) with \((f, g) \in \mathcal{H}_I\) and \(I \in \mathcal{E}\). Then, for \(a, b \in I\), one computes \(g(\lambda_a) b = g(ab) = (ga)b = \lambda_g a b\) and \(b(f \rho_a) = (bf)a = b(ga) = b\rho_g a\). Therefore, \([f, g] \mu(a) = [f \rho_a, g \lambda_a] = [\rho_g a, \lambda_g a] = \mu(ga)\) or

\[ (E.1) \quad qa = ga \quad (a \in I) . \]
This shows that \( qI \subseteq A \). Similarly,
\[
qI = \begin{cases} \frac{af}{a} \ (a \in I) \\ \end{cases}
\]
and so \( Iq \subseteq A \).

(iii) Suppose that \( q = [f, g] \) with \( (f, g) \in \mathcal{H}_f \) satisfies \( qI = 0 \). Then (E.1) shows that \( 0 = ga = ga \) for all \( a \in I \cap J \); hence, \( g|_{I \cap J} = 0 \). As we remarked above, this implies that \( (f, g) \sim (0, 0) \) or \( q = 0 \) (Exercise E.3.1). Similarly, using (E.2) instead of (E.1), \( Iq = 0 \) implies \( q = 0 \).

(iv) In view of (E.1) and (E.2) above, (iv) is clear with \( q = [f, g] \in Q \). This proves properties (i) – (iv). For uniqueness, suppose that \( Q \) and \( Q' \) both satisfy (i) – (iv) and let \( q \in Q \). Choose \( I \in \mathcal{E} \) with \( qI, Iq \subseteq A \) as in (ii) and define \( (f, g) \in \mathcal{H}_f \) by \( af = aq \) and \( ga = qa \); the associativity condition holds, because \( (af)b = (aq)b = a(qb) = a(gb) \) for \( a, b \in I \). Hence (iv) implies that there exists \( q' \in Q' \) with \( aq' = af \) and \( q'b = gb \) for all \( a, b \in I \). By (iii), the element \( q' \in Q' \) is in fact uniquely determined by the condition \( aq = aq' \) for all \( a \in I \). It is straightforward to check that \( q \mapsto q' \) gives an isomorphism \( Q \to Q' \) that is the identity on \( A \). If \( q \mapsto q'' \) is another isomorphism \( Q \to Q' \) that is the identity on \( A \), then \( aq'' = (aq)'' = aq \) for \( a \in I \) and so \( q' = q'' \).

For any \( q \in QA \), we define an ideal of \( A \) by
\[
D_q = \{ a \in A \mid qa \subseteq A \text{ and } aAq \subseteq A \}
\]
By Proposition E.1(ii), we always have \( D_q \in \mathcal{E} \).

### E.2. The Extended Center

The center of QA will be denoted by
\[
CA \overset{\text{def}}{=} \mathcal{Z}(QA)
\]
The ring \( CA \) is called the **extended center** of \( A \). The extended center coincides with the centralizer of \( A \) in QA, and hence it contains the ordinary center \( \mathcal{Z}A \):

\[
\mathcal{Z}A \subseteq CA = \{ q \in QA \mid qa = aq \forall a \in A \}
\]

Indeed, assume that \( q \in QA \) satisfies \( qa = aq \) for all \( a \in I \) \( (I \in \mathcal{E}) \) and let \( q' \in QA \) be arbitrary. Then, for any \( b \in D_{q'} \), we have \( ab, abq' \in I \) and so we compute \( (abq')q = q(abq') = abq' \). Thus, \( 1D_{q'}(q'q - qq') = 0 \), which forces \( q'q = qq' \) by Proposition E.1(iii). This proves the non-trivial inclusion in (E.4).
In terms of the original definition of QA, the extended center can be described as follows. Let $q = [f,g] \in QA$ with $(f,g) \in H_1$ and $I \in \mathcal{E}$. Then:

\begin{align}
E.5 \quad q \in C(A) \iff f = g \iff f \text{ or } g \text{ is an } (A,A)\text{-bimodule map}
\end{align}

Here, the first $\iff$ is immediate from formulas (E.1), (E.2) in conjunction with (E.4). As for the second $\iff$, the direction $\Rightarrow$ is clear and $\Leftarrow$ follows easily from associativity of the pair $(f,g)$.

**Proposition E.2.**

(a) For any $P \in \text{Spec } A$, the extended center $C(A/P)$ is a field. Furthermore, if $I$ is semiprime ideal of $A$ such that $C(A/I)$ is a domain, then $I$ is in fact prime.

(b) Assume that $P \in \text{Prim } A$, say $P = \text{Ker } V$ with $V \in \text{ Irr } A$. Then the embedding $\mathcal{Z}(A/P) \hookrightarrow \mathcal{Z}(D(V))$, $a + P \mapsto a_V$, extends to an embedding of fields $C(A/P) \hookrightarrow \mathcal{Z}(D(V))$.

**Proof.** (a) First, let us assume that $A$ is prime and show that every $0 \neq q \in CA$ has an inverse. For this, consider the ideal $D_q$ as in (E.3) and recall that $D_q \in \mathcal{E}$. Consequently, $I_q = qD_q$ is a nonzero ideal of $A$ by Proposition E.1(iii). Since $A$ is prime, it follows that $I_q \in \mathcal{E}$. Therefore, the map $f: D_q \to I_q$, $a \mapsto qa$, is an isomorphism of $(A,A)$-bimodules. Since $q = [f,f]$, the desired inverse of $q$ is given by $[f^{-1}, f^{-1}]$.

Next, assume that $A$ is semiprime but not prime. Then there exist nonzero ideals $I, J$ of $A$ such that $IJ = 0$. We may assume that $J = \text{ ann}_A(I)$. Since $A$ is semiprime, the sum $I + J$ is direct and belongs to $\mathcal{E}$ (Exercise E.3.1). Define $(A, A)$-bimodule maps $f, f': I + J \to A$ by $f(i + j) = i$ and $f'(i + j) = j$ and put $q = [f,f], q' = [f', f'] \in CA$. Then, $0 \neq q, q'$ but $qq' = 0$, whence $CA$ is not a field.

