Graduate Texts in Mathematics 129

Readings in Mathematics

SPRINGER STUDY EDITION

Graduate Texts in Mathematics
Readings in Mathematics

Ebbinghaus/Hermes/Hirzebruch/Koecher/Mainzer/Neukirch/Prestel/Remmert: *Numbers*
Fulton/Harris: *Representation Theory: A First Course*
Murty: *Problems in Analytic Number Theory*
Remmert: *Theory of Complex Functions*
Walter: *Ordinary Differential Equations*

Undergraduate Texts in Mathematics
Readings in Mathematics

Anglin: *Mathematics: A Concise History and Philosophy*
Anglin/Lambek: *The Heritage of Thales*
Bressoud: *Second Year Calculus*
Hairer/Wanner: *Analysis by Its History*
Hämmerlin/Hoffmann: *Numerical Mathematics*
Isaac: *The Pleasures of Probability*
Laubenbacher/Pengelley: *Mathematical Expeditions: Chronicles by the Explorers*
Samuel: *Projective Geometry*
Stillwell: *Numbers and Geometry*
Toth: *Glimpses of Algebra and Geometry*

William Fulton Joe Harris

Representation Theory

A First Course

With 144 Illustrations

Springer

William Fulton
Department of Mathematics
University of Michigan
Ann Arbor, MI 48109
USA
wfulton@math.lsa.umich.edu

Joe Harris
Department of Mathematics
Harvard University
Cambridge, MA 02138
USA
harris@abel.math.harvard.edu

Mathematics Subject Classification (2000): 20G05, 17B10, 17B20, 22E46

Library of Congress Cataloging-in-Publication Data
Fulton, William, 1939–
Representation theory: a first course / William Fulton and Joe Harris.
p. cm. — (Graduate texts in mathematics)
Includes bibliographical references and index.
1. Representations of groups. 2. Representations of Algebras.
3. Lie Groups. 4. Lie algebras. I. Harris, Joe. II. Title.
III. Series.
QA171.F85 1991
512′.2—dc20 90-24926

ISBN 978-0-387-97495-8 ISBN 978-1-4612-0979-9 (eBook)
DOI 10.1007/978-1-4612-0979-9 Printed on acid-free paper.

9

springeronline.com

Preface

The primary goal of these lectures is to introduce a beginner to the finite-dimensional representations of Lie groups and Lie algebras. Since this goal is shared by quite a few other books, we should explain in this Preface how our approach differs, although the potential reader can probably see this better by a quick browse through the book.

Representation theory is simple to define: it is the study of the ways in which a given group may act on vector spaces. It is almost certainly unique, however, among such clearly delineated subjects, in the breadth of its interest to mathematicians. This is not surprising: group actions are ubiquitous in 20th century mathematics, and where the object on which a group acts is not a vector space, we have learned to replace it by one that is (e.g., a cohomology group, tangent space, etc.). As a consequence, many mathematicians other than specialists in the field (or even those who think they might want to be) come in contact with the subject in various ways. It is for such people that this text is designed. To put it another way, we intend this as a book for beginners to learn from and not as a reference.

This idea essentially determines the choice of material covered here. As simple as is the definition of representation theory given above, it fragments considerably when we try to get more specific. For a start, what kind of group G are we dealing with—a finite group like the symmetric group $\mathfrak{S}_n$ or the general linear group over a finite field $\mathrm{GL}_n(\mathbb{F}_q)$, an infinite discrete group like $\mathrm{SL}_n(\mathbb{Z})$, a Lie group like $\mathrm{SL}_n\mathbb{C}$, or possibly a Lie group over a local field? Needless to say, each of these settings requires a substantially different approach to its representation theory. Likewise, what sort of vector space is G acting on: is it over $\mathbb{C}$, $\mathbb{R}$, $\mathbb{Q}$, or possibly a field of positive characteristic? Is it finite dimensional or infinite dimensional, and if the latter, what additional structure (such as norm, or inner product) does it carry? Various combinations of answers to these questions lead to areas of intense research activity in representation theory, and it is natural for a text intended to prepare students for a career in the subject to lead up to one or more of these areas. As a corollary, such a book tends to get through the elementary material as quickly as possible: if one has a semester to get up to and through Harish–Chandra modules, there is little time to dawdle over the representations of $\mathfrak{S}_4$ and $SL_3\mathbb{C}$.

By contrast, the present book focuses exactly on the simplest cases: representations of finite groups and Lie groups on finite-dimensional real and complex vector spaces. This is in some sense the common ground of the subject, the area that is the object of most of the interest in representation theory coming from outside.

The intent of this book to serve nonspecialists likewise dictates to some degree our approach to the material we do cover. Probably the main feature of our presentation is that we concentrate on examples, developing the general theory sparingly, and then mainly as a useful and unifying language to describe phenomena already encountered in concrete cases. By the same token, we for the most part introduce theoretical notions when and where they are useful for analyzing concrete situations, postponing as long as possible those notions that are used mainly for proving general theorems.

Finally, our goal of making the book accessible to outsiders accounts in part for the style of the writing. These lectures have grown from courses of the second author in 1984 and 1987, and we have attempted to keep the informal style of these lectures. Thus there is almost no attempt at efficiency: where it seems to make sense from a didactic point of view, we work out many special cases of an idea by hand before proving the general case; and we cheerfully give several proofs of one fact if we think they are illuminating. Similarly, while it is common to develop the whole semisimple story from one point of view, say that of compact groups, or Lie algebras, or algebraic groups, we have avoided this, as efficient as it may be.

It is of course not a strikingly original notion that beginners can best learn about a subject by working through examples, with general machinery only introduced slowly and as the need arises, but it seems particularly appropriate here. In most subjects such an approach means one has a few out of an unknown infinity of examples which are useful to illuminate the general situation. When the subject is the representation theory of complex semisimple Lie groups and algebras, however, something special happens: once one has worked through all the examples readily at hand—the "classical" cases of the special linear, orthogonal, and symplectic groups—one has not just a few useful examples, one has all but five "exceptional" cases.

This is essentially what we do here. We start with a quick tour through representation theory of finite groups, with emphasis determined by what is useful for Lie groups. In this regard, we include more on the symmetric groups than is usual. Then we turn to Lie groups and Lie algebras. After some preliminaries and a look at low-dimensional examples, and one lecture with some general notions about semisimplicity, we get to the heart of the course: working out the finite-dimensional representations of the classical groups.

For each series of classical Lie algebras we prove the fundamental existence theorem for representations of given highest weight by explicit construction. Our object, however, is not just existence, but to see the representations in action, to see geometric implications of decompositions of naturally occurring representations, and to see the relations among them caused by coincidences between the Lie algebras.

The goal of the last six lectures is to make a bridge between the example-oriented approach of the earlier parts and the general theory. Here we make an attempt to interpret what has gone before in abstract terms, trying to make connections with modern terminology. We develop the general theory enough to see that we have studied all the simple complex Lie algebras with five exceptions. Since these are encountered less frequently than the classical series, it is probably not reasonable in a first course to work out their representations as explicitly, although we do carry this out for one of them. We also prove the general Weyl character formula, which can be used to verify and extend many of the results we worked out by hand earlier in the book.

Of course, the point we reach hardly touches the current state of affairs in Lie theory, but we hope it is enough to keep the reader's eyes from glazing over when confronted with a lecture that begins: "Let G be a semisimple Lie group, P a parabolic subgroup, ..." We might also hope that working through this book would prepare some readers to appreciate the elegance (and efficiency) of the abstract approach.

In spirit this book is probably closer to Weyl's classic [We1] than to others written today. Indeed, a secondary goal of our book is to present many of the results of Weyl and his predecessors in a form more accessible to modern readers. In particular, we include Weyl's constructions of the representations of the general and special linear groups by using Young's symmetrizers; and we invoke a little invariant theory to do the corresponding result for the orthogonal and symplectic groups. We also include Weyl's formulas for the characters of these representations in terms of the elementary characters of symmetric powers of the standard representations. (Interestingly, Weyl only gave the corresponding formulas in terms of the exterior powers for the general linear group. The corresponding formulas for the orthogonal and symplectic groups were only given recently by D'Hoker, and by Koike and Terada. We include a simple new proof of these determinantal formulas.)

More about individual sections can be found in the introductions to other parts of the book.

Needless to say, a price is paid for the inefficiency and restricted focus of these notes. The most obvious is a lot of omitted material: for example, we include little on the basic topological, differentiable, or analytic properties of Lie groups, as this plays a small role in our story and is well covered in dozens of other sources, including many graduate texts on manifolds. Moreover, there are no infinite-dimensional representations, no Harish-Chandra or Verma modules, no Stiefel diagrams, no Lie algebra cohomology, no analysis on symmetric spaces or groups, no arithmetic groups or automorphic forms, and nothing about representations in characteristic $p > 0$. There is no consistent attempt to indicate which of our results on Lie groups apply more generally to algebraic groups over fields other than $\mathbb{R}$ or $\mathbb{C}$ (e.g., local fields). And there is only passing mention of other standard topics, such as universal enveloping algebras or Bruhat decompositions, which have become standard tools of representation theory. (Experts who saw drafts of this book agreed that some topic we omitted must not be left out of a modern book on representation theory—but no two experts suggested the same topic.)

We have not tried to trace the history of the subjects treated, or assign credit, or to attribute ideas to original sources—this is far beyond our knowledge. When we give references, we have simply tried to send the reader to sources that are as readable as possible for one knowing what is written here. A good systematic reference for the finite-group material, including proofs of the results we leave out, is Serre [Se2]. For Lie groups and Lie algebras, Serre [Se3], Adams [Ad], Humphreys [Hu1], and Bourbaki [Bour] are recommended references, as are the classics Weyl [We1] and Littlewood [Lit1].

We would like to thank the many people who have contributed ideas and suggestions for this manuscript, among them J-F. Burnol, R. Bryant, J. Carrell, B. Conrad, P. Diaconis, D. Eisenbud, D. Goldstein, M. Green, P. Griffiths, B. Gross, M. Hildebrand, R. Howe, H. Kraft, A. Landman, B. Mazur, N. Chriss, D. Petersen, G. Schwartz, J. Towber, and L. Tu. In particular, we would like to thank David Mumford, from whom we learned much of what we know about the subject, and whose ideas are very much in evidence in this book.

Had this book been written 10 years ago, we would at this point thank the people who typed it. That being no longer applicable, perhaps we should thank instead the National Science Foundation, the University of Chicago, and Harvard University for generously providing the various Macintoshes on which this manuscript was produced. Finally, we thank Chan Fulton for making the drawings.

Bill Fulton and Joe Harris

Note to the corrected fifth printing: We are grateful to S. Billey, M. Brion, R. Coleman, B. Gross, E. D'Hoker, D. Jaffe, R. Milson, K. Rumelhart, M. Reeder, and J. Willenbring for pointing out errors in earlier printings, and to many others for telling us about misprints.

Using This Book

A few words are in order about the practical use of this book. To begin with, prerequisites are minimal: we assume only a basic knowledge of standard first-year graduate material in algebra and topology, including basic notions about manifolds. A good undergraduate background should be more than enough for most of the text; some examples and exercises, and some of the discussion in Part IV may refer to more advanced topics, but these can readily be skipped. Probably the main practical requirement is a good working knowledge of multilinear algebra, including tensor, exterior, and symmetric products of finite dimensional vector spaces, for which Appendix B may help. We have indicated, in introductory remarks to each lecture, when any background beyond this is assumed and how essential it is.

For a course, this book could be used in two ways. First, there are a number of topics that are not logically essential to the rest of the book and that can be skimmed or skipped entirely. For example, in a minimal reading one could skip §§4, 5, 6, 11.3, 13.4, 15.3–15.5, 17.3, 19.5, 20, 22.1, 22.3, 23.3–23.4, 25.3, and 26.2; this might be suitable for a basic one-semester course. On the other hand, in a year-long course it should be possible to work through as much of the material as background and/or interest suggested. Most of the material in the Appendices is relevant only to such a long course. Again, we have tried to indicate, in the introductory remarks in each lecture, which topics are inessential and may be omitted.

Another aspect of the book that readers may want to approach in different ways is the profusion of examples. These are put in largely for didactic reasons: we feel that this is the sort of material that can best be understood by gaining some direct hands-on experience with the objects involved. For the most part, however, they do not actually develop new ideas; the reader whose tastes run more to the abstract and general than the concrete and special may skip many of them without logical consequence. (Of course, such a reader will probably wind up burning this book anyway.)

We include hundreds of exercises, of wildly different purposes and difficulties. Some are the usual sorts of variations of the examples in the text or are straightforward verifications of facts needed; a student will probably want to attempt most of these. Sometimes an exercise is inserted whose solution is a special case of something we do in the text later, if we think working on it will be useful motivation (again, there is no attempt at "efficiency," and readers are encouraged to go back to old exercises from time to time). Many exercises are included that indicate some further directions or new topics (or standard topics we have omitted); a beginner may best be advised to skim these for general information, perhaps working out a few simple cases. In exercises, we tried to include topics that may be hard for nonexperts to extract from the literature, especially the older literature. In general, much of the theory is in the exercises—and most of the examples in the text.

We have resisted the idea of grading the exercises by (expected) difficulty, although a "problem" is probably harder than an "exercise." Many exercises are starred: the * is not an indication of difficulty, but means that the reader can find some information about it in the section "Hints, Answers, and References" at the back of the book. This may be a hint, a statement of the answer, a complete solution, a reference to where more can be found, or a combination of any of these. We hope these miscellaneous remarks, as haphazard and uneven as they are, will be of some use.

Contents

PART I
FINITE GROUPS

Given that over three-quarters of this book is devoted to the representation theory of Lie groups and Lie algebras, why have a discussion of the representations of finite groups at all? There are certainly valid reasons from a logical point of view: many of the ideas, concepts, and constructions we will introduce here will be applied in the study of Lie groups and algebras. The real reason for us, however, is didactic, as we will now try to explain.

Representation theory is very much a 20th-century subject, in the following sense. In the 19th century, when groups were dealt with they were generally understood to be subsets of the permutations of a set, or of the automorphisms $\mathrm{GL}(V)$ of a vector space V, closed under composition and inverse. Only in the 20th century was the notion of an abstract group given, making it possible to make a distinction between properties of the abstract group and properties of the particular realization as a subgroup of a permutation group or $\mathrm{GL}(V)$. To give an analogy, in the 19th century a manifold was always a subset of $\mathbb{R}^n$; only in the 20th century did the notion of an abstract Riemannian manifold become common.

In both cases, the introduction of the abstract object made a fundamental difference to the subject. In differential geometry, one could make a crucial distinction between the intrinsic and extrinsic geometry of the manifold: which properties were invariants of the metric on the manifold and which were properties of the particular embedding in $\mathbb{R}^n$. Questions of existence or non-existence, for example, could be broken up into two parts: did the abstract manifold exist, and could it be embedded. Similarly, what would have been called in the 19th century simply "group theory" is now factored into two parts. First, there is the study of the structure of abstract groups (e.g., the classification of simple groups). Second is the companion question: given a group G, how can we describe all the ways in which G may be embedded in (or mapped to) a linear group $GL(V)$?. This, of course, is the subject matter of representation theory.

Given this point of view, it makes sense when first introducing representation theory to do so in a context where the nature of the groups G in question is itself simple, and relatively well understood. It is largely for this reason that we are starting off with the representation theory of finite groups: for those readers who are not already familiar with the motivations and goals of representation theory, it seemed better to establish those first in a setting where the structure of the groups was not itself an issue. When we analyze, for example, the representations of the symmetric and alternating groups on 3, 4, and 5 letters, it can be expected that the reader is already familiar with the groups and can focus on the basic concepts of representation theory being introduced.

We will spend the first six lectures on the case of finite groups. Many of the techniques developed for finite groups will carry over to Lie groups; indeed, our choice of topics is in part guided by this. For example, we spend quite a bit of time on the symmetric group; this is partly for its own interest, but also partly because what we learn here gives one way to study representations of the general linear group and its subgroups. There are other topics, such as the alternating group $\mathfrak{A}_d$, and the groups $SL_2(\mathbb{F}_q)$ and $GL_2(\mathbb{F}_q)$ that are studied purely for their own interest and do not appear later. (In general, for those readers primarily concerned with Lie theory, we have tried to indicate in the introductory notes to each lecture which ideas will be useful in the succeeding parts of this book.) Nonetheless, this is by no means a comprehensive treatment of the representation theory of finite groups; many important topics, such as the Artin and Brauer theorems and the whole subject of modular representations, are omitted.

LECTURE 1

Representations of Finite Groups

In this lecture we give the basic definitions of representation theory, and prove two of the basic results, showing that every representation is a (unique) direct sum of irreducible ones. We work out as examples the case of abelian groups, and the simplest nonabelian group, the symmetric group on 3 letters. In the latter case we give an analysis that will turn out not to be useful for the study of finite groups, but whose main idea is central to the study of the representations of Lie groups.

§1.1. Definitions

A *representation* of a finite group G on a finite-dimensional complex vector space V is a homomorphism $\rho: G \to \mathrm{GL}(V)$ of G to the group of automorphisms of V; we say that such a map *gives V the structure of a G-module.* When there is little ambiguity about the map ρ (and, we're afraid, even sometimes when there is) we sometimes call V itself a representation of G; in this vein we will often suppress the symbol ρ and write $g \cdot v$ or gv for $\rho(g)(v)$. The dimension of V is sometimes called the *degree* of ρ.

A *map* φ between two representations V and W of G is a vector space map $\varphi: V \to W$ such that

$$\begin{tikzcd} V \arrow[r, "\varphi"] \arrow[d, "g"'] & W \arrow[d, "g"] \\ V \arrow[r, "\varphi"] & W \end{tikzcd}$$

commutes for every $g \in G$. (We will call this a *G-linear map* when we want to distinguish it from an arbitrary linear map between the vector spaces V and W.) We can then define Ker φ, Im φ, and Coker φ, which are also G-modules.

A *subrepresentation* of a representation V is a vector subspace W of V which is invariant under G. A representation V is called *irreducible* if there is no proper nonzero invariant subspace W of V.

If V and W are representations, the *direct sum* $V \oplus W$ and the *tensor product* $V \otimes W$ are also representations, the latter via

$$g(v \otimes w) = gv \otimes gw.$$

For a representation V, the nth tensor power $V^{\otimes n}$ is again a representation of G by this rule, and the *exterior powers* $\Lambda^n(V)$ and *symmetric powers* $\mathrm{Sym}^n(V)$ are subrepresentations[1] of it. The *dual* $V^* = \mathrm{Hom}(V, \mathbb{C})$ of V is also a representation, though not in the most obvious way: we want the two representations of G to respect the natural pairing (denoted $\langle\ ,\ \rangle$) between V^* and V, so that if $\rho: G \to \mathrm{GL}(V)$ is a representation and $\rho^*: G \to \mathrm{GL}(V^*)$ is the dual, we should have

$$\langle \rho^*(g)(v^*), \rho(g)(v) \rangle = \langle v^*, v \rangle$$

for all $g \in G$, $v \in V$, and $v^* \in V^*$. This in turn forces us to define the dual representation by

$$\rho^*(g) = {}^t\rho(g^{-1}): V^* \to V^*$$

for all $g \in G$.

Exercise 1.1. Verify that with this definition of ρ^*, the relation above is satisfied.

Having defined the dual of a representation and the tensor product of two representations, it is likewise the case that if V and W are representations, then $\mathrm{Hom}(V, W)$ is also a representation, via the identification $\mathrm{Hom}(V, W) = V^* \otimes W$. Unraveling this, if we view an element of $\mathrm{Hom}(V, W)$ as a linear map φ from V to W, we have

$$(g\varphi)(v) = g\varphi(g^{-1}v)$$

for all $v \in V$. In other words, the definition is such that the diagram

$$\begin{tikzcd} V \arrow[r, "\varphi"] \arrow[d, "g"'] & W \arrow[d, "g"] \\ V \arrow[r, "g\varphi"] & W \end{tikzcd}$$

commutes. Note that the dual representation is, in turn, a special case of this:

[1] For more on exterior and symmetric powers, including descriptions as quotient spaces of tensor powers, see Appendix B.

when $W = \mathbb{C}$ is the *trivial* representation, i.e., $gw = w$ for all $w \in \mathbb{C}$, this makes V^* into a G-module, with $g\varphi(v) = \varphi(g^{-1}v)$, i.e., $g\varphi = {}^t(g^{-1})\varphi$.

Exercise 1.2. Verify that in general the vector space of G-linear maps between two representations V and W of G is just the subspace $\mathrm{Hom}(V, W)^G$ of elements of $\mathrm{Hom}(V, W)$ fixed under the action of G. This subspace is often denoted $\mathrm{Hom}_G(V, W)$.

We have, in effect, taken the identification $\mathrm{Hom}(V, W) = V^* \otimes W$ as the definition of the representation $\mathrm{Hom}(V, W)$. More generally, the usual identities for vector spaces are also true for representations, e.g.,

$$V \otimes (U \oplus W) = (V \otimes U) \oplus (V \otimes W),$$

$$\textstyle\bigwedge^k(V \oplus W) = \bigoplus_{a+b=k} \bigwedge^a V \otimes \bigwedge^b W,$$

$$\textstyle\bigwedge^k(V^*) = \bigwedge^k(V)^*,$$

and so on.

Exercise 1.3*. Let $\rho: G \to \mathrm{GL}(V)$ be any representation of the finite group G on an n-dimensional vector space V and suppose that for any $g \in G$, the determinant of $\rho(g)$ is 1. Show that the spaces $\bigwedge^k V$ and $\bigwedge^{n-k} V^*$ are isomorphic as representations of G.

If X is any finite set and G acts on the left on X, i.e., $G \to \mathrm{Aut}(X)$ is a homomorphism to the permutation group of X, there is an associated *permutation representation*: let V be the vector space with basis $\{e_x: x \in X\}$, and let G act on V by

$$g \cdot \sum a_x e_x = \sum a_x e_{gx}.$$

The *regular representation*, denoted R_G or R, corresponds to the left action of G on itself. Alternatively, R is the space of complex-valued functions on G, where an element $g \in G$ acts on a function α by $(g\alpha)(h) = \alpha(g^{-1}h)$.

