An Introduction to the Representation Theory of Groups

Emmanuel Kowalski

Graduate Studies in Mathematics
Volume 155

American Mathematical Society

An Introduction to the Representation Theory of Groups

Emmanuel Kowalski

Graduate Studies
in Mathematics
Volume 155

American Mathematical Society
Providence, Rhode Island

2010 *Mathematics Subject Classification.* Primary 20-01, 20Cxx, 22A25.

For additional information and updates on this book, visit
www.ams.org/bookpages/gsm-155

Library of Congress Cataloging-in-Publication Data

Kowalski, Emmanuel, 1969–
An introduction to the representation theory of groups / Emmanuel Kowalski.
pages cm. — (Graduate studies in mathematics ; volume 155)
Includes bibliographical references and index.
ISBN 978-1-4704-0966-1 (alk. paper)
1. Lie groups. 2. Representations of groups. 3. Group algebras. I. Title.

QA387.K69 2014
515′.7223—dc23

2014012974

---

10 9 8 7 6 5 4 3 2 1 19 18 17 16 15 14

An Introduction to the Representation Theory of Groups

Contents

Chapter 1

Introduction and motivation

This book is intended to provide a basic introduction to some of the fundamental ideas and results of *representation theory*. In this preliminary chapter, we start with some motivating remarks and provide a general overview of the rest of the text; we also include some notes on the prerequisites—which are not uniform for all parts of the notes—and discuss the basic notation that we use.

In writing this text, the objective has never been to give the shortest or slickest proof. To the extent that the author's knowledge makes this possible, the goal is rather to explain the ideas and the mechanism of thought that can lead to an understanding of "why" something is true, and not simply to the quickest line-by-line check that it holds.

The point of view is that representation theory is a fundamental theory, both for its own sake and as a tool in many other fields of mathematics; the more one knows, understands, and breathes representation theory, the better. This style (or its most ideal form) is perhaps best summarized by P. Sarnak's advice in the *Princeton Companion to Mathematics* [**24**, p. 1008]:

> One of the troubles with recent accounts of certain topics is that they can become too slick. As each new author finds cleverer proofs or treatments of a theory, the treatment evolves toward the one that contains the "shortest proofs." Unfortunately, these are often in a form that causes the new student to ponder, "How did anyone think of this?" By going back to the original sources one can

> usually see the subject evolving naturally and understand how it has reached its modern form. (There will remain those unexpected and brilliant steps at which one can only marvel at the genius of the inventor, but there are far fewer of these than you might think.) As an example, I usually recommend reading Weyl's original papers on the representation theory of compact Lie groups and the derivation of his character formula, alongside one of the many modern treatments.

So the text sometimes gives two proofs of the same result, even in cases where the arguments are fairly closely related; one may be easy to motivate ("how would one try to prove such a thing?"), while the other may recover the result by a slicker exploitation of the formalism of representation theory. To give an example, we first consider Burnside's irreducibility criterion, and its developments, using an argument roughly similar to the original one, before showing how Frobenius reciprocity leads to a quicker line of reasoning (see Sections 2.7.3 and 2.7.4).

Finally, although I have tried to illustrate many aspects of representation theory, there remains many topics that are barely mentioned or omitted altogether. Maybe the most important are:

- The representation theory of anything else than groups; in particular, Lie algebras and their representations only make passing appearances, and correspondingly those aspects of representation theory that really depend on these techniques are not developed in any detail. Here, the book [**20**] by Fulton and Harris is an outstanding resource, and the book [**18**] by Etingof, Golberg, Hensel, Liu, Schwendner, Vaintrob, and Yudovina illustrates different aspects, such as the representations of quivers.
- In a related direction, since it really depends on Lie algebraic methods, the precise classification of representations of compact Lie groups, through the theory of highest weight representations, is not considered beyond the case of $\mathrm{SU}_2(\mathbf{C})$; this is however covered in great detail in many other texts, such as [**20**] again, the book [**37**] of Knapp (especially Chapter V), or the book [**35**] of Kirillov.