(b) We may assume that $P = 0$; so $A$ embeds into $\text{End}(V)$ via $a \mapsto a_V$, and this map restricts to the canonical embedding $\mathcal{Z}A \hookrightarrow \mathcal{Z}(D(V))$ in the proposition. For a given $q \in CA$, we wish to define an endomorphism $\delta_q \in \mathcal{Z}(D(V))$ so that $\delta_q = a_V$ for $q = a \in \mathcal{Z}A$. To this end, note that $V = D_q . V$; so all elements of $V$ have the form $\sum_i d_i . v_i$ for suitable $d_i \in D_q$, $v_i \in V$. Since $q_i d_i \in A$, we may put

\[ \delta_q(\sum_i d_i . v_i) := \sum_i (qd_i). v_i \in V \]

To see that this is definition is well-defined, assume that $\sum_i d_i . v_i = \sum_j d'_j . v'_j$ with $d'_j \in D_q$, $v'_j \in V$. Then

\[ D_q . (\sum_i (qd_i). v_i - \sum_j (qd'_j). v'_j) = D_q q. (\sum_i d_i . v_i - \sum_j d'_j . v'_j) = 0 \]

By irreducibility and faithfulness of $V$, it follows that $\sum_i (qd_i). v_i = \sum_j (qd'_j). v'_j$, proving well-definedness. In particular, if $q = a \in \mathcal{Z}A$ and $v \in V$, then $\delta_a(v) =\]
E.3. Comparison with Other Rings of Quotients

\[ \delta_a(1.v) = a.v; \text{ so } \delta_a = a_V. \] The fact that, in general, \( \delta_q \in \mathcal{Z}(D(V)) \) follows from the computations, for any \( d \in D_q, v \in V, a \in A \) and \( \delta \in D(V) \),

\[ \delta_q(ad.v) = (qad).v = (aqd).v = a.\delta_q(d.v) \]

and

\[ \delta\delta_q(d.v) = \delta((qd).v) = (qd).\delta(v) = \delta_q(d.\delta(v)) = \delta_q\delta(d.v) \]

It remains to check that the map \( CA \to \mathcal{Z}(D(V)) \), \( q \mapsto \delta_q \), is a ring homomorphism; injectivity will then be automatic, because \( CA \) is a field by (a). So let \( q, q' \in C(A) \) be given. Then \( D_q \cap D_{q'} \subseteq D_{q+q'} \) and \( V = (D_q \cap D_{q'})V \). With \( d \in D_q \cap D_{q'} \) and \( v \in V \), we compute

\[ \delta_{q+q'}(d.v) = ((q + q')d).v = (qd).v + (q'd).v = \delta_q(d.v) + \delta_{q'}(d.v) \]

Thus, \( \delta_{q+q'} = \delta_q + \delta_{q'} \). Similarly, \( D_qD_{q'} \subseteq D_{qq'} \) and \( V = D_qD_{q'}V \). For \( d \in D_q, d' \in D_{q'} \) and \( v \in V \), one has

\[ \delta_{qq'}(dd'.v) = (qq'dd').v = (qd).(q'd').v = \delta_q(d.\delta_{q'}(d'.v)) = \delta_q\delta_{q'}(dd'.v) \]

This shows that \( \delta_{qq'} = \delta_q\delta_{q'} \), thereby finishing the proof of the proposition.  

E.3. Comparison with Other Rings of Quotients

The symmetric ring of quotients \( QA \) is a comparatively recent addition to the arsenal of quotient rings in ring theory. Below, we offer a few brief comments on some predecessors of \( QA \); none of this material is actually needed in this book. The main point is that, for enveloping algebras and related algebras, the extended center \( CA \) can be defined using any of the alternative rings of quotients to be discussed below.

The reader wishing to see more details is referred to Passman [165, Chapter 3] and [166, Chapters 24 and 25].

Martindale [145] (for prime rings \( A \)) and Amitsur [4] (in general), introduced right and left quotient rings that, like \( QA \), are defined as direct limits:

\[ Q_r(A) \overset{\text{def}}{=} \lim_{\longleftarrow I \in \mathcal{E}_r} \text{Hom}(I, A) \quad \text{and} \quad Q_l(A) \overset{\text{def}}{=} \lim_{\longrightarrow I \in \mathcal{E}_l} \text{Hom}(A, I) \]

Here, \( \mathcal{E}_r \) and \( \mathcal{E}_l \) denote the sets all ideals of \( A \) having zero left and right annihilator, respectively. We shall mostly be concerned with rings \( A \) where \( \mathcal{E}_l = \mathcal{E}_r = \mathcal{E} \); this holds, for example, if \( A \) is semiprime (Exercise E.3.1). The map \( [f, g] \mapsto [f] \) then embeds \( QA \) into \( Q_r(A) \); similarly \( [f, g] \mapsto [g] \) works for \( Q_l(A) \). In terms of these embeddings, the symmetric ring of quotients can be described as follows:

\[ QA \cong \{ q \in Q_r(A) \mid qI \subseteq A \text{ for some } I \in \mathcal{E}_r \} \]

\[ \cong \{ q \in Q_l(A) \mid Iq \subseteq A \text{ for some } I \in \mathcal{E}_l \} \]

These isomorphisms restrict to isomorphisms of the centers of \( QA, Q_r(A) \) and \( Q_l(A) \), because the centers are identical to the centralizer of \( A \) in each case. Thus,
if \( A \) is semiprime, then we have the following descriptions for the extended center:

\[ CA \cong \mathcal{Z}(Q_\ell(A)) \cong \mathcal{Z}(Q_r(A)) \]

For semiprime \( A \), one can similarly show that \( CA \) is also isomorphic to the center of the so-called maximal ring of quotients, \( Q_{\text{max}}(A) \).

The original literature on enveloping algebras of finite-dimensional Lie algebras employs uses yet another ring of quotients, the so-called \textit{classical ring of quotients} \( Q_{\text{cl}}(A) \). Unlike the aforementioned rings of quotients, \( Q_{\text{cl}}(A) \) need not exist for all \( A \). However, if \( A \) is semiprime noetherian, then \( Q_{\text{cl}}(A) \) does exist and is in fact isomorphic to \( Q_{\text{max}}(A) \) \cite[Proposition 4.6.2]{1}. Thus, in this case, we also have

\[ CA \cong \mathcal{Z}(Q_{\text{cl}}(A)) \]

Our main reason for favoring the symmetric ring of quotients \( QA \), besides its apparent left-right symmetry, is that it always exists and is the most economical of the above quotient rings to accommodate the extended center \( CA \). Propositions E.1 and E.2 contain all that we shall need (and more) about \( QA \) and \( CA \).

\textbf{Exercises for Section E.3}

\textbf{E.3.1} (Some details for the symmetric ring of quotients). This exercise asks you to verify some claims that were made in the course of our discussion of the symmetric ring of quotients \( QA \) in this section.

(a) Let \((f, g) \in \mathcal{H}_I\) and \((f', g') \in \mathcal{H}_{I'}\) with \( I, I' \in \mathcal{F} \). Assuming that \( f|_J = f'|_J \) or \( g|_J = g'|_J \) for some \( J \in \mathcal{E} \) with \( J \subseteq I \cap I' \), show that \( f|_{I \cap I'} = f'|_{I \cap I'} \) and \( g|_{I \cap I'} = g'|_{I \cap I'} \).

(b) Check that \( \sim \) defines an equivalence relation on \( \bigsqcup_{I \in \mathcal{E}} \mathcal{H}_I \).

(c) Assume that \( A \) is semiprime. Show that, for every ideal \( I \) of \( A \), one has \( r \cdot \text{ann}_A(I) = r \cdot \text{ann}_A(I) \). Moreover, denoting this annihilator by \( \text{ann}_A(I) \), show that the sum \( I + \text{ann}_A(I) \) is direct and belongs to \( \mathcal{E} \).