Exercise 1.4*. (a) Verify that these two descriptions of R agree, by identifying the element e_x with the characteristic function which takes the value 1 on x, 0 on other elements of G.

(b) The space of functions on G can also be made into a G-module by the rule $(g\alpha)(h) = \alpha(hg)$. Show that this is an isomorphic representation.

§1.2. Complete Reducibility; Schur's Lemma

As in any study, before we begin our attempt to classify the representations of a finite group G in earnest we should try to simplify life by restricting our search somewhat. Specifically, we have seen that representations of G can be built up out of other representations by linear algebraic operations, most simply by taking the direct sum. We should focus, then, on representations that are "atomic" with respect to this operation, i.e., that cannot be expressed as a direct sum of others; the usual term for such a representation is *indecomposable*. Happily, the situation is as nice as it could possibly be: a representation is atomic in this sense if and only if it is irreducible (i.e., contains no proper subrepresentations); and every representation is the direct sum of irreducibles, in a suitable sense uniquely so. The key to all this is

Proposition 1.5. *If W is a subrepresentation of a representation V of a finite group G, then there is a complementary invariant subspace W' of V, so that $V = W \oplus W'$.*

PROOF. There are two ways of doing this. One can introduce a (positive definite) Hermitian inner product H on V which is preserved by each $g \in G$ (i.e., such that $H(gv, gw) = H(v, w)$ for all $v, w \in V$ and $g \in G$). Indeed, if H_0 is any Hermitian product on V, one gets such an H by averaging over G:

$$H(v, w) = \sum_{g \in G} H_0(gv, gw).$$

Then the perpendicular subspace $W^\perp$ is complementary to W in V. Alternatively (but similarly), we can simply choose an arbitrary subspace U complementary to W, let $\pi_0: V \to W$ be the projection given by the direct sum decomposition $V = W \oplus U$, and average the map π_0 over G: that is, take

$$\pi(v) = \sum_{g \in G} g(\pi_0(g^{-1}v)).$$

This will then be a G-linear map from V onto W, which is multiplication by $|G|$ on W; its kernel will, therefore, be a subspace of V invariant under G and complementary to W. □

Corollary 1.6. *Any representation is a direct sum of irreducible representations.*

This property is called *complete reducibility*, or *semisimplicity*. We will see that, for continuous representations, the circle S^1, or any compact group, has this property; integration over the group (with respect to an invariant measure on the group) plays the role of averaging in the above proof. The (additive) group $\mathbb{R}$ does not have this property: the representation

$$a \mapsto \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}$$

leaves the x axis fixed, but there is no complementary subspace. We will see other Lie groups such as $\mathrm{SL}_n(\mathbb{C})$ that are semisimple in this sense. Note also that this argument would fail if the vector space V was over a field of finite characteristic since it might then be the case that $\pi(v) = 0$ for $v \in W$. The failure of complete reducibility is one of the things that makes the subject of *modular representations*, or representations on vector spaces over finite fields, so tricky.

The extent to which the decomposition of an arbitrary representation into a direct sum of irreducible ones is unique is one of the consequences of the following:

Schur's Lemma 1.7. *If V and W are irreducible representations of G and $\varphi\colon V \to W$ is a G-module homomorphism, then*

(1) *Either φ is an isomorphism, or $\varphi = 0$.*
(2) *If $V = W$, then $\varphi = \lambda \cdot I$ for some $\lambda \in \mathbb{C}$, I the identity.*

PROOF. The first claim follows from the fact that Ker φ and Im φ are invariant subspaces. For the second, since $\mathbb{C}$ is algebraically closed, φ must have an eigenvalue λ, i.e., for some $\lambda \in \mathbb{C}$, $\varphi - \lambda I$ has a nonzero kernel. By (1), then, we must have $\varphi - \lambda I = 0$, so $\varphi = \lambda I$. □

We can summarize what we have shown so far in

Proposition 1.8. *For any representation V of a finite group G, there is a decomposition*

$$V = V_1^{\oplus a_1} \oplus \cdots \oplus V_k^{\oplus a_k},$$

where the V_i are distinct irreducible representations. The decomposition of V into a direct sum of the k factors is unique, as are the V_i that occur and their multiplicities a_i.

PROOF. It follows from Schur's lemma that if W is another representation of G, with a decomposition $W = \bigoplus W_j^{\oplus b_j}$, and $\varphi\colon V \to W$ is a map of representations, then φ must map the factor $V_i^{\oplus a_i}$ into that factor $W_j^{\oplus b_j}$ for which $W_j \cong V_i$; when applied to the identity map of V to V, the stated uniqueness follows. □

In the next lecture we will give a formula for the projection of V onto $V_i^{\oplus a_i}$. The decomposition of the ith summand into a direct sum of a_i copies of V_i is not unique if $a_i > 1$, however.

Occasionally the decomposition is written

$$V = a_1 V_1 \oplus \cdots \oplus a_k V_k = a_1 V_1 + \cdots + a_k V_k, \tag{1.9}$$

especially when one is concerned only about the isomorphism classes and multiplicities of the V_i.

One more fact that will be established in the following lecture is that a finite group G admits only finitely many irreducible representations V_i up to isomorphism (in fact, we will say how many). This, then, is the framework of the classification of all representations of G: by the above, once we have described the irreducible representations of G, we will be able to describe an arbitrary representation as a linear combination of these. Our first goal, in analyzing the representations of any group, will therefore be:

(i) *Describe all the irreducible representations of G.*

Once we have done this, there remains the problem of carrying out in practice the description of a given representation in these terms. Thus, our second goal will be:

(ii) *Find techniques for giving the direct sum decomposition* (1.9), *and in particular determining the multiplicities a_i of an arbitrary representation V.*

Finally, it is the case that the representations we will most often be concerned with are those arising from simpler ones by the sort of linear- or multilinear-algebraic operations described above. We would like, therefore, to be able to describe, in the terms above, the representation we get when we perform these operations on a known representation. This is known generally as

(iii) *Plethysm*: *Describe the decompositions, with multiplicities, of representations derived from a given representation V,* such as $V \otimes V$, V^*, $\wedge^k(V)$, $\mathrm{Sym}^k(V)$, and $\wedge^k(\wedge^l V)$. Note that if V decomposes into a sum of two representations, these representations decompose accordingly; e.g., if $V = U \oplus W$, then

$$\wedge^k V = \bigoplus_{i+j=k} \wedge^i U \otimes \wedge^j W,$$

so it is enough to work out this plethysm for irreducible representations. Similarly, if V and W are two irreducible representations, we want to decompose $V \otimes W$; this is usually known as the *Clebsch–Gordan* problem.

§1.3. Examples: Abelian Groups; $\mathfrak{S}_3$

One obvious place to look for examples is with abelian groups. It does not take long, however, to deal with this case. Basically, we may observe in general that if V is a representation of the finite group G, abelian or not, each $g \in G$ gives a map $\rho(g)\colon V \to V$; but *this map is not generally a G-module homomorphism*: for general $h \in G$ we will have

$$g(h(v)) \neq h(g(v)).$$

Indeed, *$\rho(g)\colon V \to V$ will be G-linear for every ρ if (and only if) g is in the center* $Z(G)$ of G. In particular if G is abelian, and V is an irreducible representation, then by Schur's lemma every element $g \in G$ acts on V by a scalar multiple of the identity. Every subspace of V is thus invariant; so that V must be one dimensional. The irreducible representations of an abelian group G are thus simply elements of the dual group, that is, homomorphisms

$$\rho\colon G \to \mathbb{C}^*.$$

We consider next the simplest nonabelian group, $G = \mathfrak{S}_3$. To begin with, we have (as with any nontrivial symmetric group) two one-dimensional representations: we have the trivial representation, which we will denote U, and the *alternating representation* U', defined by setting

$$gv = \operatorname{sgn}(g)v$$

for $g \in G$, $v \in \mathbb{C}$. Next, since G comes to us as a permutation group, we have a natural permutation representation, in which G acts on $\mathbb{C}^3$ by permuting the coordinates. Explicitly, if $\{e_1, e_2, e_3\}$ is the standard basis, then $g \cdot e_i = e_{g(i)}$, or, equivalently,

$$g \cdot (z_1, z_2, z_3) = (z_{g^{-1}(1)}, z_{g^{-1}(2)}, z_{g^{-1}(3)}).$$

This representation, like any permutation representation, is not irreducible: the line spanned by the sum (1, 1, 1) of the basis vectors is invariant, with complementary subspace

$$V = \{(z_1, z_2, z_3) \in \mathbb{C}^3 \colon z_1 + z_2 + z_3 = 0\}.$$

This two-dimensional representation V is easily seen to be irreducible; we call it the *standard representation* of $\mathfrak{S}_3$.

Let us now turn to the problem of describing an arbitrary representation of $\mathfrak{S}_3$. We will see in the next lecture a wonderful tool for doing this, called *character theory*; but, as inefficient as this may be, we would like here to adopt a more ad hoc approach. This has some virtues as a didactic technique in the present context (admittedly dubious ones, consisting mainly of making the point that there are other and far worse ways of doing things than character theory). The real reason we are doing it is that it will serve to introduce an idea that, while superfluous for analyzing the representations of finite groups in general, will prove to be the key to understanding representations of Lie groups.

The idea is a very simple one: since we have just seen that the representation theory of a finite abelian group is virtually trivial, we will start our analysis of an arbitrary representation W of $\mathfrak{S}_3$ by looking just at the action of the abelian subgroup $\mathfrak{A}_3 = \mathbb{Z}/3 \subset \mathfrak{S}_3$ on W. This yields a very simple decomposition: if we take τ to be any generator of $\mathfrak{A}_3$ (that is, any three-cycle), the space W is spanned by eigenvectors v_i for the action of τ, whose eigenvalues are of course all powers of a cube root of unity $\omega = e^{2\pi i/3}$. Thus,

$$W = \bigoplus V_i,$$

where

$$V_i = \mathbb{C}v_i \quad \text{and} \quad \tau v_i = \omega^{\alpha_i} v_i.$$

Next, we ask how the remaining elements of $\mathfrak{S}_3$ act on W in terms of this decomposition. To see how this goes, let σ be any transposition, so that τ and σ together generate $\mathfrak{S}_3$, with the relation $\sigma\tau\sigma = \tau^2$. We want to know where σ sends an eigenvector v for the action of τ, say with eigenvalue ω^i; to answer this, we look at how τ acts on $\sigma(v)$. We use the basic relation above to write

$$\begin{aligned}\tau(\sigma(v)) &= \sigma(\tau^2(v))\\ &= \sigma(\omega^{2i}\cdot v)\\ &= \omega^{2i}\cdot\sigma(v).\end{aligned}$$

The conclusion, then, is that *if v is an eigenvector for τ with eigenvalue ω^i, then $\sigma(v)$ is again an eigenvector for τ, with eigenvalue ω^{2i}.*

Exercise 1.10. Verify that with $\sigma = (12)$, $\tau = (123)$, the standard representation has a basis $\alpha = (\omega, 1, \omega^2)$, $\beta = (1, \omega, \omega^2)$, with

$$\tau\alpha = \omega\alpha, \qquad \tau\beta = \omega^2\beta, \qquad \sigma\alpha = \beta, \qquad \sigma\beta = \alpha.$$

Suppose now that we start with such an eigenvector v for τ. If the eigenvalue of v is $\omega^i \neq 1$, then $\sigma(v)$ is an eigenvector with eigenvalue $\omega^{2i} \neq \omega^i$, and so is independent of v; and v and $\sigma(v)$ together span a two-dimensional subspace V' of W invariant under $\mathfrak{S}_3$. In fact, V' is isomorphic to the standard representation, which follows from Exercise 1.10. If, on the other hand, the eigenvalue of v is 1, then $\sigma(v)$ may or may not be independent of v. If it is not, then v spans a one-dimensional subrepresentation of W, isomorphic to the trivial representation if $\sigma(v) = v$ and to the alternating representation if $\sigma(v) = -v$. If $\sigma(v)$ and v are independent, then $v + \sigma(v)$ and $v - \sigma(v)$ span one-dimensional representations of W isomorphic to the trivial and alternating representations, respectively.

We have thus accomplished the first two of the goals we have set for ourselves above in the case of the group $G = \mathfrak{S}_3$. First, we see from the above that *the only three irreducible representations of $\mathfrak{S}_3$ are the trivial, alternating, and standard representations U, U' and V.* Moreover, for an arbitrary representation W of $\mathfrak{S}_3$ we can write

$$W = U^{\oplus a} \oplus U'^{\oplus b} \oplus V^{\oplus c};$$

and we have a way to determine the multiplicities a, b, and c: c, for example, is the number of independent eigenvectors for τ with eigenvalue ω, whereas $a + c$ is the multiplicity of 1 as an eigenvalue of σ, and $b + c$ is the multiplicity of -1 as an eigenvalue of σ.

In fact, this approach gives us as well the answer to our third problem, finding the decomposition of the symmetric, alternating, or tensor powers of a given representation W, since if we know the eigenvalues of τ on such a representation, we know the eigenvalues of τ on the various tensor powers of W. For example, we can use this method to decompose $V \otimes V$, where V is the standard two-dimensional representation. For $V \otimes V$ is spanned by the vectors $\alpha\otimes\alpha$, $\alpha\otimes\beta$, $\beta\otimes\alpha$, and $\beta\otimes\beta$; these are eigenvectors for τ with eigenvalues ω^2, 1, 1, and ω, respectively, and σ interchanges $\alpha\otimes\alpha$ with $\beta\otimes\beta$, and $\alpha\otimes\beta$ with $\beta\otimes\alpha$. Thus $\alpha\otimes\alpha$ and $\beta\otimes\beta$ span a subrepresentation isomorphic to V, $\alpha \otimes \beta + \beta \otimes \alpha$ spans a trivial representation U, and $\alpha \otimes \beta - \beta \otimes \alpha$ spans U', so

$$V \otimes V \cong U \oplus U' \oplus V.$$

Exercise 1.11. Use this approach to find the decomposition of the representations $\mathrm{Sym}^2 V$ and $\mathrm{Sym}^3 V$.

Exercise 1.12. (a) Decompose the regular representation R of $\mathfrak{S}_3$.

(b) Show that $\mathrm{Sym}^{k+6} V$ is isomorphic to $\mathrm{Sym}^k V \oplus R$, and compute $\mathrm{Sym}^k V$ for all k.

Exercise 1.13*. Show that $\mathrm{Sym}^2(\mathrm{Sym}^3 V) \cong \mathrm{Sym}^3(\mathrm{Sym}^2 V)$. Is $\mathrm{Sym}^m(\mathrm{Sym}^n V)$ isomorphic to $\mathrm{Sym}^n(\mathrm{Sym}^m V)$?

As we have indicated, the idea of studying a representation V of a group G by first restricting the action to an abelian subgroup, getting a decomposition of V into one-dimensional invariant subspaces, and then asking how the remaining generators of the group act on these subspaces, does not work well for finite G in general; for one thing, there will not in general be a convenient abelian subgroup to use. This idea will turn out, however, to be the key to understanding the representations of Lie groups, with a torus subgroup playing the role of the cyclic subgroup in this example.

Exercise 1.14*. Let V be an irreducible representation of the finite group G. Show that, up to scalars, there is a *unique* Hermitian inner product on V preserved by G.

LECTURE 2

Characters

This lecture contains the heart of our treatment of the representation theory of finite groups: the definition in §2.1 of the character of a representation, and the main theorem (proved in two steps in §2.2 and §2.4) that the characters of the irreducible representations form an orthonormal basis for the space of class functions on G. There will be more examples and more constructions in the following lectures, but this is what you need to know.

§2.1. Characters

As we indicated in the preceding section, there is a remarkably effective tool for understanding the representations of a finite group G, called *character theory*. This is in some ways motivated by the example worked out in the last section where we saw that a representation of $\mathfrak{S}_3$ was determined by knowing the eigenvalues of the action of the elements τ and $\sigma \in \mathfrak{S}_3$. For a general group G, it is not clear what subgroups and/or elements should play the role of $\mathfrak{A}_3$, τ, and σ; but the example certainly suggests that knowing all the eigenvalues of each element of G should suffice to describe the representation.

Of course, specifying all the eigenvalues of the action of each element of G is somewhat unwieldy; but fortunately it is redundant as well. For example, if we know the eigenvalues $\{\lambda_i\}$ of an element $g \in G$, then of course we know the eigenvalues $\{\lambda_i^k\}$ of g^k for each k as well. We can thus use this redundancy to simplify the data we have to specify. The key observation here is it is enough to give, for example, just the *sum* of the eigenvalues of each element of G, since knowing the sums $\sum \lambda_i^k$ of the kth powers of the eigenvalues of a given element $g \in G$ is equivalent to knowing the eigenvalues $\{\lambda_i\}$ of g themselves. This then suggests the following:

Definition. If V is a representation of G, its *character* χ_V is the complex-valued function on the group defined by

$$\chi_V(g) = \mathrm{Tr}(g|_V),$$

the trace of g on V.

In particular, we have

$$\chi_V(hgh^{-1}) = \chi_V(g),$$

so that χ_V is constant on the conjugacy classes of G; such a function is called a *class function*. Note that $\chi_V(1) = \dim V$.

Proposition 2.1. *Let V and W be representations of G. Then*

$$\chi_{V\oplus W} = \chi_V + \chi_W, \qquad \chi_{V\otimes W} = \chi_V \cdot \chi_W,$$
$$\chi_{V^*} = \overline{\chi}_V \quad \textit{and} \quad \chi_{\wedge^2 V}(g) = \tfrac{1}{2}[\chi_V(g)^2 - \chi_V(g^2)].$$

Proof. We compute the values of these characters on a fixed element $g \in G$. For the action of g, V has eigenvalues $\{\lambda_i\}$ and W has eigenvalues $\{\mu_i\}$. Then $\{\lambda_i\} \cup \{\mu_j\}$ and $\{\lambda_i \cdot \mu_j\}$ are eigenvalues for $V \oplus W$ and $V \otimes W$, from which the first two formulas follow. Similarly $\{\lambda_i^{-1} = \overline{\lambda}_i\}$ are the eigenvalues for g on V^*, since all eigenvalues are nth roots of unity, with n the order of g. Finally, $\{\lambda_i\lambda_j | i < j\}$ are the eigenvalues for g on $\wedge^2 V$, and

$$\sum_{i<j} \lambda_i\lambda_j = \frac{(\sum \lambda_i)^2 - \sum \lambda_i^2}{2};$$

and since g^2 has eigenvalues $\{\lambda_i^2\}$, the last formula follows. □

Exercise 2.2. For $\mathrm{Sym}^2 V$, verify that

$$\chi_{\mathrm{Sym}^2 V}(g) = \tfrac{1}{2}[\chi_V(g)^2 + \chi_V(g^2)].$$

Note that this is compatible with the decomposition

$$V \otimes V = \mathrm{Sym}^2 V \oplus \wedge^2 V.$$

Exercise 2.3*. Compute the characters of $\mathrm{Sym}^k V$ and $\wedge^k V$.

Exercise 2.4*. Show that if we know the character χ_V of a representation V, then we know the eigenvalues of each element g of G, in the sense that we know the coefficients of the characteristic polynomial of $g\colon V \to V$. Carry this out explicitly for elements $g \in G$ of orders 2, 3, and 4, and for a representation of G on a vector space of dimension 2, 3, or 4.

Exercise 2.5. (*The original fixed-point formula*). If V is the permutation representation associated to the action of a group G on a finite set X, show that $\chi_V(g)$ is the number of elements of X fixed by g.

As we have said, the character of a representation of a group G is really a function on the set of conjugacy classes in G. This suggests expressing the basic information about the irreducible representations of a group G in the form of a *character table*. This is a table with the conjugacy classes $[g]$ of G listed across the top, usually given by a representative g, with (for reasons that will become apparent later) the number of elements in each conjugacy class over it; the irreducible representations V of G listed on the left; and, in the appropriate box, the value of the character χ_V on the conjugacy class $[g]$.

Example 2.6. We compute the character table of $\mathfrak{S}_3$. This is easy: to begin with, the trivial representation takes the values (1, 1, 1) on the three conjugacy classes [1], [(12)], and [(123)], whereas the alternating representation has values $(1, -1, 1)$. To see the character of the standard representation, note that the permutation representation decomposes: $\mathbb{C}^3 = U \oplus V$; since the character of the permutation representation has, by Exercise 2.5, the values (3, 1, 0), we have $\chi_V = \chi_{\mathbb{C}^3} - \chi_U = (3, 1, 0) - (1, 1, 1) = (2, 0, -1)$. In sum, then, the character table of $\mathfrak{S}_3$ is

	1	3	2
$\mathfrak{S}_3$	1	(12)	(123)
trivial U	1	1	1
alternating U'	1	-1	1
standard V	2	0	-1

This gives us another solution of the basic problem posed in Lecture 1: if W is any representation of $\mathfrak{S}_3$ and we decompose W into irreducible representations $W \cong U^{\oplus a} \oplus U'^{\oplus b} \oplus V^{\oplus c}$, then $\chi_W = a\chi_U + b\chi_{U'} + c\chi_V$. In particular, since the functions χ_U, $\chi_{U'}$ and χ_V are independent, we see that W *is determined up to isomorphism by its character* χ_W.

Consider, for example, $V \otimes V$. Its character is $(\chi_V)^2$, which has values 4, 0, and 1 on the three conjugacy classes. Since $V \oplus U \oplus U'$ has the same character, this implies that $V \otimes V$ decomposes into $V \oplus U \oplus U'$, as we have seen directly. Similarly, $V \otimes U'$ has values 2, 0, and -1, so $V \otimes U' \cong V$.

Exercise 2.7*. Find the decomposition of the representation $V^{\otimes n}$ using character theory.

Characters will be similarly useful for larger groups, although it is rare to find simple closed formulas for decomposing tensor products.

§2.2. The First Projection Formula and Its Consequences

In the last lecture, we asked (among other things) for a way of locating explicitly the direct sum factors in the decomposition of a representation into irreducible ones. In this section we will start by giving an explicit formula for the projection of a representation onto the direct sum of the trivial factors in this decomposition; as it will turn out, this formula alone has tremendous consequences.