Acknowledgments. The notes were prepared in parallel with the course "Representation Theory" that I taught at ETH Zürich during the Spring Semester 2011. Thanks are obviously due to all the students who attended the course for their remarks and interest, in particular M. Lüthy, M Rüst, I. Schwabacher, M. Scheuss, and M. Tornier, and to the assistants in charge of the exercise sessions, in particular J. Ditchen who coordinated

those. Thanks also to "Anonymous Rex" for a comment on a blog post, to U. Schapira for his comments and questions during the class, and to A. Venkatesh for showing me his own notes for a (more advanced) representation theory class, from which I derived much insight.

Thanks to the reviewers for the original book proposal for suggestions and comments—in particular for some well-deserved critical comments concerning certain of the choices of notation in the first version of the text, and for pointing out that Proposition 2.3.23 is false over non-algebraically closed fields.

Finally, many thanks to E. Dunne for reading the whole manuscript carefully and making many suggestions and corrections!

1.1. Presentation

A (linear) representation of a group G is, to begin with, simply a *homomorphism*

$$\varrho : G \longrightarrow \mathrm{GL}(E),$$

where E is a vector space over some field k and $\mathrm{GL}(E)$ is the group of invertible k-linear maps on E. Thus one can guess that this should be a useful notion by noting how it involves the simplest and most ubiquitous algebraic structure, that of a group, with the powerful and flexible tools of linear algebra. Or, in other words, such a map attempts to "represent" the elements of G as symmetries of the vector space E (note that ϱ might fail to be injective, so that G is not mapped to an isomorphic group).

But even a first guess would probably not lead one to imagine how widespread and influential the concepts of representation theory turn out to be in current mathematics. Few fields of mathematics, or of mathematical physics (or chemistry), do not make use of these ideas, and many depend on representations in an essential way. We will try to illustrate this wide influence with examples, taken in particular from number theory and from basic quantum mechanics; already in Section 1.2 below we state four results, where representation theory does not appear in the statements although it is a fundamental tool in the proofs. Moreover, it should be said that representation theory is now a field of mathematics in its own right, which can be pursued without having immediate applications in mind; it does not require external influences to expand with new questions, results and concepts—but we will barely scratch such aspects.

The next chapter starts by presenting the fundamental vocabulary that is the foundation of representation theory and by illustrating it with examples. In Chapter 3, we then present a number of short sections concerning variants of the definition of representations: restrictions can be imposed on the group

G, on the type of fields or vector spaces E allowed, or additional regularity assumptions may be imposed on ϱ when this makes sense. One can also replace groups by other objects: we will mention associative algebras and Lie algebras. These variants are all important topics in their own right, but some will only reappear briefly in the rest of the book.

Continuing, Chapter 4 is an introduction to the simplest case of representation theory: the linear representations of finite groups in finite-dimensional complex vector spaces. This is also historically the first case that was studied in depth by Dirichlet (for finite abelian groups), then Frobenius, Schur, Burnside, and many others. It is a beautiful theory and has many important applications. It can also serve as a "blueprint" to many generalizations: various facts, which are extremely elementary for finite groups, remain valid, when properly framed, for important classes of infinite groups.

Among these, the compact topological groups are undoubtedly those closest to finite groups, and we consider them in Chapter 5. Then Chapter 6 presents some concrete examples of applications involving compact Lie groups (compact matrix groups, such as unitary groups $\mathrm{U}_n(\mathbf{C})$)—the most important being perhaps the way representation theory explains a lot about the way the most basic atom, hydrogen, behaves in the real world....

The final Chapter 7 has again a survey flavor, and it is intended to serve as an introduction to two other important classes of groups: algebraic groups, on the one hand, and non-compact locally compact groups, on the other hand. This last case is illustrated through the fundamental example of the group $\mathrm{SL}_2(\mathbf{R})$ of two-by-two real matrices with determinant 1. We use it primarily to illustrate some of the striking new phenomena that arise when compactness is missing.

In Appendix A, we have gathered statements and sketches of proofs for certain facts, especially the Spectral Theorem for compact self-adjoint linear operators, which are needed for rigorous treatments of unitary representations of topological groups.