\textbf{E.3.1}

\textbf{E.3.2} (Central closure). Assume that \( A \) is prime and consider the subring \( A' := A(CA) \) of the symmetric ring of quotients \( QA \); this subring is called the \textit{central closure} of \( A \). Show that \( A' \) is prime. Moreover, if \( 0 \neq x, y \in A' \) are such that \( xay = yax \) for all \( a \in A \), then \( x = zy \) for some \( z \in CA \).

\textbf{E.3.2} Primeness of \( A' \) follows from Proposition E.1: every nonzero ideal of \( A' \) has nonzero intersection with \( A \). For the second assertion, see Martindale \cite[Theorem 1]{145}. 

---

\textbf{E.3.1} (Some details for the symmetric ring of quotients). This exercise asks you to verify some claims that were made in the course of our discussion of the symmetric ring of quotients \( QA \) in this section.

(a) Let \((f, g) \in \mathcal{H}_I\) and \((f', g') \in \mathcal{H}_{I'}\) with \( I, I' \in \mathcal{F} \). Assuming that \( f|_J = f'|_J \) or \( g|_J = g'|_J \) for some \( J \in \mathcal{E} \) with \( J \subseteq I \cap I' \), show that \( f|_{I \cap I'} = f'|_{I \cap I'} \) and \( g|_{I \cap I'} = g'|_{I \cap I'} \).

(b) Check that \( \sim \) defines an equivalence relation on \( \bigsqcup_{I \in \mathcal{E}} \mathcal{H}_I \).

(c) Assume that \( A \) is semiprime. Show that, for every ideal \( I \) of \( A \), one has \( r \cdot \text{ann}_A(I) = r \cdot \text{ann}_A(I) \). Moreover, denoting this annihilator by \( \text{ann}_A(I) \), show that the sum \( I + \text{ann}_A(I) \) is direct and belongs to \( \mathcal{E} \).

\textbf{E.3.1}

\textbf{E.3.2} (Central closure). Assume that \( A \) is prime and consider the subring \( A' := A(CA) \) of the symmetric ring of quotients \( QA \); this subring is called the \textit{central closure} of \( A \). Show that \( A' \) is prime. Moreover, if \( 0 \neq x, y \in A' \) are such that \( xay = yax \) for all \( a \in A \), then \( x = zy \) for some \( z \in CA \).

\textbf{E.3.2} Primeness of \( A' \) follows from Proposition E.1: every nonzero ideal of \( A' \) has nonzero intersection with \( A \). For the second assertion, see Martindale \cite[Theorem 1]{145}. 

---

\textbf{E.3.1} (Some details for the symmetric ring of quotients). This exercise asks you to verify some claims that were made in the course of our discussion of the symmetric ring of quotients \( QA \) in this section.

(a) Let \((f, g) \in \mathcal{H}_I\) and \((f', g') \in \mathcal{H}_{I'}\) with \( I, I' \in \mathcal{F} \). Assuming that \( f|_J = f'|_J \) or \( g|_J = g'|_J \) for some \( J \in \mathcal{E} \) with \( J \subseteq I \cap I' \), show that \( f|_{I \cap I'} = f'|_{I \cap I'} \) and \( g|_{I \cap I'} = g'|_{I \cap I'} \).

(b) Check that \( \sim \) defines an equivalence relation on \( \bigsqcup_{I \in \mathcal{E}} \mathcal{H}_I \).

(c) Assume that \( A \) is semiprime. Show that, for every ideal \( I \) of \( A \), one has \( r \cdot \text{ann}_A(I) = r \cdot \text{ann}_A(I) \). Moreover, denoting this annihilator by \( \text{ann}_A(I) \), show that the sum \( I + \text{ann}_A(I) \) is direct and belongs to \( \mathcal{E} \).

\textbf{E.3.1}

\textbf{E.3.2} (Central closure). Assume that \( A \) is prime and consider the subring \( A' := A(CA) \) of the symmetric ring of quotients \( QA \); this subring is called the \textit{central closure} of \( A \). Show that \( A' \) is prime. Moreover, if \( 0 \neq x, y \in A' \) are such that \( xay = yax \) for all \( a \in A \), then \( x = zy \) for some \( z \in CA \).

\textbf{E.3.2} Primeness of \( A' \) follows from Proposition E.1: every nonzero ideal of \( A' \) has nonzero intersection with \( A \). For the second assertion, see Martindale \cite[Theorem 1]{145}. 

---

\textbf{E.3.1} (Some details for the symmetric ring of quotients). This exercise asks you to verify some claims that were made in the course of our discussion of the symmetric ring of quotients \( QA \) in this section.

(a) Let \((f, g) \in \mathcal{H}_I\) and \((f', g') \in \mathcal{H}_{I'}\) with \( I, I' \in \mathcal{F} \). Assuming that \( f|_J = f'|_J \) or \( g|_J = g'|_J \) for some \( J \in \mathcal{E} \) with \( J \subseteq I \cap I' \), show that \( f|_{I \cap I'} = f'|_{I \cap I'} \) and \( g|_{I \cap I'} = g'|_{I \cap I'} \).

(b) Check that \( \sim \) defines an equivalence relation on \( \bigsqcup_{I \in \mathcal{E}} \mathcal{H}_I \).

(c) Assume that \( A \) is semiprime. Show that, for every ideal \( I \) of \( A \), one has \( r \cdot \text{ann}_A(I) = r \cdot \text{ann}_A(I) \). Moreover, denoting this annihilator by \( \text{ann}_A(I) \), show that the sum \( I + \text{ann}_A(I) \) is direct and belongs to \( \mathcal{E} \).

\textbf{E.3.1}

\textbf{E.3.2} (Central closure). Assume that \( A \) is prime and consider the subring \( A' := A(CA) \) of the symmetric ring of quotients \( QA \); this subring is called the \textit{central closure} of \( A \). Show that \( A' \) is prime. Moreover, if \( 0 \neq x, y \in A' \) are such that \( xay = yax \) for all \( a \in A \), then \( x = zy \) for some \( z \in CA \).

\textbf{E.3.2} Primeness of \( A' \) follows from Proposition E.1: every nonzero ideal of \( A' \) has nonzero intersection with \( A \). For the second assertion, see Martindale \cite[Theorem 1]{145}.
Bibliography


61. Michel Duflo, Certaines algèbres de type fini sont des algèbres de Jacobson, J. Algebra 27 (1973), 358–365. MR 0344279 (49 #9018)


69. Walter Feit and John G. Thompson, Solvability of groups of odd order, Pacific J. Math. 13 (1963), 775–1029. MR 0166261 (29 #3538)