To start, for any representation V of a group G, we set

$$V^G = \{v \in V: gv = v \quad \forall g \in G\}.$$

We ask for a way of finding V^G explicitly. The idea behind our solution to this is already implicit in the previous lecture. We observed there that for any representation V of G and any $g \in G$, the endomorphism $g: V \to V$ is, in general, not a G-module homomorphism. On the other hand, if we take the *average* of all these endomorphisms, that is, we set

$$\varphi = \frac{1}{|G|} \sum_{g \in G} g \in \operatorname{End}(V),$$

then the endomorphism φ will be G-linear since $\sum g = \sum hgh^{-1}$. In fact, we have

Proposition 2.8. *The map φ is a projection of V onto V^G.*

Proof. First, suppose $v = \varphi(w) = (1/|G|)\sum gw$. Then, for any $h \in G$,

$$hv = \frac{1}{|G|} \sum hgw = \frac{1}{|G|} \sum gw,$$

so the image of φ is contained in V^G. Conversely, if $v \in V^G$, then $\varphi(v) = (1/|G|)\sum v = v$, so $V^G \subset \operatorname{Im}(\varphi)$; and $\varphi \circ \varphi = \varphi$. □

We thus have a way of finding explicitly the direct sum of the trivial subrepresentations of a given representation, although the formula can be hard to use if it does not simplify. If we just want to know the number m of copies of the trivial representation appearing in the decomposition of V, we can do this numerically, since this number will be just the trace of the projection φ. We have

$$m = \dim V^G = \operatorname{Trace}(\varphi)$$
$$= \frac{1}{|G|} \sum_{g \in G} \operatorname{Trace}(g) = \frac{1}{|G|} \sum_{g \in G} \chi_V(g). \tag{2.9}$$

In particular, we observe that for an irreducible representation V other than the trivial one, the sum over all $g \in G$ of the values of the character χ_V is zero.

We can do much more with this idea, however. The key is to use Exercise 1.2: if V and W are representations of G, then with $\operatorname{Hom}(V, W)$, the representation defined in Lecture 1, we have

$$\operatorname{Hom}(V, W)^G = \{G\text{-module homomorphisms from } V \text{ to } W\}.$$

If V is irreducible then by Schur's lemma $\dim \operatorname{Hom}(V, W)^G$ is the multiplicity of V in W; similarly, if W is irreducible, $\dim \operatorname{Hom}(V, W)^G$ is the multiplicity of W in V, and in the case where both V and W are irreducible, we have

$$\dim \operatorname{Hom}_G(V, W) = \begin{cases} 1 & \text{if } V \cong W \\ 0 & \text{if } V \ncong W. \end{cases}$$

But now the character $\chi_{\operatorname{Hom}(V, W)}$ of the representation $\operatorname{Hom}(V, W) = V^* \otimes W$ is given by

$$\chi_{\operatorname{Hom}(V, W)}(g) = \overline{\chi_V(g)} \cdot \chi_W(g).$$

We can now apply formula (2.9) in this case to obtain the striking

$$\frac{1}{|G|} \sum_{g \in G} \overline{\chi_V(g)} \chi_W(g) = \begin{cases} 1 & \text{if } V \cong W \\ 0 & \text{if } V \ncong W. \end{cases} \tag{2.10}$$

To express this, let

$$\mathbb{C}_{\text{class}}(G) = \{\text{class functions on } G\}$$

and define an Hermitian inner product on $\mathbb{C}_{\text{class}}(G)$ by

$$(\alpha, \beta) = \frac{1}{|G|} \sum_{g \in G} \overline{\alpha(g)} \beta(g). \tag{2.11}$$

Formula (2.10) then amounts to

Theorem 2.12. *In terms of this inner product, the characters of the irreducible representations of G are orthonormal.*

For example, the orthonormality of the three irreducible representations of $\mathfrak{S}_3$ can be read from its character table in Example 2.6. The numbers over each conjugacy class tell how many times to count entries in that column.

Corollary 2.13. *The number of irreducible representations of G is less than or equal to the number of conjugacy classes.*

We will soon show that there are no nonzero class functions orthogonal to the characters, so that equality holds in Corollary 2.13.

Corollary 2.14. *Any representation is determined by its character.*

Indeed if $V \cong V_1^{\oplus a_1} \oplus \cdots \oplus V_k^{\oplus a_k}$, with the V_i distinct irreducible representations, then $\chi_V = \sum a_i \chi_{V_i}$, and the χ_{V_i} are linearly independent.

Corollary 2.15. *A representation V is irreducible if and only if $(\chi_V, \chi_V) = 1$.*

In fact, if $V \cong V_1^{\oplus a_1} \oplus \cdots \oplus V_k^{\oplus a_k}$ as above, then $(\chi_V, \chi_V) = \sum a_i^2$.

The multiplicities a_i can be calculated via

Corollary 2.16. *The multiplicity a_i of V_i in V is the inner product of χ_V with χ_{V_i}, i.e., $a_i = (\chi_V, \chi_{V_i})$.*

We obtain some further corollaries by applying all this to the regular representation R of G. First, by Exercise 2.5 we know the character of R; it is simply

$$\chi_R(g) = \begin{cases} 0 & \text{if } g \neq e \\ |G| & \text{if } g = e. \end{cases}$$

Thus, we see first of all that R is not irreducible if $G \neq \{e\}$. In fact, if we set $R = \bigoplus V_i^{\oplus a_i}$, with V_i distinct irreducibles, then

$$a_i = (\chi_{V_i}, \chi_R) = \frac{1}{|G|} \chi_{V_i}(e) \cdot |G| = \dim V_i. \tag{2.17}$$

Corollary 2.18. *Any irreducible representation V of G appears in the regular representation* dim *V times.*

In particular, this proves again that there are only finitely many irreducible representations. As a numerical consequence of this we have the formula

$$|G| = \dim(R) = \sum_i \dim(V_i)^2. \tag{2.19}$$

Also, applying this to the value of the character of the regular representation on an element $g \in G$ other than the identity, we have

$$0 = \sum (\dim V_i) \cdot \chi_{V_i}(g) \quad \text{if } g \neq e. \tag{2.20}$$

These two formulas amount to the Fourier inversion formula for finite groups, cf. Example 3.32. For example, if all but one of the characters is known, they give a formula for the unknown character.

Exercise 2.21. The orthogonality of the rows of the character table is equivalent to an orthogonality for the columns (assuming the fact that there are as many rows as columns). Written out, this says:

(i) For $g \in G$,

$$\sum_{\chi} \overline{\chi(g)}\chi(g) = \frac{|G|}{c(g)},$$

where the sum is over all irreducible characters, and $c(g)$ is the number of elements in the conjugacy class of g.

(ii) If g and h are elements of G that are not conjugate, then

$$\sum_{\chi} \overline{\chi(g)}\chi(h) = 0.$$

Note that for $g = e$ these reduce to (2.19) and (2.20).

§2.3. Examples: $\mathfrak{S}_4$ and $\mathfrak{A}_4$

To see how the analysis of the characters of a group actually goes in practice, we now work out the character table of $\mathfrak{S}_4$. To start, we list the conjugacy classes in $\mathfrak{S}_4$ and the number of elements of $\mathfrak{S}_4$ in each. As with any symmetric group $\mathfrak{S}_d$, the conjugacy classes correspond naturally to the *partitions* of d, that is, expressions of d as a sum of positive integers $a_1, a_2, \ldots, a_k$, where the correspondence associates to such a partition the conjugacy class of a permutation consisting of disjoint cycles of length $a_1, a_2, \ldots, a_k$. Thus, in $\mathfrak{S}_4$ we have the classes of the identity element 1 ($4 = 1 + 1 + 1 + 1$), a transposition such as (12), corresponding to the partition $4 = 2 + 1 + 1$; a three-cycle (123) corresponding to $4 = 3 + 1$; a four-cycle (1234) ($4 = 4$); and the product of two disjoint transpositions (12)(34) ($4 = 2 + 2$).

Exercise 2.22. Show that the number of elements in each of these conjugacy classes is, respectively, 1, 6, 8, 6, and 3.

As for the irreducible representations of $\mathfrak{S}_4$, we start with the same ones that we had in the case of $\mathfrak{S}_3$: the trivial U, the alternating U', and the standard representation V, i.e., the quotient of the permutation representation associated to the standard action of $\mathfrak{S}_4$ on a set of four elements by the trivial subrepresentation. The character of the trivial representation on the five conjugacy classes is of course (1, 1, 1, 1, 1), and that of the alternating representation is $(1, -1, 1, -1, 1)$. To find the character of the standard representation, we observe that by Exercise 2.5 the character of the permutation representation on $\mathbb{C}^4$ is $\chi_{\mathbb{C}^4} = (4, 2, 1, 0, 0)$ and, correspondingly,

$$\chi_V = \chi_{\mathbb{C}^4} - \chi_U = (3, 1, 0, -1, -1).$$

Note that $|\chi_V| = 1$, so V is irreducible. The character table so far looks like

	1	6	8	6	3
$\mathfrak{S}_4$	1	(12)	(123)	(1234)	(12)(34)
trivial U	1	1	1	1	1
alternating U'	1	-1	1	-1	1
standard V	3	1	0	-1	-1

Clearly, we are not done yet: since the sum of the squares of the dimensions of these three representations is $1 + 1 + 9 = 11$, by (2.19) there must be additional irreducible representations of $\mathfrak{S}_4$, the squares of whose dimensions add up to $24 - 11 = 13$. Since there are by Corollary 2.13 at most two of them, there must be exactly two, of dimensions 2 and 3. The latter of these is easy to locate: if we just tensor the standard representation V with the alternating one U', we arrive at a representation V' with character $\chi_{V'} = \chi_V \cdot \chi_{U'} = (3, -1, 0, 1, -1)$. We can see that this is irreducible either from its character (since $|\chi_{V'}| = 1$) or from the fact that it is the tensor product of an irreducible representation with a one-dimensional one; since its character is not equal to that of any of the first three, this must be one of the two missing ones. As for the remaining representation of degree two, we will for now simply call it W; we can determine its character from the orthogonality relations (2.10). We obtain then the complete character table for $\mathfrak{S}_4$:

	1	6	8	6	3
$\mathfrak{S}_4$	1	(12)	(123)	(1234)	(12)(34)
trivial U	1	1	1	1	1
alternating U'	1	-1	1	-1	1
standard V	3	1	0	-1	-1
$V' = V \otimes U'$	3	-1	0	1	-1
Another W	2	0	-1	0	2

Exercise 2.23. Verify the last row of this table from (2.10) or (2.20).

We now get a dividend: we can take the character of the mystery representation W, which we have obtained from general character theory alone, and use it to describe the representation W explicitly! The key is the 2 in the last column for χ_W: this says that the action of (12)(34) on the two-dimensional vector space W is an involution of trace 2, and so must be the identity. Thus, W is really a representation of the quotient group[1]

[1] If N is a normal subgroup of a group G, a representation $\rho: G \to \mathrm{GL}(V)$ is trivial on N if and only if it factors through the quotient

$$G \to G/N \to \mathrm{GL}(V).$$

Representations of G/N can be identified with representations of G that are trivial on N.

$$\mathfrak{S}_4/\{1, (12)(34), (13)(24), (14)(23)\} \cong \mathfrak{S}_3.$$

[One may see this isomorphism by letting $\mathfrak{S}_4$ act on the elements of the conjugacy class of $(12)(34)$; equivalently, if we realize $\mathfrak{S}_4$ as the group of rigid motions of a cube (see below), by looking at the action of $\mathfrak{S}_4$ on pairs of opposite faces.] W must then be just the standard representation of $\mathfrak{S}_3$ pulled back to $\mathfrak{S}_4$ via this quotient.

Example 2.24. As we said above, the group of rigid motions of a cube is the symmetric group on four letters; $\mathfrak{S}_4$ acts on the cube via its action on the four long diagonals. It follows, of course, that $\mathfrak{S}_4$ acts as well on the set of faces, of edges, of vertices, etc.; and to each of these is associated a permutation representation of $\mathfrak{S}_4$. We may thus ask how these representations decompose; we will do here the case of the faces and leave the others as exercises.

We start, of course, by describing the character χ of the permutation representation associated to the faces of the cube. Rotation by 180° about a line joining the midpoints of two opposite edges is a transposition in $\mathfrak{S}_4$ and fixes no faces, so $\chi(12) = 0$. Rotation by 120° about a long diagonal shows $\chi(123) = 0$. Rotation by 90° about a line joining the midpoints of two opposite faces shows $\chi(1234) = 2$, and rotation by 180° gives $\chi((12)(34)) = 2$. Now $(\chi, \chi) = 3$, so χ is the sum of three distinct irreducible representations. From the table, $(\chi, \chi_U) = (\chi, \chi_{V'}) = (\chi, \chi_W) = 1$, and the inner products with the others are zero, so this representation is $U \oplus V' \oplus W$. In fact, the sums of opposite faces span a three-dimensional subrepresentation which contains U (spanned by the sum of all faces), so this representation is $U \oplus W$. The differences of opposite faces therefore span V'.

Exercise 2.25*. Decompose the permutation representation of $\mathfrak{S}_4$ on (i) the vertices and (ii) the edges of the cube.

Exercise 2.26. The alternating group $\mathfrak{A}_4$ has four conjugacy classes. Three representations U, U', and U'' come from the representations of

$$\mathfrak{A}_4/\{1, (12)(34), (13)(24), (14)(23)\} \cong \mathbb{Z}/3,$$

so there is one more irreducible representation V of dimension 3. Compute the character table, with $\omega = e^{2\pi i/3}$:

	1	4	4	3
$\mathfrak{A}_4$	1	(123)	(132)	(12)(34)
U	1	1	1	1
U'	1	ω	ω^2	1
U''	1	ω^2	ω	1
V	3	0	0	-1

Exercise 2.27. Consider the representations of $\mathfrak{S}_4$ and their restrictions to $\mathfrak{A}_4$. Which are still irreducible when restricted, and which decompose? Which pairs of nonisomorphic representations of $\mathfrak{S}_4$ become isomorphic when restricted? Which representations of $\mathfrak{A}_4$ arise as restrictions from $\mathfrak{S}_4$?

§2.4. More Projection Formulas; More Consequences

In this section, we complete the analysis of the characters of the irreducible representations of a general finite group begun in §2.2 and give a more general formula for the projection of a general representation V onto the direct sum of the factors in V isomorphic to a given irreducible representation W. The main idea for both is a generalization of the "averaging" of the endomorphisms $g\colon V \to V$ used in §2.2, the point being that instead of simply averaging all the g we can ask the question: what linear combinations of the endomorphisms $g\colon V \to V$ are G-linear endomorphisms? The answer is given by

Proposition 2.28. *Let $\alpha\colon G \to \mathbb{C}$ be any function on the group G, and for any representation V of G set*

$$\varphi_{\alpha,V} = \sum \alpha(g)\cdot g\colon V \to V.$$

Then $\varphi_{\alpha,V}$ is a homomorphism of G-modules for all V if and only if α is a class function.

Proof. We simply write out the condition that $\varphi_{\alpha,V}$ be G-linear, and the result falls out: we have

$$\begin{aligned}\varphi_{\alpha,V}(hv) &= \sum \alpha(g)\cdot g(hv)\\ &= \sum \alpha(hgh^{-1})\cdot hgh^{-1}(hv)\end{aligned}$$

(substituting hgh^{-1} for g)

$$\begin{aligned}&= h\left(\sum \alpha(hgh^{-1})\cdot g(v)\right)\\ &= h\left(\sum \alpha(g)\cdot g(v)\right)\end{aligned}$$

(if α is a class function)

$$= h(\varphi_{\alpha,V}(v)).$$

Exercise 2.29*. Complete this proof by showing that conversely if α is not a class function, then there exists a representation V of G for which $\varphi_{\alpha,V}$ fails to be G-linear. □

As an immediate consequence of this proposition, we have

Proposition 2.30. *The number of irreducible representations of G is equal to the number of conjugacy classes of G. Equivalently, their characters $\{\chi_V\}$ form an orthonormal basis for $\mathbb{C}_{\text{class}}(G)$.*

PROOF. Suppose $\alpha: G \to \mathbb{C}$ is a class function and $(\alpha, \chi_V) = 0$ for all irreducible representations V; we must show that $\alpha = 0$. Consider the endomorphism

$$\varphi_{\alpha,V} = \sum \alpha(g)\cdot g: V \to V$$

as defined above. By Schur's lemma, $\varphi_{\alpha,V} = \lambda \cdot \mathrm{Id}$; and if $n = \dim V$, then

$$\begin{aligned}\lambda &= \frac{1}{n}\cdot \mathrm{trace}(\varphi_{\alpha,V})\\ &= \frac{1}{n}\cdot \sum \alpha(g)\chi_V(g)\\ &= \frac{|G|}{n}\overline{(\alpha, \chi_{V^*})}\\ &= 0.\end{aligned}$$

Thus, $\varphi_{\alpha,V} = 0$, or $\sum \alpha(g)\cdot g = 0$ on any representation V of G; in particular, this will be true for the regular representation $V = R$. But in R the elements $\{g \in G\}$, thought of as elements of $\mathrm{End}(R)$, are linearly independent. For example, the elements $\{g(e)\}$ are all independent. Thus $\alpha(g) = 0$ for all g, as required. □

This proposition completes the description of the characters of a finite group in general. We will see in more examples below how we can use this information to build up the character table of a given group. For now, we mention another way of expressing this proposition, via the *representation ring* of the group G.

The representation ring $R(G)$ of a group G is easy to define. First, as a group we just take $R(G)$ to be the free abelian group generated by all (isomorphism classes of) representations of G, and mod out by the subgroup generated by elements of the form $V + W - (V \oplus W)$. Equivalently, given the statement of complete reducibility, we can just take all integral linear combinations $\sum a_i \cdot V_i$ of the irreducible representations V_i of G; elements of $R(G)$ are correspondingly called *virtual representations*. The ring structure is then given simply by tensor product, defined on the generators of $R(G)$ and extended by linearity.

We can express most of what we have learned so far about representations of a finite group G in these terms. To begin, the character defines a map

$$\chi: R(G) \to \mathbb{C}_{\text{class}}(G)$$

from $R(G)$ to the ring of complex-valued functions on G; by the basic formulas of Proposition 2.1, this map is in fact a ring homomorphism. The statement that a representation in determined by its character then says that χ is injective;

the images of χ are called *virtual characters* and correspond thereby to virtual representations. Finally, our last proposition amounts to the statement that *χ induces an isomorphism*

$$\chi_{\mathbb{C}}: R(G) \otimes \mathbb{C} \to \mathbb{C}_{\text{class}}(G).$$

The virtual characters of G form a lattice $\Lambda \cong \mathbb{Z}^c$ in $\mathbb{C}_{\text{class}}(G)$, in which the actual characters sit as a cone $\Lambda_0 \cong \mathbb{N}^c \subset \mathbb{Z}^c$. We can thus think of the problem of describing the characters of G as having two parts: first, we have to find Λ, and then the cone $\Lambda_0 \subset \Lambda$ (once we know Λ_0, the characters of the irreducible representations will be determined). In the following lecture we will state theorems of Artin and Brauer characterizing $\Lambda \otimes \mathbb{Q}$ and Λ.

The argument for Proposition 2.30 also suggests how to obtain a more general projection formula. Explicitly, if W is a fixed irreducible representation, then for any representation V, look at the weighted sum

$$\psi = \frac{1}{|G|} \sum_{g \in G} \overline{\chi_W(g)} \cdot g \in \operatorname{End}(V).$$

By Proposition 2.28, ψ is a G-module homomorphism. Hence, if V is irreducible, we have $\psi = \lambda \cdot \mathrm{Id}$, and

$$\begin{aligned}\lambda &= \frac{1}{\dim V} \operatorname{Trace} \psi \\ &= \frac{1}{\dim V} \cdot \frac{1}{|G|} \sum \overline{\chi_W(g)} \cdot \chi_V(g) \\ &= \begin{cases} \dfrac{1}{\dim V} & \text{if } V = W \\ 0 & \text{if } V \neq W. \end{cases}\end{aligned}$$

For arbitrary V,

$$\psi_V = \dim W \cdot \frac{1}{|G|} \sum_{g \in G} \overline{\chi_W(g)} \cdot g : V \to V \tag{2.31}$$

is the projection of V onto the factor consisting of the sum of all copies of W appearing in V. In other words, if $V = \bigoplus V_i^{\oplus a_i}$, then

$$\pi_i = \dim V_i \cdot \frac{1}{|G|} \sum_{g \in G} \overline{\chi_{V_i}(g)} \cdot g \tag{2.32}$$

is the projection of V onto $V_i^{\oplus a_i}$.

Exercise 2.33*. (a) In terms of representations V and W in $R(G)$, the inner product on $\mathbb{C}_{\text{class}}(G)$ takes the simple form

$$(V, W) = \dim \operatorname{Hom}_G(V, W).$$

(b) If $\chi \in \mathbb{C}_{\text{class}}(G)$ is a virtual character, and $(\chi, \chi) = 1$, then either χ or $-\chi$ is the character of an irreducible representation, the plus sign occurring when $\chi(1) > 0$. If $(\chi, \chi) = 2$, and $\chi(1) > 0$, then χ is either the sum or the difference of two irreducible characters.

(c) If U, V, and W are irreducible representations, show that U appears in $V \otimes W$ if and only if W occurs in $V^* \otimes U$. Deduce that this cannot occur unless $\dim U \geq \dim W/\dim V$.

We conclude this lecture with some exercises that use characters to work out some standard facts about representations.

Exercise 2.34*. Let V and W be irreducible representations of G, and L_0: $V \to W$ any linear mapping. Define L: $V \to W$ by

$$L(v) = \frac{1}{|G|} \sum_{g \in G} g^{-1} \cdot L_0(g \cdot v).$$

Show that $L = 0$ if V and W are not isomorphic, and that L is multiplication by $\operatorname{trace}(L_0)/\dim(V)$ if $V = W$.