Throughout, we also present some examples by means of exercises. These are usually not particularly difficult, but we hope they will help the reader to get acquainted with the way of thinking that representation theory often suggests for certain problems.

1.2. Four motivating statements

Below are four results, taken in very different fields, which we will discuss again later (or sometimes only sketch when very different ideas are also needed). The statements do not mention representation theory, in fact two of them do not even mention groups explicitly. Yet they are proved using

these tools, and they serve as striking illustrations of what can be done using representation theory.

Example 1.2.1 (Primes in arithmetic progressions)**.** Historically, the first triumph of representation theory is the proof by Dirichlet of the existence of infinitely many prime numbers in an arithmetic progression, whenever this is not clearly impossible:

Theorem 1.2.2 (Dirichlet)**.** *Let $q \geqslant 1$ be an integer, and let $a \geqslant 1$ be an integer coprime with q. Then there exist infinitely many prime numbers p such that*

$$p \equiv a \pmod{q},$$

i.e., such that p is of the form $p = nq + a$ for some $n \geqslant 1$.

For instance, taking $q = 10^k$ to be a power of 10, we can say that, for whichever ending pattern of digits $\boldsymbol{d} = d_{k-1}d_{k-2}\cdots d_0$ we might choose, with $d_i \in \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$, *provided* the last digit d_0 is not one of $\{0, 2, 4, 5, 6, 8\}$, there exist infinitely many prime numbers p with a decimal expansion where $\boldsymbol{d}$ are the final digits. To illustrate this, taking $q = 1000$, $\boldsymbol{d} = 237$, we find

$$1237,\ 2237,\ 5237,\ 7237,\ 8237,\ 19237,\ 25237,\ 26237,\ 31237,\ 32237,$$
$$38237,\ 40237,\ 43237,\ 46237,\ 47237,\ 52237,\ 56237,\ 58237,\ 64237,$$
$$70237,\ 71237,\ 73237,\ 77237,\ 82237,\ 85237,\ 88237,\ 89237,\ 91237,\ 92237$$

to be those prime numbers ending with 237 which are $\leqslant 100000$.

We will present the idea of the proof of this theorem in Chapter 4. As we will see, a crucial ingredient (but not the only one) is the simplest type of representation theory: that of groups that are both finite and commutative. In some sense, there is no better example to guess the power of representation theory than to see how even the simplest instance leads to such remarkable results.

Example 1.2.3 (The hydrogen atom)**.** According to current knowledge, about 75% of the observable weight of the universe is accounted for by hydrogen atoms. In quantum mechanics, the possible states of an (isolated) hydrogen atom are described in terms of combinations of "pure" states, and the latter are determined by *discrete* data, traditionally called "quantum numbers"—so that the possible energy values of the system, for instance, form a discrete set of numbers, rather than a continuous interval.

Precisely, in non-relativistic theory, there are four quantum numbers for a given pure state of hydrogen, denoted (n, ℓ, m, s)—"*principal*, *angular momentum*, *magnetic*, and *spin*" are their usual names—which are all integers,

except for s, with the restrictions

$$n \geqslant 1, \quad 0 \leqslant \ell \leqslant n-1, \quad -\ell \leqslant m \leqslant \ell, \quad s \in \{-1/2, 1/2\}.$$

It is rather striking that much of this quantum mechanical model of the hydrogen atom can be "explained" qualitatively by an analysis of the representation theory of the underlying symmetry group (see [**64**] or [**58**]) leading in particular to a natural explanation of the intricate structure of these four quantum numbers! We will attempt to explain the easiest part of this story, which only involves the magnetic and angular momentum quantum numbers, in Section 6.4.