91. ______, *Éléments de géométrie algébrique. IV. Étude locale des schémas et des morphismes de
92. Alexandre Grothendieck et al., *Revêtements étalés et groupe fondamental (SGA 1)*, Documents
algébrique du Bois Marie 1960–61, directed by A. Grothendieck, with two papers by M.
Raynaud, updated and annotated reprint of the 1971 original [Lecture Notes in Math., 224,
Springer, Berlin]. MR 2004g:14017
93. Harish-Chandra, *On some applications of the universal enveloping algebra of a semisimple Lie
Texts in Mathematics, No. 52. MR 0463157 (57 #3116)
MR MR0175950 (31 #226)
MR 0088486 (19.527a)
97. Gerhard P. Hochschild, *Basic theory of algebraic groups and Lie algebras*, Graduate Texts in
98. James E. Humphreys, *Introduction to Lie algebras and representation theory*, Springer-Verlag,
in Mathematics, No. 21. MR 0396773
100. ______, *Representations of semisimple Lie algebras in the BGG category O*, Graduate Studies
102. I. Martin Isaacs, *Character theory of finite groups*, AMS Chelsea Publishing, Providence,
MR 2270898
(1951), 5–6. MR 0044528 (13,432a)
MR 0047026 (13,816c)
105. ______, *Lectures in abstract algebra*, Graduate Texts in Mathematics, Vol. 31, Springer-Verlag,
New York-Berlin, 1975, Volume II: Linear algebra, Reprint of the 1953 edition [Van Nostrand,
Toronto, Ont.], Graduate Texts in Mathematics, No. 31. MR 0369381 (51 #5614)
original. MR 559927
MR 86d:00001
ematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)], vol. 3,
Springer-Verlag, Berlin, 1983. MR 721170
109. ______, *Representations of algebraic groups*, second ed., Mathematical Surveys and Mono-
(2004h:20061)
110. G. I. Kac, *Certain arithmetic properties of ring groups*, Funkcional. Anal. i Priložen. 6 (1972),
no. 2, 88–90. MR 0304552 (46 #3687)


156. Grigori˘ı Ol’shanski˘ı, Extension of the algebra U(g) for infinite-dimensional classical Lie algebras g, and the Yangians Y(gl(m)), Dokl. Akad. Nauk SSSR 297 (1987), no. 5, 1050–1054. MR 936073 (89g:17017)


199. Alfred Young, The collected papers of Alfred Young (1873–1940), University of Toronto Press, Toronto, Ont., Buffalo, N. Y., 1977, With a foreword by G. de B. Robinson and a biography by H. W. Turnbull, Mathematical Expositions, No. 21. MR 0439548


Index

R-points, 500
Φ-convex, 386
Φ-extremal element, 386
σ-string, 386
(co)homology
  Hochschild, 285

abelianization, 576
  of a group, 126, 501, 576
  of a Lie algebra, 502
  of an algebra, 500
action, 13, 439
  locally finite, 435, 440
  multiplicative, 307, 364
  of a Lie algebra, 249
  rational (of an algebraic group), 510
ad-nilpotent, 321, 332, 393
ad-semisimple, 321
Adams operations, 487
Adams, J. F., 487
additive group
  functor, 502
adjoint action
  on bimodules, 284
adjoint representation, 387
adjunction, 575, 576
admissible transposition, 206, 212, 221
Ado’s Theorem, 248
affine algebraic k-group, 504
affine algebraic group
  connected, 516
affine algebraic variety, 43
affine group scheme, 500
affine scheme, 500
algebra
  graded, 491
algebra, 3
  affine, 9, 36, 40, 43, 95, 269, 493
  almost commutative, 275, 286
  alternating, 22
  anticommutative, 12, 22
  artinian, 40, 472
  associated graded, 267, 269, 275, 286, 309
  augmented, 559
  central simple, 22, 76, 106
  commutative, 4, 9, 40, 135
  connected, 503
  defined over a finite field, 77
  defined over a subfield, 65, 147
  filtered, 275, 286, 309
  finite-dimensional, 6
  free, 6–9, 597
INDEX