Exercise 2.35*. Show that, if the irreducible representations of G are represented by unitary matrices [cf. Exercise 1.14], the matrix entries of these representations form an orthogonal basis for the space of *all* functions on G [with inner product given by (2.11)].

Exercise 2.36*. If G_1 and G_2 are groups, and V_1 and V_2 are representations of G_1 and G_2, then the tensor product $V_1 \otimes V_2$ is a representation of $G_1 \times G_2$, by $(g_1 \times g_2) \cdot (v_1 \otimes v_2) = g_1 \cdot v_1 \otimes g_2 \cdot v_2$. To distinguish this "external" tensor product from the internal tensor product—when $G_1 = G_2$—this *external tensor product* is sometimes denoted $V_1 \boxtimes V_2$. If χ_i is the character of V_i, then the value of the character χ of $V_1 \boxtimes V_2$ is given by the product:

$$\chi(g_1 \times g_2) = \chi_1(g_1)\chi_2(g_2).$$

If V_1 and V_2 are irreducible, show that $V_1 \boxtimes V_2$ is also irreducible and show that every irreducible representation of $G_1 \times G_2$ arises this way. In terms of representation rings,

$$R(G_1 \times G_2) = R(G_1) \otimes R(G_2).$$

In these lectures we will often be given a subgroup G of a general linear group $\mathrm{GL}(V)$, and we will look for other representations inside tensor powers of V. The following problem, which is a theorem of Burnside and Molien, shows that for a finite group G, all irreducible representations can be found this way.

Problem 2.37*. Show that if V is a faithful representation of G, i.e., $\rho\colon G \to \mathrm{GL}(V)$ is injective, then any irreducible representation of G is contained in some tensor power $V^{\otimes n}$ of V.

Problem 2.38*. Show that the dimension of an irreducible representation of G divides the order of G.

Another challenge:

Problem 2.39*. Show that the character of any irreducible representation of dimension greater than 1 assumes the value 0 on some conjugacy class of the group.

LECTURE 3

Examples; Induced Representations; Group Algebras; Real Representations

This lecture is something of a grabbag. We start in §3.1 with examples illustrating the use of the techniques of the preceding lecture. Section 3.2 is also by way of an example. We will see quite a bit more about the representations of the symmetric groups in general later; §4 is devoted to this and will certainly subsume this discussion, but this should provide at least a sense of how we can go about analyzing representations of a class of groups, as opposed to individual groups. In §§3.3 and 3.4 we introduce two basic notions in representation theory, induced representations and the group algebra. Finally, in §3.5 we show how to classify representations of a finite group on a real vector space, given the answer to the corresponding question over $\mathbb{C}$, and say a few words about the analogous question for subfields of $\mathbb{C}$ other than $\mathbb{R}$. Everything in this lecture is elementary except Exercises 3.9 and 3.32, which involve the notions of Clifford algebras and the Fourier transform, respectively (both exercises, of course, can be skipped).

§3.1. Examples: $\mathfrak{S}_5$ and $\mathfrak{A}_5$

We have found the representations of the symmetric and alternating groups for $n \leq 4$. Before turning to a more systematic study of symmetric and alternating groups, we will work out the next couple of cases.

Representations of the Symmetric Group $\mathfrak{S}_5$

As before, we start by listing the conjugacy classes of $\mathfrak{S}_5$ and giving the number of elements of each: we have 10 transpositions, 20 three-cycles, 30 four-cycles and 24 five-cycles; in addition, we have 15 elements conjugate to (12)(34) and 10 elements conjugate to (12)(345). As for the irreducible representations, we have, of course, the trivial representation U, the alternating representation U', and the standard representation V; also, as in the case of $\mathfrak{S}_4$ we can tensor the standard representation V with the alternating one to obtain another irreducible representation V' with character $\chi_{V'} = \chi_V \cdot \chi_{U'}$.

Exercise 3.1. Find the characters of the representations V and V'; deduce in particular that V and V' are distinct irreducible representations.

The first four rows of the character table are thus

	1	10	20	30	24	15	20
$\mathfrak{S}_5$	1	(12)	(123)	(1234)	(12345)	(12)(34)	(12)(345)
U	1	1	1	1	1	1	1
U'	1	-1	1	-1	1	1	-1
V	4	2	1	0	-1	0	-1
V'	4	-2	1	0	-1	0	1

Clearly, we need three more irreducible representations. Where should we look for these? On the basis of our previous experience (and Problem 2.37), a natural place would be in the tensor products/powers of the irreducible representations we have found so far, in particular in $V \otimes V$ (the other two possible products will yield nothing new: we have $V' \otimes V = V \otimes V \otimes U'$ and $V' \otimes V' = V \otimes V$). Of course, $V \otimes V$ breaks up into $\wedge^2 V$ and $\mathrm{Sym}^2 V$, so we look at these separately. To start with, by the formula

$$\chi_{\wedge^2 V}(g) = \tfrac{1}{2}(\chi_V(g)^2 - \chi_V(g^2))$$

we calculate the character of $\wedge^2 V$:

$$\chi_{\wedge^2 V} = (6, 0, 0, 0, 1, -2, 0);$$

we see from this that it is indeed a fifth irreducible representation (and that $\wedge^2 V \otimes U' = \wedge^2 V$, so we get nothing new that way).

We can now find the remaining two representations in either of two ways. First, if n_1 and n_2 are their dimensions, we have

$$5! = 120 = 1^2 + 1^2 + 4^2 + 4^2 + 6^2 + n_1^2 + n_2^2,$$

so $n_1^2 + n_2^2 = 50$. There are no more one-dimensional representations, since these are trivial on normal subgroups whose quotient group is cyclic, and $\mathfrak{A}_5$

is the only such subgroup. So the only possibility is $n_1 = n_2 = 5$. Let W denote one of these five-dimensional representations, and set $W' = W \otimes U'$. In the table, if the row giving the character of W is

$$(5 \quad \alpha_1 \quad \alpha_2 \quad \alpha_3 \quad \alpha_4 \quad \alpha_5 \quad \alpha_6),$$

that of W' is $(5 \quad -\alpha_1 \quad \alpha_2 \quad -\alpha_3 \quad \alpha_4 \quad \alpha_5 \quad -\alpha_6)$. Using the orthogonality relations or (2.20), one sees that $W' \not\cong W$; and with a little calculation, up to interchanging W and W', the last two rows are as given:

	1	10	20	30	24	15	20
$\mathfrak{S}_5$	1	(12)	(123)	(1234)	(12345)	(12)(34)	(12)(345)
U	1	1	1	1	1	1	1
U'	1	-1	1	-1	1	1	-1
V	4	2	1	0	-1	0	-1
V'	4	-2	1	0	-1	0	1
$\wedge^2 V$	6	0	0	0	1	-2	0
W	5	1	-1	-1	0	1	1
W'	5	-1	-1	1	0	1	-1

From the decomposition $V \oplus U = \mathbb{C}^5$, we have also $\wedge^4 V = \wedge^5 \mathbb{C}^5 = U'$, and $V^* = V$. The perfect pairing[1]

$$V \times \wedge^3 V \to \wedge^4 V = U',$$

taking $v \times (v_1 \wedge v_2 \wedge v_3)$ to $v \wedge v_1 \wedge v_2 \wedge v_3$ shows that $\wedge^3 V$ is isomorphic to $V^* \otimes U' = V'$.

Another way to find the representations W and W' would be to proceed with our original plan, and look at the representation $\mathrm{Sym}^2 V$. We will leave this in the form of an exercise:

Exercise 3.2. (i) Find the character of the representation $\mathrm{Sym}^2 V$.

(ii) Without using any knowledge of the character table of $\mathfrak{S}_5$, use this to show that $\mathrm{Sym}^2 V$ is the direct sum of three distinct irreducible representations.

(iii) Using our knowledge of the first five rows of the character table, show that $\mathrm{Sym}^2 V$ is the direct sum of the representations U, V, and a third irreducible representation W. Complete the character table for $\mathfrak{S}_5$.

Exercise 3.3. Find the decomposition into irreducibles of the representations $\wedge^2 W$, $\mathrm{Sym}^2 W$, and $V \otimes W$.

[1] If V and W are n-dimensional vector spaces, and U is one dimensional, a *perfect pairing* is a bilinear map $\beta\colon V \times W \to U$ such that no nonzero vector v in V has $\beta(v, W) = 0$. Equivalently, the map $V \to \mathrm{Hom}(W, U) = W^* \otimes U$, $v \mapsto (w \mapsto \beta(v, w))$, is an isomorphism.

Representations of the Alternating Group $\mathfrak{A}_5$

What happens to the conjugacy classes above if we replace $\mathfrak{S}_d$ by $\mathfrak{A}_d$? Obviously, all the odd conjugacy classes disappear; but at the same time, since conjugation by a transposition is now an outer, rather than inner, automorphism, some conjugacy classes may break into two.

Exercise 3.4. Show that the conjugacy class in $\mathfrak{S}_d$ of permutations consisting of products of disjoint cycles of lengths $b_1, b_2, \ldots$ will break up into the union of two conjugacy classes in $\mathfrak{A}_d$ if all the b_k are odd and distinct; if any b_k are even or repeated, it remains a single conjugacy class in $\mathfrak{A}_d$. (We consider a fixed point as a cycle of length 1.)

In the case of $\mathfrak{A}_5$, this means we have the conjugacy class of three-cycles (as before, 20 elements), and of products of two disjoint transpositions (15 elements); the conjugacy class of five-cycles, however, breaks up into the conjugacy classes of (12345) and (21345), each having 12 elements.

As for the representations, the obvious first place to look is at restrictions to $\mathfrak{A}_5$ of the irreducible representations of $\mathfrak{S}_5$ found above. An irreducible representation of $\mathfrak{S}_5$ may become reducible when restricted to $\mathfrak{A}_5$; or two distinct representations may become isomorphic, as will be the case with U and U', V and V', or W and W'. In fact, U, V, and W stay irreducible since their characters satisfy $(\chi, \chi) = 1$. But the character of $\wedge^2 V$ has values $(6, 0, -2, 1, 1)$ on the conjugacy classes listed above, so $(\chi, \chi) = 2$, and $\wedge^2 V$ is the sum of two irreducible representations, which we denote by Y and Z. Since the sums of the squares of all the dimensions is 60, $(\dim Y)^2 + (\dim Z)^2 = 18$, so each must be three dimensional.

Exercise 3.5*. Use the orthogonality relations to complete the character table of $\mathfrak{A}_5$:

	1	20	15	12	12
$\mathfrak{A}_5$	1	(123)	(12)(34)	(12345)	(21345)
U	1	1	1	1	1
V	4	1	0	-1	-1
W	5	-1	1	0	0
Y	3	0	-1	$\frac{1+\sqrt{5}}{2}$	$\frac{1-\sqrt{5}}{2}$
Z	3	0	-1	$\frac{1-\sqrt{5}}{2}$	$\frac{1+\sqrt{5}}{2}$

The representations Y and Z may in fact be familiar: $\mathfrak{A}_5$ can be realized as the group of motions of an icosahedron (or, equivalently, of a dodecahedron) and Y is the corresponding representation. Note that the two representations $\mathfrak{A}_5 \to \mathrm{GL}_3(\mathbb{R})$ corresponding to Y and Z have the same image, but (as you can see from the fact that their characters differ only on the conjugacy classes of (12345) and (21345)) differ by an *outer* automorphism of $\mathfrak{A}_5$.

Note also that $\wedge^2 V$ does not decompose over $\mathbb{Q}$; we could see this directly from the fact that the vertices of a dodecahedron cannot all have rational coordinates, which follows from the analogous fact for a regular pentagon in the plane.

Exercise 3.6. Find the decomposition of the permutation representation of $\mathfrak{A}_5$ corresponding to the (i) vertices, (ii) faces, and (iii) edges of the icosahedron.

Exercise 3.7. Consider the dihedral group D_{2n}, defined to be the group of isometries of a regular n-gon in the plane. Let $\Gamma \cong \mathbb{Z}/n \subset D_{2n}$ be the subgroup of rotations. Use the methods of Lecture 1 (applied there to the case $\mathfrak{S}_3 \cong D_6$) to analyze the representations of D_{2n}: that is, restrict an arbitrary representation of D_{2n} to Γ, break it up into eigenspaces for the action of Γ, and ask how the remaining generator of D_{2n} acts of these eigenspaces.

Exercise 3.8. Analyze the representations of the dihedral group D_{2n} using the character theory developed in Lecture 2.

Exercise 3.9. (a) Find the character table of the group of order 8 consisting of the quaternions $\{\pm 1, \pm i, \pm j, \pm k\}$ under multiplication. This is the case $m = 3$ of a collection of groups of order 2^m, which we denote H_m. To describe them, let C_m denote the complex Clifford algebra generated by $v_1, \ldots, v_m$ with relations $v_i^2 = -1$ and $v_i \cdot v_j = -v_j \cdot v_i$, so C_m has a basis $v_I = v_{i_1} \cdot \ldots \cdot v_{i_r}$, as $I = \{i_i < \cdots < i_r\}$ varies over subsets of $\{1, \ldots, m\}$. (See §20.1 for notation and basic facts about Clifford algebras). Set

$$H_m = \{\pm v_I \colon |I| \text{ is even}\} \subset (C_m^{\text{even}})^*.$$

This group is a 2-to-1 covering of the abelian 2-group of $m \times m$ diagonal matrices with ± 1 diagonal entries and determinant 1. The center of H_m is $\{\pm 1\}$ if m is odd and is $\{\pm 1, \pm v_{\{1,\ldots,m\}}\}$ if m is even. The other conjugacy classes consist of pairs of elements $\{\pm v_I\}$. The isomorphisms of C_m^{even} with a matrix algebra or a product of two matrix algebras give a 2^n-dimensional "spin" representation S of H_{2n+1}, and two 2^{n-1}-dimensional "spin" or "half-spin" representations S^+ and S^- of H_{2n}.

(b) Compute the characters of these spin representations and verify that they are irreducible.

(c) Deduce that the spin representations, together with the 2^{m-1} one-dimensional representations coming from the abelian group $H_m/\{\pm 1\}$ give a complete set of irreducible representations, and compute the character table for H_m.

For odd m the groups H_m are examples of *extra-special* 2-*groups*, cf. [Grie], [Qu].

Exercise 3.10. Find the character table of the group $\mathrm{SL}_2(\mathbb{Z}/3)$.

Exercise 3.11. Let $H(\mathbb{Z}/3)$ be the *Heisenberg group* of order 27:

$$H(\mathbb{Z}/3) = \left\{ \begin{pmatrix} 1 & a & b \\ 0 & 1 & c \\ 0 & 0 & 1 \end{pmatrix}, a, b, c \in \mathbb{Z}/3 \right\} \subset \mathrm{SL}_3(\mathbb{Z}/3).$$

Analyze the representations of $H(\mathbb{Z}/3)$, first by the methods of Lecture 1 (restricting in this case to the center

$$Z = \left\{ \begin{pmatrix} 1 & 0 & b \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, b \in \mathbb{Z}/3 \right\} \cong \mathbb{Z}/3$$

of $H(\mathbb{Z}/3)$), and then by character theory.

§3.2. Exterior Powers of the Standard Representation of $\mathfrak{S}_d$

How should we go about constructing representations of the symmetric groups in general? The answer to this is not immediate; it is a subject that will occupy most of the next lecture (where we will produce all the irreducible representations of $\mathfrak{S}_d$). For now, as an example of the elementary techniques developed so far we will analyze directly one of the obvious candidates:

Proposition 3.12. *Each exterior power $\wedge^k V$ of the standard representation V of $\mathfrak{S}_d$ is irreducible*, $0 \le k \le d-1$.

Proof. From the decomposition $\mathbb{C}^d = V \oplus U$, we see that V is irreducible if and only if $(\chi_{\mathbb{C}^d}, \chi_{\mathbb{C}^d}) = 2$. Similarly, since

$$\wedge^k \mathbb{C}^d = (\wedge^k V \otimes \wedge^0 U) \oplus (\wedge^{k-1} V \otimes \wedge^1 U) = \wedge^k V \oplus \wedge^{k-1} V,$$

it suffices to show that $(\chi, \chi) = 2$, where χ is the character of the representation $\wedge^k \mathbb{C}^d$. Let $A = \{1, 2, \ldots, d\}$. For a subset B of A with k elements, and $g \in G = \mathfrak{S}_d$, let

$$\{g\}_B = \begin{cases} 0 & \text{if } g(B) \neq B \\ 1 & \text{if } g(B) = B \text{ and } g|_B \text{ is an even permutation} \\ -1 & \text{if } g(B) = B \text{ and } g|_B \text{ is odd.} \end{cases}$$

Here, if $g(B) = B$, $g|_B$ denotes the permutation of the set B determined by g. Then $\chi(g) = \sum \{g\}_B$, and

$$\begin{aligned}(\chi, \chi) &= \frac{1}{d!} \sum_{g \in G} \left(\sum_B \{g\}_B\right)^2 \\ &= \frac{1}{d!} \sum_{g \in G} \sum_B \sum_C \{g\}_B \{g\}_C \\ &= \frac{1}{d!} \sum_B \sum_C \sum_g (\operatorname{sgn} g|_B) \cdot (\operatorname{sgn} g|_C),\end{aligned}$$

where the sums are over subsets B and C of A with k elements, and in the last equation, the sum is over those g with $g(B) = B$ and $g(C) = C$. Such g is given by four permutations: one of $B \cap C$, one of $B \setminus B \cap C$, one of $C \setminus B \cap C$, and one of $A \setminus B \cup C$. Letting l be the cardinality of $B \cap C$, this last sum can be written

$$\begin{aligned}&\frac{1}{d!} \sum_B \sum_C \sum_{a \in \mathfrak{S}_l} \sum_{b \in \mathfrak{S}_{k-l}} \sum_{c \in \mathfrak{S}_{k-l}} \sum_{h \in \mathfrak{S}_{d-2k+l}} (\operatorname{sgn} a)^2 (\operatorname{sgn} b)(\operatorname{sgn} c) \\ &\quad = \frac{1}{d!} \sum_B \sum_C l!(d - 2k + l)! \left(\sum_{b \in \mathfrak{S}_{k-l}} \operatorname{sgn} b\right)\left(\sum_{c \in \mathfrak{S}_{k-l}} \operatorname{sgn} c\right).\end{aligned}$$

These last sums are zero unless $k - l = 0$ or 1. The case $k = l$ gives

$$\frac{1}{d!} \sum_B k!(d-k)! = \frac{1}{d!} \binom{d}{k} k!(d-k)! = 1.$$

Similarly, the terms with $k - l = 1$ also add up to 1, so $(\chi, \chi) = 2$, as required. $\square$

Note by way of contrast that the symmetric powers of the standard representation of $\mathfrak{S}_d$ are almost never irreducible. For example, we already know that the representation $\operatorname{Sym}^2 V$ contains one copy of the trivial representation: this is just the statement that every irreducible real representation (such as V) admits an inner product (unique, up to scalars) invariant under the group action; nor is the quotient of $\operatorname{Sym}^2 V$ by this trivial subrepresentation necessarily irreducible, as witness the case of $\mathfrak{S}_5$.

§3.3. Induced Representations

If $H \subset G$ is a subgroup, any representation V of G restricts to a representation of H, denoted $\operatorname{Res}_H^G V$ or simple Res V. In this section, we describe an important construction which produces representations of G from representations of H. Suppose V is a representation of G, and $W \subset V$ is a subspace which is H-invariant. For any g in G, the subspace $g \cdot W = \{g \cdot w: w \in W\}$ depends only on the left coset gH of g modulo H, since $gh \cdot W = g \cdot (h \cdot W) = g \cdot W$; for a coset

σ in G/H, we write $\sigma \cdot W$ for this subspace of V. We say that V is *induced* by W if every element in V can be written uniquely as a sum of elements in such translates of W, i.e.,

$$V = \bigoplus_{\sigma \in G/H} \sigma \cdot W.$$

In this case we write $V = \operatorname{Ind}_H^G W = \operatorname{Ind} W$.

Example 3.13. The permutation representation associated to the left action of G on G/H is induced from the trivial one-dimensional representation W of H. Here V has basis $\{e_\sigma \colon \sigma \in G/H\}$, and $W = \mathbb{C} \cdot e_H$, with H the trivial coset.

Example 3.14. The regular representation of G is induced from the regular representation of H. Here V has basis $\{e_g \colon g \in G\}$, whereas W has basis $\{e_h \colon h \in H\}$.

We claim that, given a representation W of H, such V exists and is unique up to isomorphism. Although we will later give several fancier ways to see this, it is not hard to do it by hand. Choose a representative $g_\sigma \in G$ for each coset $\sigma \in G/H$, with e representing the trivial coset H. To see the uniqueness, note that each element of V has a unique expression $v = \sum g_\sigma w_\sigma$, for elements w_σ in W. Given g in G, write $g \cdot g_\sigma = g_\tau \cdot h$ for some $\tau \in G/H$ and $h \in H$. Then we must have

$$g \cdot (g_\sigma w_\sigma) = (g \cdot g_\sigma) w_\sigma = (g_\tau \cdot h) w_\sigma = g_\tau (h w_\sigma).$$

This proves the uniqueness and tells us how to construct $V = \operatorname{Ind}(W)$ from W. Take a copy W^σ of W for each left coset $\sigma \in G/H$; for $w \in W$, let $g_\sigma w$ denote the element of W^σ corresponding to w in W. Let $V = \bigoplus_{\sigma \in G/H} W^\sigma$, so every element of V has a unique expression $v = \sum g_\sigma w_\sigma$ for elements w_σ in W. Given $g \in G$, define

$$g \cdot (g_\sigma w_\sigma) = g_\tau (h w_\sigma) \quad \text{if } g \cdot g_\sigma = g_\tau \cdot h.$$

To show that this defines as action of G on V, we must verify that $g' \cdot (g \cdot (g_\sigma w_\sigma)) = (g' \cdot g) \cdot (g_\sigma w_\sigma)$ for another element g' in G. Now if $g' \cdot g_\tau = g_\rho \cdot h'$, then

$$g' \cdot (g \cdot (g_\sigma w_\sigma)) = g' \cdot (g_\tau (h w_\sigma)) = g_\rho (h'(h w_\sigma)).$$

Since $(g' \cdot g) \cdot g_\sigma = g' \cdot (g \cdot g_\sigma) = g' \cdot g_\tau \cdot h = g_\rho \cdot h' \cdot h$, we have

$$(g' \cdot g) \cdot (g_\sigma w_\sigma) = g_\rho ((h' \cdot h) w_\sigma) = g_\rho (h' \cdot (h w_\sigma)),$$

as required.