Example 1.2.4 ("Word" problems). For a prime number p, consider the finite group $\mathrm{SL}_2(\mathbf{F}_p)$ of square matrices of size 2 with determinant 1, and with coefficients in the finite field $\mathbf{F}_p = \mathbf{Z}/p\mathbf{Z}$. This group is generated by the two elements

$$s_1 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \qquad s_2 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \tag{1.1}$$

(this is a fairly easy fact from elementary group theory, see, e.g., [**51**, Th. 8.8] for $K = \mathbf{F}_p$ or Exercise 4.6.20). Certainly the group is also generated by the elements of the set $S = \{s_1, s_1^{-1}, s_2, s_2^{-1}\}$, and in particular, for any $g \in \mathrm{SL}_2(\mathbf{F}_p)$, there exist an integer $k \geqslant 1$ and elements $g_1, \ldots, g_k$, each of which belongs to S, such that

$$g = g_1 \cdots g_k.$$

Given g, let $\ell(g)$ be the *smallest* k for which such a representation exists. One may ask, how large can $\ell(g)$ be when g varies over $\mathrm{SL}_2(\mathbf{F}_p)$? The following result gives an answer:

Theorem 1.2.5 (Selberg, Brooks, Burger). *There exists a constant $C \geqslant 0$, independent of p, such that, with notation as above, we have*

$$\ell(g) \leqslant C \log p$$

for all $g \in \mathrm{SL}_2(\mathbf{F}_p)$.

All proofs of this result depend crucially on ideas of representation theory, among other tools. And while it may seem to be rather simple and not particularly worth notice, the following *open* question should suggest that there is something very subtle here.

Problem. *Find an* efficient algorithm *that, given p and $g \in \mathrm{SL}_2(\mathbf{F}_p)$, explicitly gives $k \leqslant C \log p$ and a sequence $(g_1, \ldots, g_k)$ in S such that*

$$g = g_1 \cdots g_k.$$

For instance, what would you do with

$$g = \begin{pmatrix} 1 & (p-1)/2 \\ 0 & 1 \end{pmatrix}$$

(for $p \geqslant 3$)? Of course, one can take $k = (p-1)/2$ and $g_i = s_1$ for all i, but when p is large, this is much larger than what the theorem claims is possible!

We will not prove Theorem 1.2.5, nor really say much more about the known proofs. However, in Section 4.7.1, we present more elementary results of Gowers [**23**] (and Nikolov and Pyber [**47**]) which are much in the same spirit, and we use the same crucial ingredient concerning representations of $\mathrm{SL}_2(\mathbf{F}_p)$. The book [**13**] of Davidoff, Sarnak, and Valette gives a complete elementary proof and is fully accessible to readers of this book.

In these three first examples, it turns out that representation theory appears in a similar manner: it is used to analyze functions on a group, in a way which is close to the theory of Fourier series or Fourier integrals; indeed, both of these can also be understood in terms of representation theory for the groups $\mathbf{R}/\mathbf{Z}$ and $\mathbf{R}$, respectively (see Section 7.3). The next motivating example is purely algebraic.

Example 1.2.6 (Burnside's $p^a q^b$ theorem)**.** Recall that a group G is called *solvable* if there is an increasing sequence of subgroups

$$1 \lhd G_k \lhd G_{k-1} \lhd \cdots \lhd G_1 \lhd G = G_0,$$

each normal in the next (but not necessarily in G), such that each successive quotient G_k/G_{k+1} is an abelian group.

Theorem 1.2.7 (Burnside)**.** *Let G be a finite group. If the order of G is divisible by at most two distinct prime numbers, then G is solvable.*

This beautiful result is sharp in some sense: it is well known that the symmetric group $\mathfrak{S}_5$ of order $5! = 120$ is *not* solvable, and since 120 is divisible only by the primes 2, 3 and 5, we see that the analogue statement with 2 prime factors replaced with 3 is not true. (Also it is clear that the converse is not true either: any abelian group is solvable, and there are such groups of any order.)

This theorem of Burnside will be proved using representation theory of finite groups in Section 4.7.2 of Chapter 4, in much the same way as Burnside proceeded in the early 20th century. It was only in the late 1960s that a proof not using representation theory was found, first by Goldschmidt when the primes p and q are odd, and then independently by Bender and Matsuyama for the general case. There is a full account of this in [**29**, §7D], and although it is not altogether overwhelming in length, the reader who

compares them will probably agree that the proof based on representation theory is significantly easier to digest.