free commutative, 9
Frobenius, 96–109, 540, 559
graded, 7, 8, 21, 177
graded commutative, 22
graded-commutative, 12
local, 38, 96
noetherian, 19, 40, 269
opposite, 14
PI (polynomial identity), 50, 276, 437
prime, 63
primitive, 63
self-injective, 559
semiprime, 63, 64
semiprimitive, 48, 63
semisimple, 51, 57, 63, 64
separable, 65, 76, 105, 147, 549
simple, 22, 169
split semisimple, 59, 102
symmetric, 96–109, 122
symmetric (of a vector space), 10, 138
algebra of invariants, 174, 280
algebra of polynomial functions, 174, 446, 593
algebraic \( n \)-torus, 507
algebraic group, 499–535
linear, 513–515
alternating, 11
ambiguity, 599
resolvable, 599
resolvable relative to \( \leq \), 602, 603
Amitsur, S.A., 611
anti-invariant, 366
anticommutativity, 244
antipode, 135, 277, 448, 450
of an enveloping algebra, 274
antisymmetrization, 179
antisymmetrizer, 179, 366, 480
Artin-Tate Lemma, 20, 40, 276, 496
ascending central series, 252
associative law, 4, 15, 568
for tensor products, 582, 584
augmentation ideal, 132, 273
relative, 141, 275, 459, 475
augmentation map, 132, 273
automorphism
inner, 39
of graphs, 196
automorphism group, 568
averaging operator, 134
Axiom of Choice, 19, 574
balanced, 580, 584
base
for the root system \( A_n \), 348
for the root system \( B_n \), 348
for the root system \( C_n \), 348
for the root system \( D_n \), 349
of a root system, 347
Bergman, G., 51, 521, 597, 599
bi-commutant, 36
bi-ideal, 445
bialgebra, 444
bifunctor, 20, 136, 463, 465, 574, 580, 582, 585, 587
bilinear form
alternating, 244, 563
associative, 314, 430
\( G \)-invariant, 564
invariant, 195, 563
non-degenerate, 314, 318, 430, 563
skew-symmetric, 563
symmetric, 563
bilinear map, 584
bimodule, 15, 570, 581
regular, 15, 95, 97, 570, 582
Borel subalgebra, 374
Bourbaki, 261, 364
braid relation, 212
branching graph, 185, 190, 191
branching rule, 198, 221, 226, 231
Brauer, R., 121, 124
Brauer-Fowler Theorem, 163–165, 561
Bruhat order, 532, 533, 535
Burnside’s $p^aq^b$-Theorem, 162, 163
Burnside’s Lemma, 134
Burnside’s Theorem, 60–63, 70, 77, 127, 129, 161, 308, 559
Burnside, W., 163, 476
Cartan homomorphism, 86, 88, 92
Cartan integers, 326, 331, 334, 337, 340
Cartan matrix
  of a finite-dimensional algebra, 92
  of a root system, 356, 360
  of the Taft algebra, 481
  of type $A_n - D_n$, 356
  of type $A_n$, 334
  of type $B_n$, 337
Cartan subalgebra, 324
Cartan’s Criterion, 261–263, 315
Cartan, E., 261
Cartier-Gabriel-Kostant Theorem, 492
  of a representation, 319
  universal, 319
Casimir operator, 100, 101, 105
Casimir operator and Higman trace, 110
Casimir trace, 101, 102, 123, 553, 554
Catalan number, 203
category, 567
  concrete, 570
  locally small, 568
  small, 568
Cauchon diagram, 532
Cauchon, G., 531
Cauchy-Frobenius Lemma, 134
center, 246
  of a bimodule, 464
  of a Lie algebra, 246
  of an algebra, 3
  central character, 381
  central closure, 612
  centralizer, 36, 130, 187, 221, 222, 235, 246, 398, 406
  centralizing homomorphism, 50, 497
chain condition
  ascending (ACC), 19, 41
  descending (DCC), 41, 472, 598
color
  central, 290
  complex, 148
  formal, 389, 412
    of an $sl_2$-module, 305
    of Verma modules, 413, 414
  irreducible, 70
    of a representation, 66
    of an induced representation, 123
  regular, 66, 76, 100, 102, 103, 106, 130, 146
character table, 121, 224
  of $A_4$, 156
  of $A_5$, 158
  of $D_4$, 128
  of $S_3$, 129
  of $S_4$, 154
  of $S_5$, 155
characteristic polynomial, 31, 173, 495
Chevalley property, 479, 480, 493, 522
Chevalley Restriction Theorem, 397
Chevalley’s Theorem, 479, 525
Chevalley, C., 522, 525
Chin, W., 527
Chinese Remainder Theorem, 63, 71, 126, 127
class equation, 556–558
class function, 120, 164, 564
class sum, 133, 162
classical Lie algebras, 332
classification of finite simple groups, 164, 481, 493
Clifford theory, 166–168
Clifford’s Theorem, 157, 161, 167
coaction, 439
coalgebra, 426
  cocommutative, 426
  coinside, 427, 439
  group-like, 427, 433, 437, 440, 446, 491
  simple, 441
coisassociativity, 135, 274, 426, 429
coisassociativity, 430
cobordism diagram, 429, 430
coform, 430
coidal, 426
  left, 440
  right, 440, 445
coidal subalgebra
  left, 467, 559
  right, 445, 463, 471
coinduction functor, 27, 95, 117
coinvariants, 140
  of a comodule, 462
comodule diagram, 568
comodule, 13, 438
  irreducible, 441
  regular, 440
  trivial, 462
complement, 18, 52
complete flag, 258
composition
  of a positive integer, 227, 237
composition (weak)
  of a positive integer, 180
composition factor, 35, 65, 91
composition series, 32, 72, 91
comultiplication, 135, 274, 277, 426
conditional expectation, 40
conjugacy classes, 120
  of \( A_n \), 155
  of \( S_n \), 153
  \( p \)-regular, 121, 483
Conjugacy Theorem, 393, 394, 397, 402
conjugate elements in a group, 120
conjugate partition, 196
conjugation, 120
connected component (of a
topological space), 519
content
  of a \( \lambda/\mu \)-tableau, 220
content of a Young tableau, 204
convolution, 432
  algebra, 432, 437
core, 294
  \( G \)-core, 526
  \( H \)-core, 496, 534, 535
  \( p \)-core, 148
coroot, 347
counit, 132, 273, 407, 426
counit laws, 426, 430, 431
Coxeter generators, 190, 533
crossed product, 166, 169
cyclotomic polynomial, 126
Dedekind’s Lemma, 459
defining representation
  of \( sl_2 \), 299
  of \( sl_{n+1} \), 386
degree
  of a representation, 24
  of a word, 600
degree-lexicographic order, 600
deleted permutation representation,
  128, 131
derivation, 249, 279
  inner, 249
  left \( \sigma \)-, 20
  of an algebra, 20
  outer, 249, 251, 255, 316
derived series, 253
descending central series, 252
determinant, 235, 448
diagonal action, 137
diagonal map, 274
diagonalizable, 254
diamond algebra, 253, 255
diamond conditions, 599, 603
Diamond Lemma, 529, 597–605
diamond Lie algebra, 250, 251
dihedral group, 127
dimension
  of a representation, 24
dimension augmentation, 485
direct product
  of algebras, 56
  of Lie algebras, 250, 272
direct sum
  of representations, 24
distinguished group-like element, 539
division algebra, 30
Dixmier, J., 291, 295
Dixmier-Mœglin equivalence, 293, 295, 297, 403
  for enveloping algebras, 295, 607
domain, 20, 268, 275
double centralizer, 36, 60
Double Centralizer Theorem, 232
doubly transitive group action, 131
Drinfel’d, V., 470
dual bases, 81
  for a Frobenius form, 98, 99
Dual Bases Lemma, 81
dual basis, 587
Dual Basis Lemma, 83
dual numbers, 523
dual root system, 347
duality functor, 571
Duflo isomorphism, 391
Duflo, M., 286
Dynkin diagram, 356
  of type $A_n - D_n$, 356
eigenspace, 54
  generalized, 260
Eilenberg swindle, 95
element
  group-like, 427, 445, 447, 539
  primitive, 445, 447, 449
  right modular, 539, 540, 545
elementary automorphism (of a Lie algebra), 324, 393, 406
elementary divisor, 37
Engel’s Theorem, 255–258, 290, 295
Engel, F., 256
enveloping algebra, 243, 264–266, 428, 433
  comultiplication in, 428
  Hopf ideals of, 459
  of $\mathfrak{sl}_2$, 597, 601
  restricted, 312, 446
equivalence
  of categories, 14, 574, 575
  of representations, 25, 39
Erdős-Kaplansky Theorem, 465, 587
Erdmann, K., 141
Etingof, P., 549
exact sequence, 177, 585
exponent, 124
extended center, 293, 609
extension of scalars, 29, 39, 64, 65, 76
exterior algebra, 11, 138
Feit-Thompson Theorem, 164
Ferrers diagram, 196
field
  of algebraic numbers, 219
field extension
  finite, 66
  finite separable, 76
  separable, 64, 66
filtration
  of a representation, 41
  of a vector space, 267
of an algebra, 267, 269
finite dual, 68, 434
finite rank, 588
Finiteness Theorem
for coalgebras and comodules, 441, 472, 561
fixed point, 130
forgetful functor, 7, 570, 577
form
of degree $n$, 182
formal character, 173, 305
formal linear combinations, 577
Frobenius endomorphism, 69
Frobenius extension, 117
of algebras, 110
Frobenius form, 98, 540
Frobenius reciprocity, 28, 61, 119, 144, 198, 221, 228
Frobenius’ Divisibility Theorem, 160, 161, 554, 555, 557
Frobenius’ formula, 168
Frobenius, F. G., 160, 185
Frobenius-Schur indicator, 165, 168, 561, 562, 566
functor, 570, 580, 582, 585
adjoint, 7, 10, 115, 264, 435, 438, 449, 450, 575
left, 575
right, 575
contravariant, 571, 587
covariant, 571
essentially surjective, 574
exact, 81, 124, 581
faithful, 570
full, 570
left exact, 81, 585
right exact, 581
Fundamental Theorem
of $S_n$-invariants, 369, 402
Fundamental Theorem on Coalgebras, 440
fundamental weights, 372
for the root system $A_n$, 361
$G$-algebra, 138, 280, 488
$g$-algebra, 280, 488
$G$-ideal, 526
$G$-prime ideal, 527
$G$-set, 130, 140, 141, 149
Galois descent, 64
Geiger, J., 531
Gelaki, S., 549
Gelfand-Zetlin (GZ) algebra, 186, 187, 193
Gelfand-Zetlin (GZ) basis, 192, 194, 195, 215
of $V_{n-1}$, 192, 195, 216, 218, 219
general linear group
functor, 503
general linear Lie algebra, 245
generalized matrix algebra, 16
generating function, 169–171
generators
of a module, 19
of an algebra, 6
Generic Flatness Lemma, 287, 593, 594
generic matrix, 454, 456
generic point, 46, 48, 289
Goodearl, K., 528, 533
Goodearl-Letzter stratification, 528
Graph Isomorphism Theorem, 185, 186, 196, 197, 203, 212–214
Grothendieck group
of finite-dimensional projectives, 85
of finite-dimensional representations, 72, 283, 306, 388, 481
of finitely generated projectives, 84
Grothendieck, A., 486, 494, 593
group
alternating, 175
group algebra, 113, 114, 363, 449
  Hopf subalgebras and Hopf ideals of, 458
group functor, 500
group ring, 116
  integral, 482
group scheme, 499
group-like element, 449, 454–456, 471, 545
groupoid, 569