Example 3.15. If $W = \bigoplus W_i$, then $\operatorname{Ind} W = \bigoplus \operatorname{Ind} W_i$.

The existence of the induced representation follows from Examples 3.14 and 3.15 since any W is a direct sum of summands of the regular representation.

Exercise 3.16. (a) If U is a representation of G and W a representation of H, show that (with all tensor products over $\mathbb{C}$)

$$U \otimes \operatorname{Ind} W = \operatorname{Ind}(\operatorname{Res}(U) \otimes W).$$

In particular, $\operatorname{Ind}(\operatorname{Res}(U)) = U \otimes P$, where P is the permutation representation of G on G/H. For a formula for $\operatorname{Res}(\operatorname{Ind}(W))$, for W a representation of H, see [Se2, p. 58].

(b) Like restriction, induction is transitive: if $H \subset K \subset G$ are subgroups, show that

$$\operatorname{Ind}_H^G(W) = \operatorname{Ind}_K^G(\operatorname{Ind}_H^K W).$$

Note that Example 3.15 says that the map Ind gives a group homomorphism between the representation rings $R(H)$ and $R(G)$, in the opposite direction from the ring homomorphism Res: $R(G) \to R(H)$ given by restriction; Exercise 3.16(a) says that this map satisfies a "push–pull" formula $\alpha \cdot \operatorname{Ind}(\beta) = \operatorname{Ind}(\operatorname{Res}(\alpha) \cdot \beta)$ with respect to the restriction map.

Proposition 3.17. *Let W be a representation of H, U a representation of G, and suppose $V = \operatorname{Ind} W$. Then any H-module homomorphism $\varphi\colon W \to U$ extends uniquely to a G-module homomorphism $\tilde{\varphi}\colon V \to U$. i.e.,*

$$\operatorname{Hom}_H(W, \operatorname{Res} U) = \operatorname{Hom}_G(\operatorname{Ind} W, U).$$

In particular, this universal property determines Ind W *up to canonical isomorphism.*

Proof. With $V = \bigoplus_{\sigma \in G/H} \sigma \cdot W$ as before, define $\tilde{\varphi}$ on $\sigma \cdot W$ by

$$\sigma \cdot W \xrightarrow{g_\sigma^{-1}} W \xrightarrow{\varphi} U \xrightarrow{g_\sigma} U,$$

which is independent of the representative g_σ for σ since φ is H-linear. □

To compute the character of $V = \operatorname{Ind} W$, note that $g \in G$ maps σW to $g\sigma W$, so the trace is calculated from those cosets σ with $g\sigma = \sigma$, i.e., $s^{-1}gs \in H$ for $s \in \sigma$. Therefore,

$$\chi_{\operatorname{Ind} W}(g) = \sum_{g\sigma=\sigma} \chi_W(s^{-1}gs) \qquad (s \in \sigma \text{ arbitrary}). \tag{3.18}$$

Exercise 3.19. (a) If C is a conjugacy class of G, and $C \cap H$ decomposes into conjugacy classes $D_1, \ldots, D_r$ of H, (3.18) can be rewritten as: the value of the character of Ind W on C is

$$\chi_{\operatorname{Ind} W}(C) = \frac{|G|}{|H|} \sum_{i=1}^{r} \frac{|D_i|}{|C|} \chi_W(D_i).$$

(b) If W is the trivial representation of H, then

$$\chi_{\operatorname{Ind} W}(C) = \frac{[G:H]}{|C|} \cdot |C \cap H|.$$

Corollary 3.20 (Frobenius Reciprocity). *If W is a representation of H, and U a representation of G, then*

$$(\chi_{\operatorname{Ind} W}, \chi_U)_G = (\chi_W, \chi_{\operatorname{Res} U})_H.$$

PROOF. It suffices by linearity to prove this when W and U are irreducible. The left-hand side is the number of times U appears in $\operatorname{Ind} W$, which is the dimension of $\operatorname{Hom}_G(\operatorname{Ind} W, U)$. The right-hand side is the dimension of $\operatorname{Hom}_H(W, \operatorname{Res} U)$. These dimensions are equal by the proposition. □

If W and U are irreducible, Frobenius reciprocity says: *the number of times U appears in* Ind W *is the same as the number of times W appears in* Res U.

Frobenius reciprocity can be used to find characters of G if characters of H are known.

Example 3.21. We compute $\operatorname{Ind}_H^G W$, when $H = \mathfrak{S}_2 \subset G = \mathfrak{S}_3$, $W = V_2$ (the standard representation) $= U_2'$ (the alternating representation). We know the irreducible represenatations of $\mathfrak{S}_3$: U_3, U_3', V_3, which restrict to U_2, $U_2' = V_2$, $U_2 \oplus U_2'$, respectively. Thus, by Frobenius, $\operatorname{Ind} V_2 = U_3' \oplus V_3$.

Example 3.22. Consider next $H = \mathfrak{S}_3 \subset G = \mathfrak{S}_4$, $W = V_3$. Again we know the irreducible representations, and $\operatorname{Res} U_4 = U_3$, $\operatorname{Res} U_4' = U_3'$, $\operatorname{Res} V_4 = U_3 \oplus V_3$ [the vector

$$(1, 1, 1, -3) \in V_4 = \{(x_1, x_2, x_3, x_4) : \textstyle\sum x_i = 0\}$$

is fixed by H], $\operatorname{Res} V_4' = U_3' \oplus V_3'$, with $V_3' = V_3$, and $\operatorname{Res} W_4 = V_3$ (as one may see directly). Hence, $\operatorname{Ind} V_3 = V_4 \oplus V_4' \oplus W_4$. (Note that the isomorphism $\operatorname{Res} W_4 = V_3$ actually follows, since one W_4 is all that could be added to $V_4 \oplus V_4'$ to get $\operatorname{Ind} V_3$.)

Exercise 3.23. Determine the isomorphism classes of the representations of $\mathfrak{S}_4$ induced by (i) the one-dimensional representation of the group generated by (1234) in which $(1234)\cdot v = iv$, $i = \sqrt{-1}$; (ii) the one-dimensional representation of the group generated by (123) in which $(123)\cdot v = e^{2\pi i/3}v$.

Exercise 3.24. Let $H = \mathfrak{A}_5 \subset G = \mathfrak{S}_5$. Show that $\operatorname{Ind} U = U \oplus U'$, $\operatorname{Ind} V = V \oplus V'$, and $\operatorname{Ind} W = W \oplus W'$, whereas $\operatorname{Ind} Y = \operatorname{Ind} Z = \wedge^2 V$.

Exercise 3.25*. Which irreducible representations of $\mathfrak{S}_d$ remain irreducible when restricted to $\mathfrak{A}_d$? Which are induced from $\mathfrak{A}_d$? How much does this tell you about the irreducible representations of $\mathfrak{A}_d$?

Exercise 3.26*. There is a unique nonabelian group of order 21, which can be realized as the group of affine transformations $x \mapsto \alpha x + \beta$ of the line over the field with seven elements, with α a cube root of unity in that field. Find the irreducible representations and character table for this group.

Now that we have introduced the notion of induced representation, we can state two important theorems describing the characters of representations of a finite group. In the preceding lecture we mentioned the notion of *virtual character*; this is just an element of the image Λ of the character map

$$\chi\colon R(G) \to \mathbb{C}_{\text{class}}(G)$$

from the representation ring $R(G)$ of virtual representations. The following two theorems both state that in order to generate $\Lambda \otimes \mathbb{Q}$ (resp. Λ) it is enough to consider the simplest kind of induced representations, namely, those induced from cyclic (respective elementary) subgroups of G. For the proofs of these theorems we refer to [Se2, §9, 10]. We will not need them in these lectures.

Artin's Theorem 3.27. *The characters of induced representations from cyclic subgroups of G generate a lattice of finite index in Λ.*

A subgroup H of G is *p-elementary* if $H = A \times B$, with A cyclic of order prime to p and B a p-group.

Brauer's Theorem 3.28. *The characters of induced representations from elementary subgroups of G generate the lattice Λ.*

§3.4. The Group Algebra

There is an important notion that we have already dealt with implicitly but not explicitly; this is the group algebra $\mathbb{C}G$ associated to a finite group G. This is an object that for all intents and purposes can completely replace the group G itself; any statement about the representations of G has an exact equivalent statement about the group algebra. Indeed, to a large extent the choice of language is a matter of taste.

The underlying vector space of the group algebra of G is the vector space with basis $\{e_g\}$ corresponding to elements of the group G, that is, the underlying vector space of the regular representation. We define the algebra structure on this vector space simply by

$$e_g \cdot e_h = e_{gh}.$$

By a representation of the algebra $\mathbb{C}G$ on a vector space V we mean simply an algebra homomorphism

$$\mathbb{C}G \to \operatorname{End}(V),$$

so that a representation V of $\mathbb{C}G$ is the same thing as a left $\mathbb{C}G$-module. Note that a representation $\rho\colon G \to \operatorname{Aut}(V)$ will extend by linearity to a map $\tilde{\rho}\colon \mathbb{C}G \to \operatorname{End}(V)$, so that representations of $\mathbb{C}G$ correspond exactly to representations of G; the left $\mathbb{C}G$-module given by $\mathbb{C}G$ itself corresponds to the regular representation.

If $\{W_i\}$ are the irreducible representations of G, then we have seen that the regular representation R decomposes

$$R = \bigoplus (W_i)^{\oplus \dim(W_i)}.$$

We can now refine this statement in terms of the group algebra: we have

Proposition 3.29. *As algebras,*

$$\mathbb{C}G \cong \bigoplus \operatorname{End}(W_i).$$

Proof. As we have said, for any representation W of G, the map $G \to \operatorname{Aut}(W)$ extends by linearity to a map $\mathbb{C}G \to \operatorname{End}(W)$; applying this to each of the irreducible representations W_i gives us a canonical map

$$\varphi\colon \mathbb{C}G \to \bigoplus \operatorname{End}(W_i).$$

This is injective since the representation on the regular representation is faithful. Since both have dimension $\sum(\dim W_i)^2$, the map is an isomorphism. □

A few remarks are in order about the isomorphism φ of the proposition. First, φ can be interpreted as the Fourier transform, cf. Exercise 3.32. Note also that Proposition 2.28 has a natural interpretation in terms of the group algebra: it says that the center of $\mathbb{C}G$ consists of those $\sum \alpha(g)e_g$ for which α is a class function.

Next, we can thinkof φ as the decomposition of the semisimple algebra $\mathbb{C}G$ into a product of matrix algebras. It implies that the matrix entries of the irreducible representations give a basis for the space of all functions on G, cf. Exercise 2.35.

Note in particular that any irreducible representation is isomorphic to a (minimal) left ideal in $\mathbb{C}G$. These left ideals are generated by idempotents. In fact, we can interpret the projection formulas of the last lecture in the language of the group algebra: the formulas say simply that the elements

$$\dim W \cdot \frac{1}{|G|} \sum_{g \in G} \overline{\chi_W(g)} \cdot e_g \in \mathbb{C}G$$

are the idempotents in the group algebra corresponding to the direct sum factors in the decomposition of Proposition 3.29. To locate the irreducible representations W_i of a group G [not just a direct sum of $\dim(W_i)$ copies], we want to find other idempotents of $\mathbb{C}G$. We will see this carried out for the symmetric groups in the following lecture.

The group algebra also gives us another description of induced representations: if W is a representation of a subgroup H of G, then the induced representation may be constructed simply by

$$\operatorname{Ind} W = \mathbb{C}G \otimes_{\mathbb{C}H} W,$$

where G acts on the first factor: $g \cdot (e_{g'} \otimes w) = e_{gg'} \otimes w$. The isomorphism of the reciprocity theorem is then a special case of a general formula for a change of rings $\mathbb{C}H \to \mathbb{C}G$:

$$\mathrm{Hom}_{\mathbb{C}H}(W, U) = \mathrm{Hom}_{\mathbb{C}G}(\mathbb{C}G \otimes_{\mathbb{C}H} W, U).$$

Exercise 3.30*. The induced representation $\mathrm{Ind}(W)$ can also be realized concretely as a space of W-valued functions on G, which can be useful to produce matrix realizations, or when trying to decompose $\mathrm{Ind}(W)$ into irreducible pieces. Show that $\mathrm{Ind}(W)$ is isomorphic to

$$\mathrm{Hom}_H(\mathbb{C}G, W) \cong \{f\colon G \to W : f(hg) = hf(g), \quad \forall h \in H, g \in G\},$$

where G acts by $(g' \cdot f)(g) = f(gg')$.

Exercise 3.31. If $\mathbb{C}G$ is identified with the space of functions on G, the function φ corresponding to $\sum_{g \in G} \varphi(g) e_g$, show that the product in $\mathbb{C}G$ corresponds to the convolution $*$ of functions:

$$(\varphi * \psi)(g) = \sum_{h \in G} \varphi(h)\psi(h^{-1}g).$$

(With integration replacing summation, this indicates how one may extend the notion of regular representation to compact groups.)

Exercise 3.32*. If $\rho\colon G \to \mathrm{GL}(V_\rho)$ is a representation, and φ is a function on G, define the *Fourier transform* $\hat{\varphi}(\rho)$ in $\mathrm{End}(V_\rho)$ by the formula

$$\hat{\varphi}(\rho) = \sum_{g \in G} \varphi(g) \cdot \rho(g).$$

(a) Show that $\widehat{\varphi * \psi}(\rho) = \widehat{\varphi}(\rho) \cdot \widehat{\psi}(\rho)$.

(b) Prove the *Fourier inversion formula*

$$\varphi(g) = \frac{1}{|G|} \sum \dim(V_\rho) \cdot \mathrm{Trace}(\rho(g^{-1}) \cdot \hat{\varphi}(\rho)),$$

the sum over the irreducible representations ρ of G. This formula is equivalent to formulas (2.19) and (2.20).

(c) Prove the *Plancherel formula* for functions φ and ψ on G:

$$\sum_{g \in G} \varphi(g^{-1})\psi(g) = \frac{1}{|G|} \sum_{\rho} \dim(V_\rho) \cdot \mathrm{Trace}(\hat{\varphi}(\rho)\hat{\psi}(\rho)).$$

Our choice of left action of a group on a space has been perfectly arbitrary, and the entire story is the same if G acts on the *right* instead. Moreover, there is a standard way to change a right action into a left action, and vice versa: Given a right action of G on V, define the left action by

$$g \cdot v = v \cdot (g^{-1}), \qquad g \in G, v \in V.$$

If $A = \mathbb{C}G$ is the group algebra, a right action of G on V makes V a right A-module. To turn right modules into left modules, we can use the anti-involution $a \mapsto \hat{a}$ of A defined by $(\sum a_g e_g)^\wedge = \sum a_g e_{g^{-1}}$. A right A-module is then turned into a left A-module by setting $a \cdot v = v \cdot \hat{a}$.

The following exercise will take you back to the origins of representation theory in the 19th century, when Frobenius found the characters by factoring this determinant.

Exercise 3.33*. Given a finite group G of order n, take a variable x_g for each element g in G, and order the elements of G arbitrarily. Let F be the determinant of the $n \times n$ matrix whose entry in the row labeled by g and column labeled by h is $x_{g \cdot h^{-1}}$. This is a form of degree n in the n variables x_g, which is independent of the ordering. Normalize the factors of F to take the value 1 when $x_e = 1$ and $x_g = 0$ for $g \neq e$. Show that the irreducible factors of F correspond to the irreducible representations of G. Moreover, if F_ρ is the factor corresponding to the representation ρ, show that the degree of F_ρ is the degree $d(\rho)$ of the representation ρ, and that each F_ρ occurs in F $d(\rho)$ times. If χ_ρ is the character of ρ, and $g \neq e$, show that $\chi_\rho(g)$ is the coefficient of $x_g \cdot x_e^{d(\rho)-1}$ in F_ρ.

§3.5. Real Representations and Representations over Subfields of $\mathbb{C}$

If a group G acts on a real vector space V_0, then we say the corresponding complex representation of $V = V_0 \otimes_{\mathbb{R}} \mathbb{C}$ is *real*. To the extent that we are interested in the action of a group G on real rather than complex vector spaces, the problem we face is to say which of the complex representations of G we have studied are in fact real.

Our first guess might be that a representation is real if and only if its character is real-valued. This turns out not to be the case: the character of a real representation is certainly real-valued, but the converse need not be true. To find an example, suppose $G \subset \mathrm{SU}(2)$ is a finite, nonabelian subgroup. Then G acts on $\mathbb{C}^2 = V$ with a real-valued character since the trace of any matrix in $\mathrm{SU}(2)$ is real. If V were a real representation, however, then G would be a subgroup of $\mathrm{SO}(2) = S^1$, which is abelian. To produce such a group, note that $\mathrm{SU}(2)$ can be identified with the unit quaternions. Set $G = \{\pm 1, \pm i, \pm j, \pm k\}$. Then $G/\{\pm 1\}$ is abelian, so has four one-dimensional representations, which give four one-dimensional representations of G. Thus, G has one irreducible two-dimensional representation, whose character is real, but which is not real.

Exercise 3.34*. Compute the character table for this quaternion group G, and compare it with the character table of the dihedral group of order 8.

A more successful approach is to note that if V is a real representation of G, coming from V_0 as above, then one can find a positive definite symmetric bilinear form on V_0 which is preserved by G. This gives a symmetric bilinear form on V which is preserved by G. Not every representation will have such a form since degeneracies may arise when one tries to construct one following the construction of Proposition 1.5. In fact,

Lemma 3.35. *An irreducible representation V of G is real if and only if there is a nondegenerate symmetric bilinear form B on V preserved by G.*

PROOF. If we have such B, and an arbitrary nondegenerate Hermitian form H, also G-invariant, then

$$V \xrightarrow{B} V^* \xrightarrow{H} V$$

gives a conjugate linear isomorphism φ from V to V: given $x \in V$, there is a unique $\varphi(x) \in V$ with $B(x, y) = H(\varphi(x), y)$, and φ commutes with the action of G. Then $\varphi^2 = \varphi \circ \varphi$ is a complex linear G-module homomorphism, so $\varphi^2 = \lambda \cdot \mathrm{Id}$. Moreover,

$$H(\varphi(x), y) = B(x, y) = B(y, x) = H(\varphi(y), x) = \overline{H(x, \varphi(y))},$$

from which it follows that $H(\varphi^2(x), y) = H(x, \varphi^2(y))$, and therefore λ is a positive real number. Changing H by a scalar, we may assume $\lambda = 1$, so $\varphi^2 = \mathrm{Id}$. Thus, V is a sum of real eigenspaces V_+ and V_- for φ corresponding to eigenvalues 1 and -1. Since φ commutes with G, V_+ and V_- are G-invariant subspaces. Finally, $\varphi(ix) = -i\varphi(x)$, so $iV_+ = V_-$, and $V = V_+ \otimes \mathbb{C}$. □

Note from the proof that a real representation is also characterized by the existence of a conjugate linear endomorphism of V whose square is the identity; if $V = V_0 \otimes_{\mathbb{R}} \mathbb{C}$, it is given by conjugation: $v_0 \otimes \lambda \mapsto v_0 \otimes \bar{\lambda}$.

A warning is in order here: an irreducible representation of G on a vector space over $\mathbb{R}$ may become reducible when we extend the group field to $\mathbb{C}$. To give the simplest example, the representation of $\mathbb{Z}/n$ on $\mathbb{R}^2$ given by

$$\rho\colon k \mapsto \begin{pmatrix} \cos\dfrac{2\pi k}{n} & -\sin\dfrac{2\pi k}{n} \\ \sin\dfrac{2\pi k}{n} & \cos\dfrac{2\pi k}{n} \end{pmatrix}$$

is irreducible over $\mathbb{R}$ for $n > 2$ (no line in $\mathbb{R}^2$ is fixed by the action of $\mathbb{Z}/n$), but will be reducible over $\mathbb{C}$. Thus, classifying the irreducible representations of G over $\mathbb{C}$ that are real does not mean that we have classified all the irreducible real representations. However, we will see in Exercise 3.39 below how to finish the story once we have found the real representations of G that are irreducible over $\mathbb{C}$.

Suppose V is an irreducible representation of G with χ_V real. Then there is a G-equivariant isomorphism $V \cong V^*$, i.e., there is a G-equivariant (nondegenerate) bilinear form B on V; but, in general, B need not be symmetric. Regarding B in

$$V^* \otimes V^* = \mathrm{Sym}^2 V^* \oplus \wedge^2 V^*,$$

and noting the uniqueness of B up to multiplication by scalars, we see that B is either symmetric or skew-symmetric. If B is skew-symmetric, proceeding as above one can scale so $\varphi^2 = -\mathrm{Id}$. This makes V "quaternionic," with φ becoming multiplication[2] by j:

Definition 3.36. A *quaternionic* representation is a (complex) representation V which has a G-invariant homomorphism $J\colon V \to V$ that is conjugate linear, and satisfies $J^2 = -\mathrm{Id}$. Thus, a skew-symmetric nondegenerate G-invariant B determines a quaternionic structure on V.

Summarizing the preceding discussion we have the

Theoem 3.37. *An irreducible representation V is one and only one of the following*:

(1) Complex: *χ_V is not real-valued; V does not have a G-invariant nondegenerate bilinear form.*

(2) Real: *$V = V_0 \otimes \mathbb{C}$, a real representation; V has a G-invariant symmetric nondegenerate bilinear form.*

(3) Quaternionic: *χ_V is real, but V is not real; V has a G-invariant skew-symmetric nondegeneate bilinear form.*

Exercise 3.38. Show that for V irreducible,

$$\frac{1}{|G|}\sum_{g \in G} \chi_V(g^2) = \begin{cases} 0 & \text{if } V \text{ is complex} \\ 1 & \text{if } V \text{ is real} \\ -1 & \text{if } V \text{ is quaternionic.} \end{cases}$$

This verifies that the three cases in the theorem are mutually exclusive. It also implies that if the order of G is odd, all nontrivial representations must be complex.