Remark 1.2.8. There are even more striking results which are much more difficult. For instance, the famous "Odd-order Theorem" of Feit and Thompson states that if G has *odd* order, then G is necessarily solvable.

1.3. Prerequisites and notation

In Chapters 2 and 4, we depend only on the content of a basic graduate course in algebra: basic group theory, abstract linear algebra over fields, polynomial rings, finite fields, modules over rings, bilinear forms, and the tensor product and its variants. In later chapters, other structures are involved: groups are considered with a topology, measure spaces and integration theory is involved, as well as basic Hilbert space theory and functional analysis. All these are used at the level of introductory graduate courses.

We will use the following notation:

(1) For a set X, $|X| \in [0, +\infty]$ denotes its cardinality, with $|X| = \infty$ if X is infinite. There is no distinction in this text between the various infinite cardinals.

(2) We denote by $\mathbf{R}^{+,\times}$ the interval $]0, +\infty[$ seen as a subgroup of the multiplicative group $\mathbf{R}^{\times}$.

(3) If k is a field and $d \geqslant 1$ an integer, an element of $\mathrm{GL}_d(k)$ (or of $\mathrm{GL}(E)$ where E is a finite-dimensional k-vector space) is called *unipotent* if there exists $n \geqslant 1$ such that $(u - \mathrm{Id}_k)^n = 0$.

(4) Given a ring A, with a unit $1 \in A$, and A-modules M and N, we denote by $\mathrm{Hom}(M, N)$ or $\mathrm{Hom}_A(M, N)$ the space of A-linear maps from M to N.

(5) If E is a vector space over a field k, E' denotes the dual space $\mathrm{Hom}_k(E, k)$. We often use the duality bracket notation for evaluating linear maps on vectors, i.e., for $v \in E$ and $\lambda \in E'$, we write

$$\langle \lambda, v \rangle = \lambda(v).$$

(6) For $f : M \to N$, a map of A-modules, $\mathrm{Ker}(f)$ and $\mathrm{Im}(f)$ denote the kernel and the image of f, respectively.

(7) A *projection* $f : M \longrightarrow M$ is a linear map such that $f \circ f = f$. If f is such a projection, we have $M = \mathrm{Im}(f) \oplus \mathrm{Ker}(f)$; we also say that f is the projection on $\mathrm{Im}(f)$ with kernel $\mathrm{Ker}(f)$.

(8) Given A and M, N as above, $M \otimes N$ or $M \otimes_A N$ denotes the tensor product of M and N. Recall that $M \otimes N$ can be characterized up to

isomorphism by the existence of canonical isomorphisms

$$\mathrm{Hom}_A(M \otimes N, N_1) \simeq \mathrm{Bil}(M \times N, N_1)$$

for any A-module N_1, where the right-hand side is the A-module of all A-bilinear maps

$$\beta : M \times N \to N_1.$$

In particular, there is a bilinear map

$$\beta_0 : M \times N \longrightarrow M \otimes N$$

that corresponds to $N_1 = M \otimes N$ and to the identity map in $\mathrm{Hom}_A(M \otimes N, N_1)$. One writes $v \otimes w$ instead of $\beta_0(v, w)$.

The elements of the type $v \otimes w$ in $M \otimes N$ are called *pure tensors*. Note that usually not all elements in the tensor product are pure tensors and that one can have $v \otimes w = v' \otimes w'$ even if $(v, w) \neq (v', w')$.

If $A = k$ is a field and (e_i), (f_j) are bases of the k-vector spaces M and N, respectively, then $(e_i \otimes f_j)$ is a basis of $M \otimes N$. Moreover, any $v \in M \otimes N$ has a *unique* expression

$$v = \sum_j v_j \otimes f_j$$

with $v_j \in M$ for all j.