$H$-comodule algebra, 488
$H$-ideal, 496
$H$-module algebra, 280, 488
$H$-prime, 496
Harish-Chandra, 399
Harish-Chandra homomorphism, 398
Harish-Chandra isomorphism, 399, 402, 406
Hattori’s Lemma, 124
Hattori, A., 83
Hattori-Stallings rank, 87, 92, 94, 124
Hattori-Stallings trace, 83
head, 89
  of a representation, 379
heart, 294
height
  of a root, 354
Heisenberg Lie algebra, 248, 250–252, 260, 263, 276, 291
Hermite reciprocity, 311
Hermitian inner product, 146
Hessian, 312
highest root, 387
highest weight representation, 377
highest weight vector, 378
Higman ideal, 123
Higman trace, 100, 101
Higman, D. G., 100, 105
Hilbert Basis Theorem, 19, 20, 270
Hilbert series, 174, 177
Hilbert-Serre Theorem, 177
homogeneous
  component (of a representation), 54–56, 60, 167, 195
  element (in a graded algebra), 8
homomorphism
  of algebras, 4
  of bialgebras, 445
  of coalgebras, 426
  of comodules, 439
  of graded algebras, 8
  of Hopf algebras, 448
  of Lie algebras, 244
  of modules, 13
  of representations, 25
hook, 199
  hook length, 199
  hook walk proof, 199
  hook-length formula, 186, 199
hook partition, 203
Hook-Length Formula, 199
Hopf algebra, 134, 272, 448
  almost cocommutative, 470, 482
  bi-semisimple, 549
  cocommutative, 135, 274, 463, 480, 482, 486, 487, 490
  involutory, 451, 457, 459, 471, 477, 544, 549–552
  unimodular, 538, 539, 544, 557
Hopf ideal, 448
Hopf kernel (of a representation of a
  Hopf algebra), 475
Hopf module, 467
  relative, 467, 471, 558
Hopf subalgebra, 448
Hopkins-Levitzki Theorem, 473

icosahedron, 175
ideal
  cofinite, 62, 434
completely prime, 308, 309, 403, 529
locally closed prime, 288, 294, 309
maximal, 44
nil, 88, 287
nilpotent, 49, 62, 287
of a Lie algebra, 246
prime, 44
primitive, 42, 285
  right vs. left, 42, 292
rational, 292, 294
semiprime, 45, 49, 285, 286, 592
idempotent, 75, 81, 95, 105, 553
primitive, 558
primitive central, 60, 103, 146, 193, 555
identity component (of an affine algebraic group), 519
identity element, xi, 4, 569
identity functor, 570
induction, 166
induction functor, 27, 39, 117, 123, 124
inflation (of a representation), 27, 48, 58, 73
inner product, 330
  Hermitian, 149, 565
integral
  in a Hopf algebra, 538
  of the representation ring, 486
integral closure, 104
integrality, 104, 160–162, 495, 555
intertwine, 25
invariant subalgebra
  of an H-module algebra, 491
invariant theory, 169, 174
  multiplicative, 307
  noncommutative, 488
invariants, 64, 396
  of a group representation, 132
of a Hopf algebra representation, 462
of a Hopf representation, 462
of a Lie algebra representation, 256, 277
of a permutation representation, 132
of an sl₂-representation, 311
polynomial, 174
inversion
  of a permutation, 206, 214
inversion
  of a permutation, 354
involution, 164, 391, 486
  standard involution of a group algebra, 135, 190
  standard involution of an enveloping algebra, 274
irreducible character, 121
irreducible component (of a topological space), 517
irreducible components
  of a root system, 356
irreducible constituent, 55, 57
irreducible representations
  of Lie algebras, 245
  of the Lie algebra so₅, 415
  of the symmetric group Sₙ, 185
irreducible topological space, 50
isomorphism, 568
  functorial, 573
  natural, 573
  of categories, 574
  of functors, 110, 573
  of graphs, 196
  of root systems, 340
isomorphism theorem
  for Lie algebras, 246
isotropy group, 141, 159
Itô’s Theorem, 160, 166, 168
Itô, N., 161
Jacobi identity, 244, 245, 249, 281, 605
Jacobson property, 285, 286, 288, 294, 298, 518, 592
Jacobson radical, 48, 49, 62, 69, 479, 492
Jacobson, N., 276
Jacobson-Morozov Lemma, 312
Jacobson-Zariski Topology, 527
Jacobson-Zariski topology, 45–50, 288, 289, 291, 297, 404
Jordan canonical form, 26, 260–261
Jordan decomposition, 261
abstract, 322, 325
preservation of, 322
Jordan-Chevalley decomposition, 261
Jordan-Hölder Theorem, 33–35, 85, 91, 296
Jucys-Murphy (JM) elements, 187
$k$-field, 5, 22, 294
Kac, G., 556
Kaplansky, I., 549, 554, 557
Kashina, Y., 561
kernel (of a representation), 24
Kharchenko, V., 293, 607
Killing form, 314, 319, 407
Klein 4-group, 153
Kolchin’s Theorem, 256
Koszul sign rule, 22
Kronecker product, 590
Krull-Schmidt Theorem, 38, 40, 85, 91, 559, 560

Lagrange’s Theorem, 559
$\lambda$-ring, 486
$\lambda$-tableau, 198
$\lambda/\mu$-tableau, 219
Larson, R. G., 537, 549, 550
lattice, 115
Launois, S., 531
Laurent polynomial algebra, 115, 364
Le-diagram, 532
Leibniz formula, 251, 393, 405
Leibniz identity, 244, 249
Leibniz product rule, 249, 279, 280
Lenagan, T., 533
length
of a $\mathfrak{sl}_2$-representation, 301
of a completely reducible representation, 53
of a permutation, 206, 214, 353
of a representation, 34
of a Weyl group element, 352
Letzter, E., 528
Levi decomposition, 254
Levitzki’s Theorem, 256
Levitzki, J., 287
lexicographic order, 600
Lie $p$-algebra, 446, 522
Lie algebra, 244
2-dimensional non-abelian, 263, 264, 276, 288, 291, 297
abelian, 244
linear, 248
nilpotent, 252, 263
of an affine algebraic group, 522
reductive, 321, 323
semisimple, 254
simple, 254
solvable, 253, 263
Lie bracket, 244
Lie commutator, 59, 67, 244, 249
Lie subalgebra, 246
Lie’s Theorem, 258–260, 262, 263, 277, 295–297, 321, 377
linear dual, 587
linear form, 587
linear recursion, 453
Liu, C.-H., 472
locally finite part, 440
longest element (of a Weyl group), 354