Exercise 3.39. Let V_0 be a real vector space on which G acts irreducibly, $V = V_0 \otimes \mathbb{C}$ the corresponding real representation of G. Show that if V is not irreducible, then it has exactly two irreducible factors, and they are conjugate complex representations of G.

[2] See §7.2 for more on quaternions and quaternonic representations.

Exercise 3.40. Classify the real representations of $\mathfrak{A}_4$.

Exercise 3.41*. The group algebra $\mathbb{R}G$ is a product of simple $\mathbb{R}$-algebras corresponding to the irreducible representations over $\mathbb{R}$. These simple algebras are matrix algebras over $\mathbb{C}$, $\mathbb{R}$, or the quaternions $\mathbb{H}$ according as the representation is complex, real, or quaternionic.

Exercise 3.42*. (a) Show that all characters of a group are real if and only if every element is conjugate to its inverse.

(b) Show that an element σ in a split conjugacy class of $\mathfrak{A}_d$ is conjugate to its inverse if and only if the number of cycles in σ whose length is congruent to 3 modulo 4 is even.

(c) Show that the only d's for which every character of $\mathfrak{A}_d$ is real-valued are $d = 1, 2, 5, 6, 10$, and 14.

Exercise 3.43*. Show that: (i) the tensor product of two real or two quaternionic representations is real; (ii) for any V, $V^* \otimes V$ is real; (iii) if V is real, so are all $\bigwedge^k V$; (iv) if V is quaternionic, $\bigwedge^k V$ is real for k even, quaternionic for k odd.

Representations over Subfields of $\mathbb{C}$ in General

We consider next the generalization of the preceding problem to more general subfields of $\mathbb{C}$. Unfortunately, our results will not be nearly as strong in general, but we can at least express the problem neatly in terms of the representation ring of G.

To begin with, our terminology in this general setting is a little different. Let $K \subset \mathbb{C}$ be any subfield. We define a K-representation of G to be a vector space V_0 over K on which G acts; in this case we say that the complex representation $V = V_0 \otimes \mathbb{C}$ is *defined over K*.

One way to measure how many of the representations of G are defined over a field K is to introduce the *representation ring $R_K(G)$ of G over K*. This is defined just like the ordinary representation ring; that is, it is just the group of formal linear combinations of K-representations of G modulo relations of the form $V + W - (V \oplus W)$, with multiplication given by tensor product.

Exercise 3.44*. Describe the representation ring of G over $\mathbb{R}$ for some of the groups G whose complex representation we have analyzed above. In particular, is the rank of $R_{\mathbb{R}}(G)$ always the same as the rank of $R(G)$?

Exercise 3.45*. (a) Show that $R_K(G)$ is the subring of the ring of class functions on G generated (as an additive group) by characters of representations defined over K.

(b) Show that the characters of irreducible representations over K form an orthogonal basis for $R_K(G)$.

(c) Show that a complex representation of G can be defined over K if and only if its character belongs to $R_K(G)$.

For more on the relation between $R_K(G)$ and $R(G)$, see [Se2].

LECTURE 4

Representations of $\mathfrak{S}_d$: Young Diagrams and Frobenius's Character Formula

In this lecture we get to work. Specifically, we give in §4.1 a complete description of the irreducible representations of the symmetric group, that is, a construction of the representations (via Young symmetrizers) and a formula (Frobenius' formula) for their characters. The proof that the representations constructed in §4.1 are indeed the irreducible representations of the symmetric group is given in §4.2; the proof of Frobenius' formula, as well as a number of others, in §4.3. Apart from their intrinsic interest (and undeniable beauty), these results turn out to be of substantial interest in Lie theory: analogs of the Young symmetrizers will give a construction of the irreducible representations of $\mathrm{SL}_n\mathbb{C}$. At the same time, while the techniques of this lecture are completely elementary (we use only a few identities about symmetric polynomials, proved in Appendix A), the level of difficulty is clearly higher than in preceding lectures. The results in the latter half of §4.3 (from Corollary 4.39 on) in particular are quite difficult, and inasmuch as they are not used later in the text may be skipped by readers who are not symmetric group enthusiasts.

§4.1. Statements of the Results

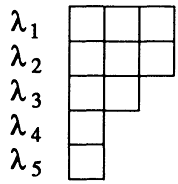
The number of irreducible representaton of $\mathfrak{S}_d$ is the number of conjugacy classes, which is the number $p(d)$ of partitions[1] of d: $d = \lambda_1 + \cdots + \lambda_k$, $\lambda_1 \geq \cdots \geq \lambda_k \geq 1$. We have

[1] It is sometimes convenient, and sometimes a nuisance, to have partitions that end in one or more zeros; if convenient, we allow some of the λ_i on the end to be zero. Two sequences define the same partition, of course, if they differ only by zeros at the end.

$$\sum_{d=0}^{\infty} p(d)t^d = \prod_{n=1}^{\infty}\left(\frac{1}{1-t^n}\right)$$
$$= (1 + t + t^2 + \cdots)(1 + t^2 + t^4 + \cdots)(1 + t^3 + \cdots)\cdots$$

which converges exactly in $|t| < 1$. This partition number is an interesting arithmetic function, whose congruences and growth behavior as a function of d have been much studied (cf. [Har], [And]). For example, $p(d)$ is asymptotically equal to $(1/\alpha d)e^{\beta\sqrt{d}}$, with $\alpha = 4\sqrt{3}$ and $\beta = \pi\sqrt{2/3}$.

To a partition $\lambda = (\lambda_1, \ldots, \lambda_k)$ is associated a *Young diagram* (sometimes called a Young frame or Ferrers diagram)



with λ_i boxes in the ith row, the rows of boxes lined up on the left. The *conjugate partition* $\lambda' = (\lambda'_1, \ldots, \lambda'_r)$ to the partition λ is defined by interchanging rows and columns in the Young diagram, i.e., reflecting the diagram in the 45° line. For example, the diagram above is that of the partition (3, 3, 2, 1, 1), whose conjugate is (5, 3, 2). (Without reference to the diagram, the conjugate partition to λ can be defined by saying λ'_i is the number of terms in the partition λ that are greater than or equal to i.)

Young diagrams can be used to describe projection operators for the regular representation, which will then give the irreducible representations of $\mathfrak{S}_d$. For a given Young diagram, number the boxes, say consecutively as shown:

1	2	3
4	5	
6	7	
8		

More generally, define a *tableau* on a given Young diagram to be a numbering of the boxes by the integers $1, \ldots, d$. Given a tableau, say the canonical one shown, define two subgroups[2] of the symmetric group

[2] If a tableau other than the canonical one were chosen, one would get different groups in place of P and Q, and different elements in the group ring, but the representations constructed this way will be isomorphic.

$$P = P_\lambda = \{g \in \mathfrak{S}_d: g \text{ preserves each row}\}$$

and

$$Q = Q_\lambda = \{g \in \mathfrak{S}_d: g \text{ preserves each column}\}.$$

In the group algebra $\mathbb{C}\mathfrak{S}_d$, we introduce two elements corresponding to these subgroups: we set

$$a_\lambda = \sum_{g \in P} e_g \quad \text{and} \quad b_\lambda = \sum_{g \in Q} \operatorname{sgn}(g) \cdot e_g. \tag{4.1}$$

To see what a_λ and b_λ do, observe that if V is any vector space and $\mathfrak{S}_d$ acts on the dth tensor power $V^{\otimes d}$ by permuting factors, the image of the element $a_\lambda \in \mathbb{C}\mathfrak{S}_d \to \operatorname{End}(V^{\otimes d})$ is just the subspace

$$\operatorname{Im}(a_\lambda) = \operatorname{Sym}^{\lambda_1} V \otimes \operatorname{Sym}^{\lambda_2} V \otimes \cdots \otimes \operatorname{Sym}^{\lambda_k} V \subset V^{\otimes d},$$

where the inclusion on the right is obtained by grouping the factors of $V^{\otimes d}$ according to the rows of the Young tableaux. Similarly, the image of b_λ on this tensor power is

$$\operatorname{Im}(b_\lambda) = \textstyle\bigwedge^{\mu_1} V \otimes \bigwedge^{\mu_2} V \otimes \cdots \otimes \bigwedge^{\mu_l} V \subset V^{\otimes d},$$

where μ is the conjugate partition to λ.

Finally, we set

$$c_\lambda = a_\lambda \cdot b_\lambda \in \mathbb{C}\mathfrak{S}_d; \tag{4.2}$$

this is called a *Young symmetrizer*. For example, when $\lambda = (d)$, $c_{(d)} = a_{(d)} = \sum_{g \in \mathfrak{S}_d} e_g$, and the image of $c_{(d)}$ on $V^{\otimes d}$ is $\operatorname{Sym}^d V$. When $\lambda = (1, \ldots, 1)$, $c_{(1,\ldots,1)} = b_{(1,\ldots,1)} = \sum_{g \in \mathfrak{S}_d} \operatorname{sgn}(g) e_g$, and the image of $c_{(1,\ldots,1)}$ on $V^{\otimes d}$ is $\bigwedge^d V$. We will eventually see that the image of the symmetrizers c_λ in $V^{\otimes d}$ provide essentially all the finite-dimensional irreducible representations of $\operatorname{GL}(V)$. Here we state the corresponding fact for representations of $\mathfrak{S}_d$:

Theorem 4.3. *Some scalar multiple of c_λ is idempotent, i.e., $c_\lambda^2 = n_\lambda c_\lambda$, and the image of c_λ (by right multiplication on $\mathbb{C}\mathfrak{S}_d$) is an irreducible representation V_λ of $\mathfrak{S}_d$. Every irreducible representation of $\mathfrak{S}_d$ can be obtained in this way for a unique partition.*

We will prove this theorem in the next section. Note that, as a corollary, each irreducible representation of $\mathfrak{S}_d$ can be defined over the rational numbers since c_λ is in the rational group algebra $\mathbb{Q}\mathfrak{S}_d$. Note also that the theorem gives a direct correspondence between conjugacy classes in $\mathfrak{S}_d$ and irreducible representations of $\mathfrak{S}_d$, something which has never been achieved for general groups.

For example, for $\lambda = (d)$,

$$V_{(d)} = \mathbb{C}\mathfrak{S}_d \cdot \sum_{g \in \mathfrak{S}_d} e_g = \mathbb{C} \cdot \sum_{g \in \mathfrak{S}_d} e_g$$

is the trivial representation U, and when $\lambda = (1, \ldots, 1)$,

$$V_{(1,\ldots,1)} = \mathbb{C}\mathfrak{S}_d \cdot \sum_{g \in \mathfrak{S}_d} \operatorname{sgn}(g) e_g = \mathbb{C} \cdot \sum_{g \in \mathfrak{S}_d} \operatorname{sgn}(g) e_g$$

is the alternating representation U'. For $\lambda = (2, 1)$,

$$c_{(2,1)} = (e_1 + e_{(12)}) \cdot (e_1 - e_{(13)}) = 1 + e_{(12)} - e_{(13)} - e_{(132)}$$

in $\mathbb{C}\mathfrak{S}_3$, and $V_{(2,1)}$ is spanned by $c_{(2,1)}$ and $(13) \cdot c_{(2,1)}$, so $V_{(2,1)}$ is the standard representation of $\mathfrak{S}_3$.

Exercise 4.4*. Set $A = \mathbb{C}\mathfrak{S}_d$, so $V_\lambda = Ac_\lambda = Aa_\lambda b_\lambda$.

(a) Show that $V_\lambda \cong Ab_\lambda a_\lambda$.

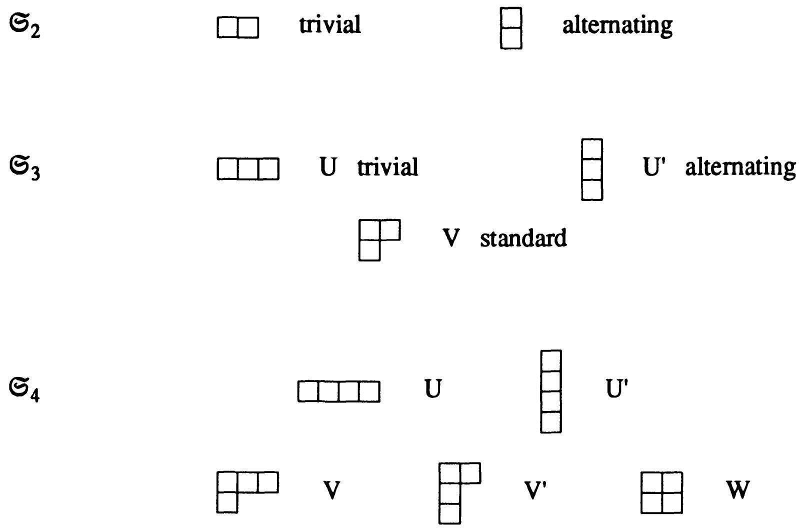
(b) Show that V_λ is the image of the map from Aa_λ to Ab_λ given by right multiplication by b_λ. By (a), this is isomorphic to the image of $Ab_\lambda \to Aa_\lambda$ given by right multiplication by a_λ.

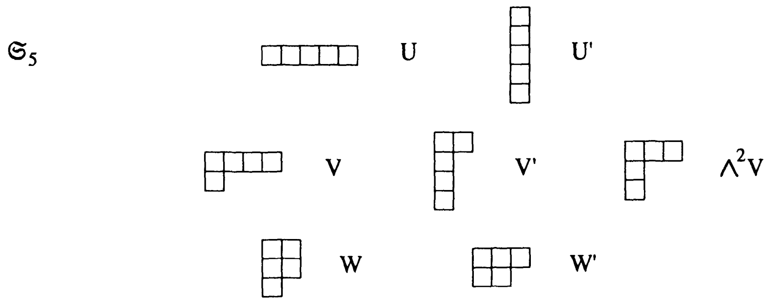
(c) Using (a) and the description of V_λ in the theorem show that

$$V_{\lambda'} = V_\lambda \otimes U',$$

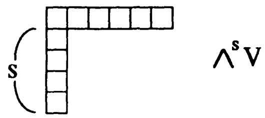
where λ' is the conjugate partition to λ and U' is the alternating representation.

Examples 4.5. In earlier lectures we described the irreducible representations of $\mathfrak{S}_d$ for $d \leq 5$. From the construction of the representation corresponding to a Young diagram it is not hard to work out which representations come from which diagrams:





Exercise 4.6*. Show that for general d, the standard representation V corresponds to the partition $d = (d - 1) + 1$. As a challenge, you can try to prove that the exterior powers of the standard representation V are represented by a "hook":



Note that this recovers our theorem that the $\Lambda^s V$ are irreducible.

Next we turn to Frobenius's formula for the character χ_λ of V_λ, which includes a formula for its dimension. Let $C_{\mathbf{i}}$ denote the conjugacy class in $\mathfrak{S}_d$ determined by a sequence

$$\mathbf{i} = (i_1, i_2, \ldots, i_d) \quad \text{with} \sum \alpha i_\alpha = d:$$

$C_{\mathbf{i}}$ consists of those permutations that have i_1 1-cycles, i_2 2-cycles, $\ldots$, and i_d d-cycles.

Introduce independent variables $x_1, \ldots, x_k$, with k at least as large as the number of rows in the Young diagram of λ. Define the *power sums* $P_j(x)$, $1 \le j \le d$, and the *discriminant* $\Delta(x)$ by

$$\begin{aligned} P_j(x) &= x_1^j + x_2^j + \cdots + x_k^j, \\ \Delta(x) &= \prod_{i<j} (x_i - x_j). \end{aligned} \tag{4.7}$$

If $f(x) = f(x_1, \ldots, x_k)$ is a formal power series, and $(l_1, \ldots, l_k)$ is a k-tuple of non-negative integers, let

$$[f(x)]_{(l_1, \ldots, l_k)} = \text{coefficient of } x_1^{l_1} \cdot \cdots \cdot x_k^{l_k} \text{ in } f. \tag{4.8}$$

Given a partition λ: $\lambda_1 \ge \cdots \ge \lambda_k \ge 0$ of d, set

$$l_1 = \lambda_1 + k - 1, \qquad l_2 = \lambda_2 + k - 2, \ldots, l_k = \lambda_k, \tag{4.9}$$

a strictly decreasing sequence of k non-negative integers. The character of V_λ evaluated on $g \in C_{\mathbf{i}}$ is given by the remarkable

Frobenius Formula 4.10

$$\chi_\lambda(C_{\mathbf{i}}) = \left[\Delta(x) \cdot \prod_j P_j(x)^{i_j}\right]_{(l_1, \ldots, l_k)}.$$

For example, if $d = 5$, $\lambda = (3, 2)$, and $C_{\mathbf{i}}$ is the conjugacy class of $(12)(345)$, i.e., $i_1 = 0$, $i_2 = 1$, $i_3 = 1$, then

$$\chi_{(3,2)}(C_{\mathbf{i}}) = [(x_1 - x_2) \cdot (x_1^2 + x_2^2)(x_1^3 + x_2^3)]_{(4,2)} = 1.$$

Other entries in our character tables for $\mathfrak{S}_3$, $\mathfrak{S}_4$, and $\mathfrak{S}_5$ can be verified as easily, verifying the assertions of Examples 4.5.

In terms of certain symmetric functions S_λ called *Schur polynomials*, Frobenius's formula can be expressed by

$$\prod_j P_j(x)^{i_j} = \sum \chi_\lambda(C_{\mathbf{i}}) S_\lambda,$$

the sum over all partitions λ of d in at most k parts (cf. Proposition 4.37 and (A.27)). Although we do not use Schur polynomials explicitly in this lecture, they play the central role in the algebraic background developed in Appendix A.

Let us use the Frobenius formula to compute the dimension of V_λ. The conjugacy class of the identity corresponds to $\mathbf{i} = (d)$, so

$$\dim V_\lambda = \chi_\lambda(C_{(d)}) = [\Delta(x) \cdot (x_1 + \cdots + x_k)^d]_{(l_1, \ldots, l_k)}.$$

Now $\Delta(x)$ is the Vandermonde determinant:

$$\begin{vmatrix} 1 & x_k & \cdots & x_k^{k-1} \\ \vdots & \vdots & & \vdots \\ 1 & x_1 & \cdots & x_1^{k-1} \end{vmatrix} = \sum_{\sigma \in \mathfrak{S}_k} (\operatorname{sgn} \sigma) x_k^{\sigma(1)-1} \cdot \ldots \cdot x_1^{\sigma(k)-1}.$$

The other term is

$$(x_1 + \cdots + x_k)^d = \sum \frac{d!}{r_1! \cdot \ldots \cdot r_k!} x_1^{r_1} x_2^{r_2} \cdot \ldots \cdot x_k^{r_k},$$

the sum over k-tuples $(r_1, \ldots, r_k)$ that sum to d. To find the coefficient of $x_1^{l_1} \cdot \ldots \cdot x_k^{l_k}$ in the product, we pair off corresponding terms in these two sums, getting

$$\sum \operatorname{sgn}(\sigma) \cdot \frac{d!}{(l_1 - \sigma(k) + 1)! \cdots (l_k - \sigma(1) + 1)!},$$

the sum over those σ in $\mathfrak{S}_k$ such that $l_{k-i+1} - \sigma(i) + 1 \geq 0$ for all $1 \leq i \leq k$. This sum can be written as

$$\frac{d!}{l_1!\cdots l_k!}\sum_{\sigma\in\mathfrak{S}_k}\operatorname{sgn}(\sigma)\prod_{j=1}^{k} l_j(l_j-1)\cdot\ldots\cdot(l_j-\sigma(k-j+1)+2)$$

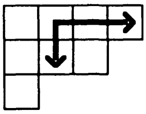
$$=\frac{d!}{l_1!\cdots l_k!}\begin{vmatrix} 1 & l_k & l_k(l_k-1) & \cdots \\ \vdots & \vdots & \vdots & \vdots \\ 1 & l_1 & l_1(l_1-1) & \cdots \end{vmatrix}.$$

By column reduction this determinant reduces to the Vandermonde determinant, so

$$\dim V_\lambda = \frac{d!}{l_1!\cdots l_k!}\prod_{i<j}(l_i-l_j), \tag{4.11}$$

with $l_i = \lambda_i + k - i$.

There is another way of expressing the dimensions of the V_λ. The *hook length* of a box in a Young diagram is the number of squares directly below or directly to the right of the box, including the box once.



In the following diagram, each box is labeled by its hook length:

6	4	3	1
4	2	1	
1			

Hook Length Formula 4.12.

$$\dim V_\lambda = \frac{d!}{\prod(\text{Hook lengths})}.$$

For the above partition $4 + 3 + 1$ of 8, the dimension of the corresponding representation of $\mathfrak{S}_8$ is therefore $8!/6\cdot 4\cdot 4\cdot 2\cdot 3 = 70$.

Exercise 4.13*. Deduce the hook length formula from the Frobenius formula (4.11).

Exercise 4.14*. Use the hook length formula to show that the only irreducible representations of $\mathfrak{S}_d$ of dimension less than d are the trivial and alternating representations U and U' of dimension 1, the standard representation V and $V' = V\otimes U'$ of dimension $d-1$, and three other examples: the two-dimensional representation of $\mathfrak{S}_4$ corresponding to the partition $4 = 2 + 2$, and the two five-dimensional representations of $\mathfrak{S}_6$ corresponding to the partitions $6 = 3 + 3$ and $6 = 2 + 2 + 2$.

Exercise 4.15*. Using Frobenius's formula or otherwise, show that:

$$\chi_{(d-1,1)}(C_{\mathbf{i}}) = i_1 - 1;$$
$$\chi_{(d-2,1,1)}(C_{\mathbf{i}}) = \tfrac{1}{2}(i_1 - 1)(i_1 - 2) - i_2;$$
$$\chi_{(d-2,2)}(C_{\mathbf{i}}) = \tfrac{1}{2}(i_1 - 1)(i_1 - 2) + i_2 - 1.$$

Can you continue this list?