(9) Given a ring A and A-modules given with linear maps

$$M' \xrightarrow{f} M \xrightarrow{g} M',$$

the sequence is said to be *exact* if $\mathrm{Im}(f) = \mathrm{Ker}(g)$ in M. In particular, a sequence

$$0 \longrightarrow M' \xrightarrow{f} M$$

is exact if and only if $\mathrm{Ker}(f) = 0$, which means that f is injective, and a sequence

$$M \xrightarrow{g} M'' \longrightarrow 0$$

is exact if and only if $\mathrm{Im}(g) = \mathrm{Ker}(0) = M''$, i.e., if and only if g is surjective.

A sequence

$$0 \longrightarrow M' \xrightarrow{f} M \xrightarrow{g} M'' \longrightarrow 0,$$

where all three intermediate 3-term sequences are exact, is called a *short exact sequence*. This means that f is injective, g is surjective and the image of f coincides with the kernel of g. It is also usual to say that M is an *extension* of M'' by M'. Note that there is no typo here: this is indeed the standard terminology, instead of speaking of extensions of M'.

(10) Given a vector space E over a field k and a family $(F_i)_{i\in I}$ of linear subspaces of E, we say that the subspaces F_i *are in direct sum* if the subspace they span is a direct sum of the F_i, or in other words, if

$$F_i \cap \Big(\sum_{\substack{j\in I\\ j\neq i}} F_j\Big) = 0$$

for all $i \in I$ (equivalently, any family $(f_i)_{i\in I}$ of vectors in F_i, which are zero for all but finitely many indices i, is linearly independent).

(11) Given a group G, we denote by $[G, G]$ the *commutator group* (or *derived subgroup*) of G, which is generated by all commutators $[g, h] = ghg^{-1}h^{-1}$. Note that not all elements of $[G, G]$ are themselves commutators; see Remark 4.4.5 for examples. The subgroup $[G, G]$ is normal in G, and the quotient group $G/[G, G]$ is abelian; it is called the *abelianization* of G.

(12) We denote by $\mathbf{F}_p$ the finite field $\mathbf{Z}/p\mathbf{Z}$, for p prime and, more generally, by $\mathbf{F}_q$ a finite field with q elements, where $q = p^n$, $n \geqslant 1$, is a power of p. In Chapter 4, we need some simple facts about these, in particular the fact that for each $n \geqslant 1$, there is—up to isomorphism—a unique extension $k/\mathbf{F}_p$ of degree n, i.e., a finite field k of order $q = p^n$. An element $x \in k$ is in $\mathbf{F}_p$ if and only if $x^p = x$ (e.g., because the equation $X^p - X = 0$ has at most p roots, and all $x \in \mathbf{F}_p$ are roots). The group homomorphism

$$N = N_{k/\mathbf{F}_p} : \begin{cases} k^\times & \longrightarrow & \mathbf{F}_p^\times \\ x & \mapsto & \prod_{j=0}^{n-1} x^{p^j} \end{cases}$$

(called the *norm* from k to $\mathbf{F}_p$) is well defined and surjective. Indeed, it is well defined because one checks that $N(x)^p = N(x)$, and surjective, e.g., because the kernel is defined by a non-zero polynomial equation of degree at most $1 + p + p + \cdots + p^{n-1} = (p^n - 1)/(p - 1)$, and hence contains at most that many elements, so the image has at least $p - 1$ elements. Moreover, the kernel of the norm is the set of all x which can be written as y/y^p for some $y \in k^\times$.

Similarly, the homomorphism of abelian groups

$$\mathrm{Tr} = \mathrm{Tr}_{k/\mathbf{F}_p} : \begin{cases} \mathbf{F}_q & \longrightarrow & \mathbf{F}_p \\ x & \mapsto & x + x^p + \cdots + x^{p^{n-1}} \end{cases}$$

is well defined and is surjective; it is called the *trace* from k to $\mathbf{F}_p$.

(13) When considering a normed vector space E, we usually denote the norm by $\|v\|$, and sometimes write $\|v\|_E$, when more than one space (or norm) are considered simultaneously.

(14) When considering a Hilbert space H, we speak synonymously of an inner product or of a positive-definite hermitian form, which we denote $\langle \cdot, \cdot \rangle$ or $\langle \cdot, \cdot \rangle_{\mathrm{H}}$ if more than one space might be understood. We use the convention