MacMahon Master Theorem, 171
Manin, Y., 456
marked cycle shape, 188
Markov chain, 200
Martindale, W.S., 611
Maschke's Theorem, 116, 117, 123, 143, 147–150, 163, 178, 186, 235, 479, 480, 483, 548–549
for Hopf algebras, 549, 551, 552, 554, 561, 562
relative, 148, 525
Masuoka, A., 558
matrix
   orthogonal, 217
maximal vector, 377
maximum condition, 19
   on semiprime ideals, 50
Michler's Theorem, 493
Michler, G., 493
minimal polynomial, 52
minimum condition, 41
misordering index, 604
module
   artinian, 41
   cyclic, 19, 53
   finite length, 41
   finitely generated, 17, 19
   flat, 581, 593
   free, 17, 19, 80, 469, 576, 581, 583
   graded, 21, 177
   graded free, 177
   injective, 81, 94, 107
   left, 13
   locally finite, 440
   noetherian, 19, 41
   over a PID, 37, 38
   projective, 79, 80, 107, 581
Molien’s Theorem, 174–177, 406

monoid, 116, 491, 569
monoid algebra, 116, 445, 446, 491
monoid ring, 116
   total, 116
monoidal category, 137
monomial matrix, 521
monomials, 6, 597
Morita context, 95, 545
morphism
   functorial, 572
   natural, 572
   of functors, 572
multilinear map, 584
   symmetric, 182
multiplication, 4
multiplicative group
   functor, 502
multiplicative invariant algebra, 307, 364
multiplicative invariants
   of the weights lattice for $A_n$, 369
   of the weights lattice for $C_n$, 369
multiplicity, 55, 145, 375
   of a weight, 300, 305
   of composition factor of a
   representation, 35, 91
Multiplicity-Freeness Theorem, 185, 190, 191
Murnaghan-Nakayama Rule, 186, 219, 224–228, 230, 231

Nakayama automorphism, 98–100, 542
Nakayama relations, 119
Nakayama twist, 108
natural equivalence, 573
natural transformation, 571
naturality condition, 572
Newton identities, 170, 171, 225, 417
Nichols, W., 458, 558
Nichols-Zoeller Theorem, 537, 554, 557, 559, 560
nilpotency class, 252
nilpotent
  element, 256
nilradical
  of a finite-dimensional Lie algebra, 255
Noether’s Finiteness Theorem, 493
Noether-Deuring Theorem, 29, 39
normal
element, 47, 275, 288, 296, 454
normalizer, 246
  of a Lie subalgebra, 257
Nullstellensatz, 43, 294, 298, 381, 506, 513, 517, 518, 592
  for enveloping algebras, 285–288, 290, 294, 309, 404
  weak, 36, 40, 41, 45, 285–287, 289, 294, 517, 592
Oberst, U., 543
On-Line Encyclopedia of Integer Sequences, 195
open problem, 270, 549, 554
opposite
group, 123, 135
  Lie algebra, 272
opposite category, 571
opposite Lie algebra, 272
orbit sum, 133, 187, 364
orthogonal idempotents, 88
orthogonality, 195
orthogonality relations, 92, 143–145, 149, 158, 165
column, 148
generalized, 149
orthonormal basis, 144

\textit{p}-core, 480
\textit{p}-regular
  conjugacy class, 121, 168
  element, 121
group, 484
  pairing, 86, 447
partially ordered set, 19
partition, 153
partitions, 195
Passman, D. S., 149, 168
permutation matrices, 128
PID, 219
pipe dream, 532
place permutations, 178
plethysm, 136
Poincaré-Birkhoff-Witt Theorem, 270, 604
Poincaré series, 174
Poincaré-Birkhoff-Witt Theorem,
  267–273, 276, 281, 284, 286, 291, 300, 309, 378, 398, 399, 428, 603
polarization, 182–184, 392
polynomial function, 593
  invariant, 392
polynomial functions
  on \( \text{GL}_n \), 454
  on \( \text{SL}_n \), 455
positive cone, 75
power series, 116, 433
power sum, 170
principal indecomposable representations, 91, 94, 559
principal open subset, 594
probabilistic proof, 199
probability, 200
product
  of categories, 574
projective cover, 90
projective representations, 168
projectives, 79
pseudo-reflection, 402
pseudo-reflection group, 402
pullback, 117
pulling back (a representation), 26, 48, 73
pushing forward (a representation), 27

$q$-binomial coefficient, 460
quantum $GL_n, 457$
quantum $SL_n, 457$
quantum affine $n$-space, 456, 529
quantum affine $n$-torus, 530
quantum binomial formula, 460, 544
quantum determinant, 456
quantum group, 297
quantum matrix algebra, 601
quantum plane, 23, 46, 50, 96, 298, 456, 457, 529
quantum torus, 23
quasi-homeomorphism, 298
quasi-inverse, 574
quaternion algebra, 131, 565
quaternion group, 131
Quillen, D., 286
quotient representation, 24

Rabinowitsch trick, 287, 592
Radford’s formula, 537, 545, 557
Radford, D. E., 537, 545, 549, 550
radical
Jacobson, 287
of a finite-dimensional Lie algebra, 254
of a representation, 63
of an ideal, 49, 591
of the Killing form, 314
random walk, 200
rank
of a finitely generated projective, 83
of a free module, 583
of a root system, 340
of a semisimple Lie algebra, 324, 330
rational canonical form, 26
reduced trace, 106
reduced word, 533
reduction, 598
reduction system, 598
reflection, 151, 342, 402
simple, 352
regular
elements of a Lie algebra, 394
representation
G-faithful, 475
g-faithful, 267, 282, 475
absolutely irreducible, 61, 145, 359, 373
adjoint, 249, 309, 318
of a group, 130, 133, 149, 493
of a Hopf algebra, 489, 492, 553
of a Lie algebra, 245, 247, 249, 254, 281, 392
artinian, 41
completely reducible, 51–53, 64, 141
complex, 113
faithful, 24, 64, 300
of a group, 475
of a Lie algebra, 475
finite length, 41
finite-dimensional, 24
finite-length, 33, 52
finitely generated, 52
fundamental, 385
of $sl_{n+1}, 385$
indecomposable, 37, 91, 304
induced, 123
injective, 559
inner faithful (of a Hopf algebra), 475
irreducible, 30, 245
locally finite, 301, 318, 394, 435, 440
modular, 113, 121
noetherian, 41
of a group, 113
of a Lie algebra, 245, 255
of an algebra, 24
INDEX