Exercise 4.16*. If g is a cycle of length d in $\mathfrak{S}_d$, show that $\chi_\lambda(g)$ is ± 1 if λ is a hook, and zero if λ is not a hook:

$$\chi_\lambda(g) = \begin{cases} (-1)^s & \text{if } \lambda = (d - s, 1, \ldots, 1), 0 \le s \le d - 1 \\ 0 & \text{otherwise.} \end{cases}$$

Exercise 4.17. Frobenius [Fro1] used his formula to compute the value of χ_λ on a cycle of length $m \le d$.

(a) Following the procedure that led to (4.11)—which was the case $m = 1$—show that

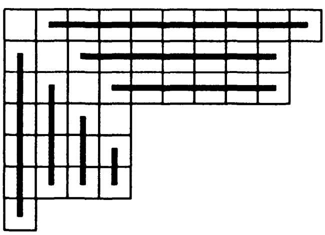
$$\chi_\lambda((12 \ldots m)) = \frac{\dim V_\lambda}{-m^2 h_m} \sum_{p=1}^{k} \frac{\psi(l_p)}{\varphi'(l_p)}, \tag{4.18}$$

where $h_m = d!/(d - m)!m$ is the number of cycles of length m (if $m > 1$), and

$$\varphi(x) = \prod_{i=1}^{k} (x - l_i), \qquad \psi(x) = \varphi(x - m) \prod_{j=1}^{m} (x - j + 1).$$

The sum in (4.18) can be realized as the coefficient of x^{-1} in the Laurent expansion of $\psi(x)/\varphi(x)$ at $x = \infty$.

Define the *rank* r of a partition to be the length of the diagonal of its Young diagram, and let a_i and b_i be the number of boxes below and to the right of the ith box of the diagonal, reading from lower right to upper left. Frobenius called $\begin{pmatrix} a_1 a_2 \ldots a_r \\ b_1 b_2 \ldots b_r \end{pmatrix}$ the *characteristics* of the partition. (Many writers now use a reverse notation for the characteristics, writing $(b_r, \ldots, b_1 | a_r, \ldots, a_r)$ instead.) For the partition (10, 9, 9, 4, 4, 4, 1):



$r = 4$

$$\text{characteristics} = \begin{pmatrix} 2 & 3 & 4 & 6 \\ 0 & 6 & 7 & 9 \end{pmatrix}$$

Algebraically, r and the characteristics $a_1 < \cdots < a_r$ and $b_1 < \cdots < b_r$ are determined by requiring the equality of the two sets

$$\{l_1, \ldots, l_k, k-1-a_1, \ldots, k-1-a_r\} \quad \text{and}$$
$$\{0, 1, \ldots, k-1, k+b_1, \ldots, k+b_r\}.$$

(b) Show that $\psi(x)/\varphi(x) = g(y)/f(y)$, where $y = x - d$ and

$$f(y) = \frac{\prod_{i=1}^{r}(y - b_i)}{\prod_{i=1}^{r}(y + a_i + 1)}, \qquad g(y) = f(y-m)\prod_{j=1}^{m}(y - j + 1).$$

Deduce that the sum in (4.18) is the coefficient of x^{-1} in $g(x)/f(x)$.

(c) When $m = 2$, use this to prove the formula

$$\chi_\lambda((12)) = \frac{\dim V_\lambda}{d(d-1)} \sum_{i=1}^{r} (b_i(b_i + 1) - a_i(a_i + 1)).$$

Hurwitz [Hur] used this formula of Frobenius to calculate the number of ways to write a given permutation as a product of transpositions. From this he gave a formula for the number of branched coverings of the Riemann sphere with a given number of sheets and given simple branch points. Ingram [In] has given other formulas for $\chi_\lambda(g)$, when g is a somewhat more complicated conjugacy class.

Exercise 4.19*. If V is the standard representation of $\mathfrak{S}_d$, prove the decompositions into irreducible representations:

$$\mathrm{Sym}^2 V \cong U \oplus V \oplus V_{(d-2,2)},$$
$$V \otimes V = \mathrm{Sym}^2 V \oplus \Lambda^2 V \cong U \oplus V \oplus V_{(d-2,2)} \oplus V_{(d-2,1,1)}.$$

Exercise 4.20*. Suppose λ is symmetric, i.e., $\lambda = \lambda'$, and let $q_1 > q_2 > \cdots > q_r > 0$ be the lengths of the symmetric hooks that form the diagram of λ; thus, $q_1 = 2\lambda_1 - 1$, $q_2 = 2\lambda_2 - 3, \ldots$. Show that if g is a product of disjoint cycles of lengths $q_1, q_2, \ldots, q_r$, then

$$\chi_\lambda(g) = (-1)^{(d-r)/2}.$$

§4.2. Irreducible Representations of $\mathfrak{S}_d$

We show next that the representations V_λ constructed in the first section are exactly the irreducible representations of $\mathfrak{S}_d$. This proof appears in many standard texts (e.g. [C-R], [Ja-Ke], [N-S], [Wel]), so we will be a little concise.

Let $A = \mathbb{C}\mathfrak{S}_d$ be the group ring of $\mathfrak{S}_d$. For a partition λ of d, let P and Q be the corresponding subgroups preserving the rows and columns of a Young tableau T corresponding to λ, let $a = a_\lambda$, $b = b_\lambda$, and let $c = c_\lambda = ab$ be the corresponding Young symmetrizer, so $V_\lambda = Ac_\lambda$ is the corresponding representation. (These groups and elements should really be subscripted by T to denote dependence on the tableau chosen, but the assertions made depend only on the partition, so we usually omit reference to T.)

Note that $P \cap Q = \{1\}$, so an element of $\mathfrak{S}_d$ can be written in at most one way as a product $p \cdot q$, $p \in P$, $q \in Q$. Thus, c is the sum $\sum \pm e_g$, the sum over all g that can be written as $p \cdot q$, with coefficient ± 1 being $\operatorname{sgn}(q)$; in particular, the coefficient of e_1 in c is 1.

Lemma 4.21. (1) *For $p \in P$, $p \cdot a = a \cdot p = a$.*

(2) *For $q \in Q$, $(\operatorname{sgn}(q)q) \cdot b = b \cdot (\operatorname{sgn}(q)q) = b$.*

(3) *For all $p \in P$, $q \in Q$, $p \cdot c \cdot (\operatorname{sgn}(q)q) = c$, and, up to multiplication by a scalar, c is the only such element in A.*

Proof. Only the last assertion is not obvious. If $\sum n_g e_g$ satisfies the condition in (3), then $n_{pgq} = \operatorname{sgn}(q) n_g$ for all g, p, q; in particular, $n_{pq} = \operatorname{sgn}(q) n_1$. Thus, it suffices to verify that $n_g = 0$ if $g \notin PQ$. For such g it suffices to find a transposition t such that $p = t \in P$ and $q = g^{-1}tg \in Q$; for then $g = pgq$, so $n_g = -n_g$. If $T' = gT$ is the tableau obtained by replacing each entry i of T by $g(i)$, the claim is that there is are two distinct integers that appear in the same row of T and in the same column of T'; t is then the transposition of these two integers. We must verify that if there were no such pair of integers, then one could write $g = p \cdot q$ for some $p \in P$, $q \in Q$. To do this, first take $p_1 \in P$ and $q_1' \in Q' = gQg^{-1}$ so that $p_1 T$ and $q_1' T'$ have the same first row; repeating on the rest of the tableau, one gets $p \in P$ and $q' \in Q'$ so that $pT = q'T'$. Then $pT = q'gT$, so $p = q'g$, and therefore $g = pq$, where $q = g^{-1}(q')^{-1}g \in Q$, as required. □

We order partitions *lexicographically*:

$$\lambda > \mu \quad \text{if the first nonvanishing } \lambda_i - \mu_i \text{ is positive.} \tag{4.22}$$

Lemma 4.23. (1) *If $\lambda > \mu$, then for all $x \in A$, $a_\lambda \cdot x \cdot b_\mu = 0$. In particular, if $\lambda > \mu$, then $c_\lambda \cdot c_\mu = 0$.*

(2) *For all $x \in A$, $c_\lambda \cdot x \cdot c_\lambda$ is a scalar multiple of c_λ. In particular, $c_\lambda \cdot c_\lambda = n_\lambda c_\lambda$ for some $n_\lambda \in \mathbb{C}$.*

Proof. For (1), we may take $x = g \in \mathfrak{S}_d$. Since $g \cdot b_\mu \cdot g^{-1}$ is the element constructed from gT', where T' is the tableau used to construct b_μ, it suffices to show that $a_\lambda \cdot b_\mu = 0$. One verifies that $\lambda > \mu$ implies that there are two integers in the same row of T and the same column of T'. If t is the transposition of these integers, then $a_\lambda \cdot t = a_\lambda$, $t \cdot b_\mu = -b_\mu$, so $a_\lambda \cdot b_\mu = a_\lambda \cdot t \cdot t \cdot b_\mu = -a_\lambda \cdot b_\mu$, as required. Part (2) follows from Lemma 4.21 (3). □

Exercise 4.24*. Show that if $\lambda \neq \mu$, then $c_\lambda \cdot A \cdot c_\mu = 0$; in particular, $c_\lambda \cdot c_\mu = 0$.

Lemma 4.25. (1) *Each V_λ is an irreducible representation of $\mathfrak{S}_d$.*
(2) *If $\lambda \neq \mu$, then V_λ and V_μ are not isomorphic.*

PROOF. For (1) note that $c_\lambda V_\lambda \subset \mathbb{C}c_\lambda$ by Lemma 4.23. If $W \subset V_\lambda$ is a subrepresentation, then $c_\lambda W$ is either $\mathbb{C}c_\lambda$ or 0. If the first is true, then $V_\lambda = A \cdot c_\lambda \subset W$. Otherwise $W \cdot W \subset A \cdot c_\lambda W = 0$, but this implies $W = 0$. Indeed, a projection from A onto W is given by right multiplication by an element $\varphi \in A$ with $\varphi = \varphi^2 \in W \cdot W = 0$. This argument also shows that $c_\lambda V_\lambda \neq 0$.

For (2), we may assume $\lambda > \mu$. Then $c_\lambda V_\lambda = \mathbb{C}c_\lambda \neq 0$, but $c_\lambda V_\mu = c_\lambda \cdot Ac_\mu = 0$, so they cannot be isomorphic A-modules. □

Lemma 4.26. *For any λ, $c_\lambda \cdot c_\lambda = n_\lambda c_\lambda$, with $n_\lambda = d!/\dim V_\lambda$.*

PROOF. Let F be right multiplication by c_λ on A. Since F is multiplication by n_λ on V_λ, and zero on $\operatorname{Ker}(c_\lambda)$, the trace of F is n_λ times the dimension of V_λ. But the coefficient of e_g in $e_g \cdot c_\lambda$ is 1, so trace$(F) = |\mathfrak{S}_d| = d!$. □

Since there are as many irreducible representations V_λ as conjugacy classes of $\mathfrak{S}_d$, these must form a complete set of isomorphism classes of irreducible representations, which completes the proof of Theorem 4.3. In the next section we will prove Frobenius's formula for the character of V_λ, and, in a series of exercises, discuss a little of what else is known about them: how to decompose tensor products or induced or restricted representations, how to find a basis for V_λ, etc.

§4.3. Proof of Frobenius's Formula

For any partition λ of d, we have a subgroup, often called a *Young subgroup*,

$$\mathfrak{S}_\lambda = \mathfrak{S}_{\lambda_1} \times \cdots \times \mathfrak{S}_{\lambda_k} \hookrightarrow \mathfrak{S}_d. \tag{4.27}$$

Let U_λ be the representation of $\mathfrak{S}_d$ induced from the trivial representation of $\mathfrak{S}_\lambda$. Equivalently, $U_\lambda = A \cdot a_\lambda$, with a_λ as in the preceding section. Let

$$\psi_\lambda = \chi_{U_\lambda} = \text{character of } U_\lambda. \tag{4.28}$$

Key to this investigation is the relation between U_λ and V_λ, i.e., between ψ_λ and the character χ_λ of V_λ. Note first that V_λ appears in U_λ, since there is a surjection

$$U_\lambda = Aa_\lambda \twoheadrightarrow V_\lambda = Aa_\lambda b_\lambda, \quad x \mapsto x \cdot b_\lambda. \tag{4.29}$$

Alternatively,

$$V_\lambda = Aa_\lambda b_\lambda \cong Ab_\lambda a_\lambda \subset Aa_\lambda = U_\lambda,$$

by Exercise 4.4. For example, we have

$$U_{(d-1,1)} \cong V_{(d-1,1)} \oplus V_{(d)}$$

which expresses the fact that the permutation representation $\mathbb{C}^d$ of $\mathfrak{S}_d$ is the sum of the standard representation and the trivial representation. Eventually we will see that every U_λ contains V_λ with multiplicity one, and contains only other V_μ for $\mu > \lambda$.

The character of U_λ is easy to compute directly since U_λ is an induced representation, and we do this next.

For $\mathbf{i} = (i_1, \ldots, i_d)$ a d-tuple of non-negative integers with $\sum \alpha i_\alpha = d$, denote by

$$C_\mathbf{i} \subset \mathfrak{S}_d$$

the conjugacy class consisting of elements made up of i_1 1-cycles, i_2 2-cycles, $\ldots$, i_d d-cycles. The number of elements in $C_\mathbf{i}$ is easily counted to be

$$|C_\mathbf{i}| = \frac{d!}{1^{i_1} i_1! 2^{i_2} i_2! \cdot \ldots \cdot d^{i_d} i_d!}. \tag{4.30}$$

By the formula for characters of induced representations (Exercise 3.19),

$$\begin{aligned}\psi_\lambda(C_\mathbf{i}) &= \frac{1}{|C_\mathbf{i}|}[\mathfrak{S}_d : \mathfrak{S}_\lambda] \cdot |C_\mathbf{i} \cap \mathfrak{S}_\lambda| \\ &= \frac{1^{i_1} i_1! \cdot \ldots \cdot d^{i_d} i_d!}{d!} \cdot \frac{d!}{\lambda_1! \cdot \ldots \cdot \lambda_k!} \cdot \sum \prod_{p=1}^{k} \frac{\lambda_p!}{1^{r_{p1}} r_{p1}! \cdot \ldots d^{r_{pd}} r_{pd}!},\end{aligned}$$

where the sum is over all collections $\{r_{pq}: 1 \le p \le k, 1 \le q \le d\}$ of non-negative integers satisfying

$$\begin{aligned} i_q &= r_{1q} + r_{2q} + \cdots + r_{kq}, \\ \lambda_p &= r_{p1} + 2r_{p2} + \cdots + dr_{pd}. \end{aligned}$$

(To count $C_\mathbf{i} \cap \mathfrak{S}_\lambda$, write the pth component of an element of $\mathfrak{S}_\lambda$ as a product of r_{p1} 1-cycles, r_{p2} 2-cycles, ...) Simplifying,

$$\psi_\lambda(C_\mathbf{i}) = \sum \prod_{q=1}^{d} \frac{i_q!}{r_{1q}! r_{2q}! \cdot \ldots \cdot r_{kq}!}, \tag{4.31}$$

the sum over the same collections of integers $\{r_{pq}\}$.

This sum is exactly the coefficient of the monomial $X^\lambda = x_1^{\lambda_1} \cdot \ldots \cdot x_k^{\lambda_k}$ in the power sum symmetric polynomial

$$P^{(\mathbf{i})} = (x_1 + \cdots + x_k)^{i_1} \cdot (x_1^2 + \cdots + x_k^2)^{i_2} \cdot \ldots \cdot (x_1^d + \cdots + x_k^d)^{i_d}. \tag{4.32}$$

So we have the formula

$$\psi_\lambda(C_\mathbf{i}) = [P^{(\mathbf{i})}]_\lambda = \text{coefficient of } X^\lambda \text{ in } P^{(\mathbf{i})}. \tag{4.33}$$

To prove Frobenius's formula, we need to compare these coefficients with the coefficients $\omega_\lambda(\mathbf{i})$ defined by

$$\omega_\lambda(\mathbf{i}) = [\Delta \cdot P^{(\mathbf{i})}]_l, \quad l = (\lambda_1 + k - 1, \lambda_2 + k - 2, \ldots, \lambda_k). \tag{4.34}$$

Our goal, Frobenius's formula, is the assertion that $\chi_\lambda(C_\mathbf{i}) = \omega_\lambda(\mathbf{i})$.

There is a general identity, valid for any symmetric polynomial P, relating such coefficients:

$$[P]_\lambda = \sum_\mu K_{\mu\lambda}[\Delta \cdot P]_{(\mu_1+k-1, \mu_2+k-2, \ldots, \mu_k)},$$

where the coefficients $K_{\mu\lambda}$ are certain universally defined integers, called *Kostka numbers.* For any partitions λ and μ of d, the integer $K_{\mu\lambda}$ may be defined combinatorially as the number of ways to fill the boxes of the Young diagram for μ with λ_1 1's, λ_2 2's, up to λ_k k's, in such a way that the entries in each row are nondecreasing, and those in each column are strictly increasing; such are called *semistandard tableaux on μ of type λ.* In particular,

$$K_{\lambda\lambda} = 1, \quad \text{and } K_{\mu\lambda} = 0 \text{ for } \mu < \lambda.$$

The integer $K_{\mu\lambda}$ may be also be defined to be the coefficient of the monomial $X^\lambda = x_1^{\lambda_1} \cdot \ldots \cdot x_k^{\lambda_k}$ in the Schur polynomial S_μ corresponding to μ. For the proof that these are equivalent definitions, see (A.9) and (A.19) of Appendix A. In the present case, applying Lemma A.26 to the polynomial $P = P^{(\mathbf{i})}$, we deduce

$$\psi_\lambda(C_\mathbf{i}) = \sum_\mu K_{\mu\lambda}\omega_\mu(\mathbf{i}) = \omega_\lambda(\mathbf{i}) + \sum_{\mu > \lambda} K_{\mu\lambda}\omega_\mu(\mathbf{i}). \tag{4.35}$$

The result of Lemma A.28 can be written, using (4.30), in the form

$$\frac{1}{d!}\sum_\mathbf{i} |C_\mathbf{i}|\omega_\lambda(\mathbf{i})\omega_\mu(\mathbf{i}) = \delta_{\lambda\mu}. \tag{4.36}$$

This indicates that the functions ω_λ, regarded as functions on the conjugacy classes of $\mathfrak{S}_d$, satisfy the same orthogonality relations as the irreducible characters of $\mathfrak{S}_d$. In fact, one can deduce formally from these equations that the ω_λ must be the irreducible characters of $\mathfrak{S}_d$, which is what Frobenius proved. A little more work is needed to see that ω_λ is actually the character of the representation V_λ, that is, to prove

Proposition 4.37. *Let $\chi_\lambda = \chi_{V_\lambda}$ be the character of V_λ. Then for any conjugacy class $C_\mathbf{i}$ of $\mathfrak{S}_d$,*

$$\chi_\lambda(C_\mathbf{i}) = \omega_\lambda(\mathbf{i}).$$

PROOF. We have seen in (4.29) that the representation U_λ, whose character is ψ_λ, contains the irreducible representation V_λ. In fact, this is all that we need to know about the relation between U_λ and V_λ. It implies that we have

$$\psi_\lambda = \sum_\mu n_{\lambda\mu}\chi_\mu, \quad n_{\lambda\lambda} \geq 1, \text{ all } n_{\lambda\mu} \geq 0. \tag{4.38}$$

Consider this equation together with (4.35). We deduce first that each ω_λ is a virtual character: we can write

$$\omega_\lambda = \sum m_{\lambda\mu}\chi_\mu, \quad m_{\lambda\mu} \in \mathbb{Z}.$$

But the ω_λ, like the χ_λ, are orthonormal by (4.36), so

$$1 = (\omega_\lambda, \omega_\lambda) = \sum_\mu m_{\lambda\mu}^2,$$

and hence ω_λ is $\pm\chi$ for some irreducible character χ. (It follows from the hook length formula that the plus sign holds here, but we do not need to assume this.)

Fix λ, and assume inductively that $\chi_\mu = \omega_\mu$ for all $\mu > \lambda$, so by (4.35)

$$\psi_\lambda = \omega_\lambda + \sum_{\mu > \lambda} K_{\mu\lambda}\chi_\mu.$$

Comparing this with (4.38), and using the linear independence of characters, the only possibility is that $\omega_\lambda = \chi_\lambda$. □

Corollary 4.39 (Young's rule). *The integer $K_{\mu\lambda}$ is the multiplicity of the irreducible representation V_μ in the induced representation U_λ:*

$$U_\lambda \cong V_\lambda \oplus \bigoplus_{\mu > \lambda} K_{\mu\lambda} V_\mu, \qquad \psi_\lambda = \chi_\lambda + \sum_{\mu > \lambda} K_{\mu\lambda}\chi_\mu.$$

Note that when $\lambda = (1, \ldots, 1)$, U_λ is just the regular representation, so $K_{\mu(1,\ldots,1)} = \dim V_\mu$. This shows that *the dimension of V_λ is the number of standard tableaux on λ*, i.e., the number of ways to fill the Young diagram of λ with the numbers from 1 to d, such that all rows and columns are increasing. The hook length formula gives another combinatorial formula for this dimension. Frame, Robinson, and Thrall proved that these two numbers are equal. For a short and purely combinatorial proof, see [G-N-W]. For another proof that the dimension of V_λ is the number of standard tableaux, see [Jam]. The latter leads to a canonical decomposition of the group ring $A = \mathbb{C}\mathfrak{S}_d$ as the direct sum of left ideals Ae_T, summing over all standard tableaux, with $e_T = (\dim V_\lambda/d!)\cdot c_T$, and c_T the Young symmetrizer corresponding to T, cf. Exercises 4.47 and 4.50. This, in turn, leads to explicit calculation of matrices of the representations V_λ with integer coefficients.