of finite length, 382
permutation, 130, 132, 137, 140, 141
projective, 79
rational (of an affine algebraic group), 508, 524
regular, 25, 30, 66, 130, 146
bimodule, 25, 97
right, 25
self-dual, 137, 150, 195, 465
semisimple, 51
simple, 30
trivial, 462
representation algebra, 481, 552
representation ring
of \( \mathfrak{sl}_2 \), 306
of a Hopf algebra, 481
of a Lie algebra, 284, 306, 388
representations
of \( \mathfrak{sl}_2 \), 298
representative function, 459
representative functions, 452
restitution, 182–184
restriction, 44, 117, 166
restriction functor, 27, 39, 117
Reynolds operator, 40
Rieffel’s Theorem, 476
Rieffel, M., 476
ring
of algebraic integers, 219
ring of quotients
classical, 612
left, 611
maximal, 612
right, 611
symmetric, 293, 607, 611
root
negative, 348
positive, 348
simple, 348
root lattice, 359
for the root system \( C_n \), 368
for the root system \( D_n \), 368
root space decomposition, 323, 326, 374
root spaces, 326
root string, 326
root system, 340
irreducible, 329, 354
of rank 2, 341
of type \( A_1 \), 341
of type \( A_2 \), 334
root systems, 331
roots, 326
roots of unity
functor, 502
Rowen, L., 437
Schneider, H., 543
Schneider, H.-J., 537
Schur division algebra, 36, 41, 56, 58, 124, 232, 286
Schur functor, 236, 417
Schur’s Double Centralizer Theorem, 235, 236, 238, 239, 417
Schur’s Lemma, 35, 36, 55, 56, 58, 61, 92, 117, 143, 144, 146, 162, 192, 224, 286, 319, 320, 322, 525, 553, 563–565
Schur, I., 160, 168, 185
semi-invariant, 132, 277
semidirect product, 276
of Lie algebras, 250
semigroup partial order, 598
semiprime radical, 49
of an ideal, 67
semisimple, 313
semisimplification, 63, 69
Serre, J.-P., 477
Shephard-Todd-Chevalley Theorem, 402, 406
short exact sequence, 18, 19, 33, 41,
66, 81, 124, 443, 487, 581
split, 18
sign representation, 128
of Weyl groups, 366
simple, 313
skeleton, 575
skew hook, 224
skew Laurent polynomial algebra,
20, 23
skew polynomial algebra, 19, 276
skew shape, 219
sl2-triples, 186, 210, 298, 326, 334
smash product, 284, 491
snake relations, 431
socle, 54, 108, 375
socle series, 110
Sommerhäuser, Y., 561
special linear group, 160
functor, 503
special linear Lie algebra, 247, 248
splitting field, 36, 37, 59, 61, 65, 70,
74, 77, 92, 93, 117, 124, 131,
144, 148, 155, 216, 493, 553
splitting map, 18
splitting principle, 484
stable isomorphism, 85, 96
Stallings, J., 83
standard filtration, 269
standard monomial, 11, 12, 172,
271, 428, 433, 457, 458
standard monomials, 271
standard permutation representation
of $S_n$, 128, 130, 150, 214
standard representation
of $S_3$, 147
of $S_n$, 128, 192, 195, 213, 214,
216, 218, 219
of the Weyl algebra, 26, 263
standard representation of $S_n$, 131,
150
standard Young tableau
of shape $\lambda$, 198, 219
Stirling numbers, 532
Structure Theorem for Hopf
Modules, 467–470, 472, 474
subalgebra, 4
cofinite, 40
subbialgebra, 445
subcategory, 569
full, 569
subcoalggebra, 426
subcomodule, 439
subgroup scheme, 502
subrepresentation, 24, 245
Sweedler algebra, 96, 109, 458
Sweedler dual, 434
Sweedler Hopf algebra, 470, 496
Sweedler notation, 429
Sweedler power, 487
Sweedler powers, 561
Sweedler, M., 458
Sweedler-Heyneman notation, 429
symbol, 268, 271, 309
symmetric polynomial, 169
complete, 169
elementary, 169, 390
symmetrization, 179
symmetrization map, 282
symmetrizer, 179, 480
symmetry group, 127, 129, 175
Taft Hopf algebra, 538, 539, 544
Taft Hopf algebras, 457, 481
Taft, E., 457
Tate, J., 161
tensor
algebra, 6–9, 138, 279, 490
antisymmetric, 178, 480
power, 8, 138, 279, 490, 583, 584
product, 570, 579
product formula (for group
representations), 142
symmetric, 178, 480
tensor category, 137, 278

tensor power, 6

tensor product
  of algebras, 5, 15
  outer, 67, 77, 138, 140, 227

topological space
  connected, 516
  irreducible, 516

total degree, 8, 9

totally non-negative, 549
trace, 26, 65
trace form, 68, 120, 122, 539, 543
trace map, 247, 471
  relative, 148
transporter, 246
transpose of a linear map, 139, 571, 587
triangular decomposition, 373, 374
trigonometric Hopf algebra, 460
trivial representation, 117, 130, 277, 462
  of the representation algebra, 556
twist
  of a representation, 28, 39, 142, 166

unipotent, 256
unit, 4
  of an adjunction, 265, 275
unit axioms, 568
unit laws, 4
universal property, 18, 114
  of enveloping algebras, 264
  of the free algebra, 9
  of the polynomial algebra, 9
  of the tensor algebra, 7
  of the tensor product, 580
universal trace, 68, 76, 100

Vandermonde determinant, 410
vector partition function, 378
vector space
  graded, 21

Verma module, 379, 413

walls, 351
Wedderburn’s Structure Theorem, 3, 30, 51, 57, 59, 60, 62, 63, 91, 102, 103, 146, 153, 154, 233, 314, 549, 556
weight, 54, 195, 305
  dominant, 362
  for the root system $B_n$, 368, 415
  for the root system $C_n$, 368
  for the root system $D_n$, 368
  fundamental, 360
  strongly dominant, 362, 366
weight diagram, 386
weight lattice, 360, 363, 372, 376
weight space, 54, 132, 142, 147, 258, 296, 511
  for $sl_2$-modules, 300
  generalized, 301
weight spaces, 375
weight vector, 132, 194, 258, 277, 375
weights, 194, 208, 214, 375
  for $sl_2$-representations, 300
Weyl algebra, 12–13, 23, 26, 31, 41, 203, 250, 260, 291, 312, 597, 600
  standard representation of, 26, 30, 263
Weyl chamber, 351
  fundamental, 352, 362
Weyl group, 343, 373, 396
  of $A_n$, 344
  of $B_n$, 346
  of $C_n$, 346
  of $D_n$, 347
Weyl’s Theorem, 318, 321, 322, 371, 378, 389, 403, 416, 525
winding automorphism, 142, 399, 471
Witt algebra, 251, 270
words, 6, 597
Yakimov, M., 531
Young diagram, 185, 195
Young graph, 185, 195, 196
Young module, 180
Young subgroup, 152, 180, 227, 228, 236, 420
Young’s orthogonal form, 218, 222
Young’s orthogonal form, 186
Young, A., 152, 185
Zariski dense, 375, 594
Zariski topology, 45, 181, 375, 397, 400, 512, 517, 593–595
Zariski’s Lemma, 592
Zhang, J., 472
Zhu, Y., 556, 557, 561
Zoeller, M. B., 558
Zorn’s Lemma, 19, 53, 561