For another example of Young's rule, we have a decomposition

$$U_{(d-a,a)} = \bigoplus_{i=0}^{a} V_{(d-i,i)}.$$

In fact, the only μ whose diagrams can be filled with $d - a$ 1's and a 2's, nondecreasing in rows and strictly increasing in columns, are those with at most two rows, with the second row no longer than a; and such a diagram has only one such tableau, so there are no multiplicities.

Exercise 4.40*. The characters ψ_λ of $\mathfrak{S}_d$ have been defined only when λ is a partition of d. Extend the definition to any k-tuple $a = (a_1, \ldots, a_k)$ of integers that add up to d by setting $\psi_a = 0$ if any of the a_i are negative, and otherwise $\psi_a = \psi_\lambda$, where λ is the reordering of $a_1, \ldots, a_k$ in descending order. In this case ψ_a is the character of the representation induced from the trivial representation by the inclusion of $\mathfrak{S}_{a_1} \times \cdots \times \mathfrak{S}_{a_k}$ in $\mathfrak{S}_d$. Use (A.5) and (A.9) of Appendix A to prove the *determinantal formula* for the irreducible characters χ_λ in terms of the induced characters ψ_μ:

$$\chi_\lambda = \sum_{\tau \in \mathfrak{S}_k} \operatorname{sgn}(\tau)\psi_{(\lambda_1+\tau(1)-1,\, \lambda_2+\tau(2)-2, \ldots, \lambda_k+\tau(k)-k)}.$$

If one writes ψ_a as a formal product $\psi_{a_1} \cdot \psi_{a_2} \cdot \ldots \cdot \psi_{a_k}$, the preceding formula can be written

$$\chi_\lambda = |\psi_{\lambda_i+j-i}| = \begin{vmatrix} \psi_{\lambda_1} & \psi_{\lambda_1+1} & \psi_{\lambda_1+k-1} \\ \psi_{\lambda_2-1} & \psi_{\lambda_2} \cdots & \\ \vdots & & \vdots \\ \psi_{\lambda_k-k+1} \cdots & & \psi_{\lambda_k} \end{vmatrix}.$$

The formal product of the preceding exercise is the character version of an "outer product" of representations. Given any non-negative integers $d_1, \ldots, d_k$, and representations V_i of $\mathfrak{S}_{d_i}$, denote by $V_1 \circ \cdots \circ V_k$ the (isomorphism class of the) representation of $\mathfrak{S}_d$, $d = \sum d_i$, induced from the tensor product representation $V_1 \boxtimes \cdots \boxtimes V_k$ of $\mathfrak{S}_{d_1} \times \cdots \times \mathfrak{S}_{d_k}$ by the inclusion of $\mathfrak{S}_{d_1} \times \cdots \times \mathfrak{S}_{d_k}$ in $\mathfrak{S}_d$ (see Exercise 2.36). This product is commutative and associative. It will turn out to be useful to have a procedure for decomposing such a representation into its irreducible pieces. For this it is enough to do the case of two factors, and with the individual representations V_i irreducible. In this case, one has, for V_λ the representation of $\mathfrak{S}_d$ corresponding to the partition λ of d and V_μ the representation of $\mathfrak{S}_m$ corresponding to the partition μ of m,

$$V_\lambda \circ V_\mu = \sum N_{\lambda\mu\nu} V_\nu, \tag{4.41}$$

the sum over all partitions ν of $d + m$, with $N_{\lambda\mu\nu}$ the coefficients given by the *Littlewood–Richardson rule* (A.8) of Appendix A. Indeed, by the exercise, the character of $V_\lambda \circ V_\mu$ is the product of the corresponding determinants, and, by (A.8), that is the sum of the characters $N_{\lambda\mu\nu}\chi_\nu$.

When $m = 1$ and $\mu = (m)$, V_μ is trivial; this gives

$$\operatorname{Ind}_{\mathfrak{S}_d}^{\mathfrak{S}_{d+1}} V_\lambda = \sum V_\nu, \tag{4.42}$$

the sum over all ν whose Young diagram is obtained from that of λ by adding one box. This formula uses only a simpler form of the Littlewood–Richardson rule known as Pieri's formula, which is proved in (A.7).

Exercise 4.43*. Show that the Littlewood–Richardson number $N_{\lambda\mu\nu}$ is the multiplicity of the irreducible representation $V_\lambda \boxtimes V_\mu$ in the restriction of V_ν from $\mathfrak{S}_{d+m}$ to $\mathfrak{S}_d \times \mathfrak{S}_m$. In particular, taking $m = 1$, $\mu = (1)$, Pieri's formula (A.7) gives

$$\operatorname{Res}_{\mathfrak{S}_d}^{\mathfrak{S}_{d+1}} V_\nu = \sum V_\lambda,$$

the sum over all λ obtained from ν by removing one box. This is known as the "branching theorem," and is useful for inductive proofs and constructions, particularly because the decomposition is multiplicity free. For example, you can use it to reprove the fact that the multiplicity of V_λ in U_μ is the number of semistandard tableaux on μ of type λ. It can also be used to prove the assertion made in Exercise 4.6 that the representations corresponding to hooks are exterior powers of the standard representation.

Exercise 4.44* (Pieri's rule). Regard $\mathfrak{S}_d$ as a subgroup of $\mathfrak{S}_{d+m}$ as usual. Let λ be a partition of d and ν a partition of $d + m$. Use Exercise 4.40 to show that the multiplicity of V_ν in the induced representation $\mathrm{Ind}(V_\lambda)$ is zero unless the Young diagram of λ is contained in that of ν, and then it is the number of ways to number the skew diagram lying between them with the numbers from 1 to m, increasing in both row and column. By Frobenius reciprocity, this is the same as the multiplicity of V_λ in $\mathrm{Res}(V_\nu)$.

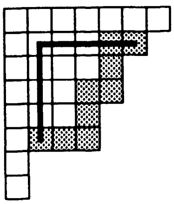
When applied to $d = 0$ (or 1), this implies again that the dimension of V_ν is the number of standard tableaux on the Young diagram of ν.

For a sampling of the many applications of these rules, see [Dia §7, §8].

Problem 4.45*. The *Murnaghan–Nakayama rule* gives an efficient inductive method for computing character values: If λ is a partition of d, and $g \in \mathfrak{S}_d$ is written as a product of an m-cycle and a disjoint permutation $h \in \mathfrak{S}_{d-m}$, then

$$\chi_\lambda(g) = \sum (-1)^{r(\mu)} \chi_\mu(h),$$

where the sum is over all partitions μ of $d - m$ that are obtained from λ by removing a skew hook of length m, and $r(\mu)$ is the number of vertical steps in the skew hook, i.e., one less than the number of rows in the hook. A *skew hook* for λ is a connected region of boundary boxes for its Young diagram such that removing them leaves a smaller Young diagram; there is a one-to-one correspondence between skew hooks and ordinary hooks of the same size, as indicated:



$\lambda = (7, 6, 5, 5, 4, 4, 1, 1)$

$\mu = (7, 4, 4, 3, 3, 1, 1, 1)$

hook length $= 9$, $r = 4$

For example, if λ has no hooks of length m, then $\chi_\lambda(g) = 0$.

The Murnaghan–Nakayama rule may be written inductively as follows: If g is a written as a product of disjoint cycles of lengths $m_1, m_2, \ldots, m_p$, with the lengths m_i taken in any order, then $\chi_\lambda(g)$ is the sum $\sum (-1)^{r(s)}$, where the sum is over all ways s to decompose the Young diagram of λ by successively removing p skew hooks of lengths $m_1, \ldots, m_p$, and $r(s)$ is the total number of vertical steps in the hooks of s.

(a) Deduce the Murnaghan–Nakayama rule from (4.41) and Exercise 4.16, using the Littlewood–Richardson rule. Or:

(b) With the notation of Exercise 4.40, show that

$$\psi_{a_1}\psi_{a_2}\cdot\ldots\cdot\psi_{a_k}(g) = \sum_{i=1}^{k} \psi_{a_1}\psi_{a_2}\cdot\ldots\cdot\psi_{a_i-m}\psi_{a_{i+1}}\cdot\ldots\cdot\psi_{a_k}(h).$$

Exercise 4.46*. Show that Corollary 4.39 implies the "Snapper conjecture": the irreducible representation V_μ occurs in the induced representation U_λ if and only if

$$\sum_{i=1}^{j} \lambda_i \leq \sum_{i=1}^{j} \mu_i \quad \text{for all } j \geq 1.$$

Problem 4.47*. There is a more intrinsic construction of the irreducible representation V_λ, called a *Specht module*, which does not involve the choice of a tableau; it is also useful for studying representations of $\mathfrak{S}_d$ in positive characteristic. Define a *tabloid* $\{T\}$ to be an equivalence class of tableaux (numberings by the integers 1 to d) on λ, two being equivalent if the rows are the same up to order. Then $\mathfrak{S}_d$ acts by permutations on the tabloids, and the corresponding representation, with basis the tabloids, is isomorphic to U_λ. For each tableau T, define an element E_T in this representation space, by

$$E_T = b_T\{T\} = \sum \operatorname{sgn}(q)\{qT\},$$

the sum over the q that preserve the columns of T. The span of all E_T's is isomorphic to V_λ, and the E_T's, where T varies over the standard tableaux, form a basis.

Another construction of V_λ is to take the subspace of the polynomial ring $\mathbb{C}[x_1, \ldots, x_d]$ spanned by all polynomials F_T, where $F_T = \prod(x_i - x_j)$, the product over all pairs $i < j$ which occur in the same column in the tableau T.

Exercise 4.48*. Let U'_λ be the representation $A \cdot b_\lambda$, which is the representation of $\mathfrak{S}_d$ induced from the tensor product of the alternating representations on the subgroup $\mathfrak{S}_\mu = \mathfrak{S}_{\mu_1} \times \cdots \times \mathfrak{S}_{\mu_r}$, where $\mu = \lambda'$ is the conjugate partition. Show that the decomposition of U'_λ is

$$U'_\lambda = \sum_\mu K_{\mu'\lambda'} V_\mu.$$

Deduce that V_λ is the only irreducible representation that occurs in both U_λ and U'_λ, and it occurs in each with multiplicity one.

Note, however, that in general $A \cdot c_\lambda \neq A \cdot a_\lambda \cap A \cdot b_\lambda$ since $A \cdot c_\lambda$ may not be contained in $A \cdot a_\lambda$.

Exercise 4.49*. With notation as in (4.41), if $U' = V_{(1,\ldots,1)}$ is the alternating representation of $\mathfrak{S}_m$, show that $V_\lambda \circ V_{(1,\ldots,1)}$ decomposes into a direct sum $\bigoplus V_\pi$, the sum over all π whose Young diagram can be obtained from that of λ by adding m boxes, with no two in the same row.

Exercise 4.50. We have seen that $A = \mathbb{C}\mathfrak{S}_d$ is isomorphic to a direct sum of m_λ copies of $V_\lambda = Ac_\lambda$, where $m_\lambda = \dim V_\lambda$ is the number of standard tableaux on λ. This can be seen explicitly as follows. For each standard tableau T on each λ, let c_T be the element of $\mathbb{C}\mathfrak{S}_d$ constructed from T. Then $A = \bigoplus A \cdot c_T$. Indeed, an argument like that in Lemma 4.23 shows that $c_T \cdot c_{T'} = 0$ whenever T and T' are tableaux on the same diagram and $T > T'$, i.e., the first entry (reading from left to right, then top to bottom) where the tableaux differ has the entry of T larger than that of T'. From this it follows that the sum $\Sigma A \cdot c_T$ is direct. A dimension count concludes the proof. (This also gives another proof that the dimension of V_λ is the number of standard tableaux on λ, provided one verifies that the sum of the squares of the latter numbers is $d!$, cf. [Boe] or [Ke].)

Exercise 4.51*. There are several methods for decomposing a tensor product of two representations of $\mathfrak{S}_d$, which amounts to finding the coefficients $C_{\lambda\mu\nu}$ in the decomposition

$$V_\lambda \otimes V_\mu \cong \Sigma_\nu C_{\lambda\mu\nu} V_\nu,$$

for λ, μ, and ν partitions of d. Since one knows how to express V_μ in terms of the induced representations U_ν, it suffices to compute $V_\lambda \otimes U_\nu$, which is isomorphic to $\mathrm{Ind}(\mathrm{Res}(V_\lambda))$, restricting and inducing from the subgroup $\mathfrak{S}_\nu = \mathfrak{S}_{\nu_1} \times \mathfrak{S}_{\nu_2} \times \cdots$; this restriction and induction can be computed by the Littlewood–Richardson rule. For $d \leq 5$, you can work out these coefficients using only restriction to $\mathfrak{S}_{d-1}$ and Pieri's formula.

(a) Prove the following closed-form formula for the coefficients, which shows in particular that they are independent of the ordering of the subscripts λ, μ, and ν:

$$C_{\lambda\mu\nu} = \sum_{\mathbf{i}} \frac{1}{z(\mathbf{i})} \omega_\lambda(\mathbf{i})\omega_\mu(\mathbf{i})\omega_\nu(\mathbf{i}),$$

the sum over all $\mathbf{i} = (i_1, \ldots, i_d)$ with $\Sigma \alpha i_\alpha = d$, and with $\omega_\lambda(\mathbf{i}) = \chi_\lambda(C_\mathbf{i})$ and $z(\mathbf{i}) = i_1!1^{i_1} \cdot i_2!2^{i_2} \cdot \ldots \cdot i_d!d^{i_d}$.

(b) Show that

$$C_{\lambda\mu(d)} = \begin{cases} 1 & \text{if } \mu = \lambda \\ 0 & \text{otherwise,} \end{cases} \qquad C_{\lambda\mu(1,\ldots,1)} = \begin{cases} 1 & \text{if } \mu = \lambda' \\ 0 & \text{otherwise.} \end{cases}$$

Exercise 4.52*. Let $R_d = R(\mathfrak{S}_d)$ denote the representation ring, and set $R = \bigoplus_{d=0}^{\infty} R_d$. The outer product of (4.41) determines maps

$$R_n \otimes R_m \to R_{n+m},$$

which makes R into a commutative, graded $\mathbb{Z}$-algebra. Restriction determines maps

$$R_{n+m} = R(\mathfrak{S}_{n+m}) \to R(\mathfrak{S}_n \times \mathfrak{S}_m) = R_n \otimes R_m,$$

which defines a *co-product* $\delta\colon R \to R \otimes R$. Together, these make R into a (graded) Hopf algebra. (This assertion implies many of the formulas we have proved in this lecture, as well as some we have not.)

(a) Show that, as an algebra,

$$R \cong \mathbb{Z}[H_1, \ldots, H_d, \ldots],$$

where H_d is an indeterminate of degree d; H_d corresponds to the trivial representation of $\mathfrak{S}_d$. Show that the co-product δ is determined by

$$\delta(H_n) = H_n \otimes 1 + H_{n-1} \otimes H_1 + \cdots + 1 \otimes H_n.$$

If we set $\Lambda = \mathbb{Z}[H_1, \ldots, H_d, \ldots] = \bigoplus \Lambda_d$, we can identify Λ_d with the symmetric polynomials of degree d in $k \geq d$ variables. The basic symmetric polynomials in Λ_d defined in Appendix A therefore correspond to virtual representations of $\mathfrak{S}_d$.

(b) Show that E_d corresponds to the alternating representation U', and

$$H_\lambda \leftrightarrow U_\lambda, \qquad S_\lambda \leftrightarrow V_\lambda, \qquad E_\lambda \leftrightarrow U'_{\lambda'}.$$

(c) Show that the scalar product $\langle\ ,\ \rangle$ defined on Λ_d in (A.16) corresponds to the scalar product defined on class functions in (2.11).

(d) Show that the involution ϑ of Exercise A.32 corresponds to tensoring a representation with the alternating representation U'.

(e) Show that the inverse map from R_d to Λ_d takes a representation W to

$$\sum_{\mathbf{i}} \frac{1}{z(\mathbf{i})} \chi_W(C_{(\mathbf{i})}) P^{(\mathbf{i})},$$

where $z(\mathbf{i}) = i_1! 1^{i_1} \cdot i_2! 2^{i_2} \cdot \ldots \cdot i_d! d^{i_d}$.

The (inner) tensor product of representations of $\mathfrak{S}_d$ gives a map $R_d \otimes R_d \to R_d$ which corresponds to an "inner product" on symmetric functions, sometimes denoted $*$.

(f) Show that

$$P^{(\mathbf{i})} * P^{(\mathbf{j})} = \begin{cases} 0 & \text{for } \mathbf{j} \neq \mathbf{i} \\ z(\mathbf{i}) P^{(\mathbf{i})} & \text{if } \mathbf{j} = \mathbf{i}. \end{cases}$$

Since these $P^{(\mathbf{i})}$ form a basis for $\Lambda_d \otimes \mathbb{Q}$, this formula determines the inner product.

LECTURE 5

Representations of $\mathfrak{A}_d$ and $GL_2(\mathbb{F}_q)$

In this lecture we analyze the representation of two more types of groups: the alternating groups $\mathfrak{A}_d$ and the linear groups $GL_2(\mathbb{F}_q)$ and $SL_2(\mathbb{F}_q)$ over finite fields. In the former case, we prove some general results relating the representations of a group to the representations of a subgroup of index two, and use what we know about the symmetric group; this should be completely straightforward given just the basic ideas of the preceding lecture. In the latter case we start essentially from scratch. The two sections can be read (or not) independently; neither is logically necessary for the remainder of the book.

§5.1. Representations of $\mathfrak{A}_d$

The alternating groups $\mathfrak{A}_d$, $d \geq 5$, form one of the infinite families of simple groups. In this section, continuing the discussion of §3.1, we describe their irreducible representations. The basic method for analyzing representations of $\mathfrak{A}_d$ is by restricting the representations we know from $\mathfrak{S}_d$.

In general when H is a subgroup of index two in a group G, there is a close relationship between their representations. We will see this phenomenon again in Lie theory for the subgroups SO_n of the orthogonal groups O_n.

Let U and U' denote the trivial and nontrivial representation of G obtained from the two representations of G/H. For any representation V of G, let $V' = V \otimes U'$; the character of V' is the same as the character of V on elements of H, but takes opposite values on elements not in H. In particular, $\mathrm{Res}_H^G V' = \mathrm{Res}_H^G V$.

If W is any representation of H, there is a *conjugate* representation defined by conjugating by any element t of G that is not in H; if ψ is the character of W, the character of the conjugate is $h \mapsto \psi(tht^{-1})$. Since t is unique up to multiplication by an element of H, the conjugate representation is unique up to isomorphism.

Proposition 5.1. *Let V be an irreducible representation of G, and let $W = \mathrm{Res}_H^G V$ be the restriction of V to H. Then exactly one of the following holds:*

(1) *V is not isomorphic to V'; W is irreducible and isomorphic to its conjugate;* $\mathrm{Ind}_H^G W \cong V \oplus V'$.

(2) *$V \cong V'$; $W = W' \oplus W''$, where W' and W'' are irreducible and conjugate but not isomorphic;* $\mathrm{Ind}_H^G W' \cong \mathrm{Ind}_H^G W'' \cong V$.

Each irreducible representation of H arises uniquely in this way, noting that in case (1) *V' and V determine the same representation.*

Proof. Let χ be the character of V. We have

$$|G| = 2|H| = \sum_{h \in H} |\chi(h)|^2 + \sum_{t \notin H} |\chi(t)|^2.$$

Since the first sum is an integral multiple of $|H|$, this multiple must be 1 or 2, which are the two cases of the proposition. This shows that W is either irreducible or the sum of two distinct irreducible representations W' and W''. Note that the second case happens when $\chi(t) = 0$ for all $t \notin H$, which is the case when V' is isomorphic to V. In the second case, W' and W'' must be conjugate since W is self-conjugate, and if W' and W'' were self-conjugate V would not be irreducible. The other assertions in (1) and (2) follow from the isomorphism $\mathrm{Ind}(\mathrm{Res}\ V) = V \otimes (U \oplus U')$ of Exercise 3.16. Similarly, for any representation W of H, $\mathrm{Res}(\mathrm{Ind}\ W)$ is the direct sum of W and its conjugate—as follows say from Exercise 3.19—from which the last statement follows readily. □

Most of this discussion extends with little change to the case where H is a normal subgroup of arbitrary prime index in G, cf. [B-tD, pp. 293–296]. Clifford has extended much of this proposition to arbitrary normal subgroups of finite index, cf. [Dor, §14].

There are two types of conjugacy classes c in H: those that are also conjugacy classes in G, and those such that $c \cup c'$ is a conjugacy class in G, where $c' = tct^{-1}$, $t \notin H$; the latter are called *split*. When W is irreducible, its character assumes the same values—those of the character of the representation V of G that restricts to W—on pairs of split conjugacy classes, whereas in the other case the characters of W' and W'' agree on nonsplit classes, but they must disagree on some split classes. If $\chi_{W'}(c) = \chi_{W''}(c') = x$, and $\chi_{W'}(c') = \chi_{W''}(c) = y$, we know the sum $x + y$, since it is the value of the character of the representation V that gives rise to W' and W'' on $c \cup c'$. Often the exact values of x and y can be determined from orthogonality considerations.

Exercise 5.2*. Show that the number of split conjugacy classes is equal to the number of irreducible representations V of G that are isomorphic to V', or to the number of irreducible representations of H that are not isomorphic to their conjugates. Equivalently, the number of nonsplit classes in H is same as the number of conjugacy classes of G that are not in H.

We apply these considerations to the alternating subgroup of the symmetric group. Consider restrictions of the representations V_λ from $\mathfrak{S}_d$ to $\mathfrak{A}_d$. Recall that if λ' is the conjugate partition to λ, then

$$V_{\lambda'} = V_\lambda \otimes U',$$

with U' the alternating representation. The two cases of the proposition correspond to the cases (1) $\lambda' \neq \lambda$ and (2) $\lambda' = \lambda$. If $\lambda' \neq \lambda$, let W_λ be the restriction of V_λ to $\mathfrak{A}_d$. If $\lambda' = \lambda$, let W_λ' and W_λ'' be the two representations whose sum is the restriction of V_λ. We have

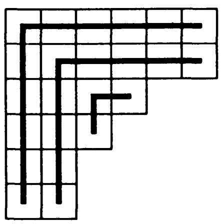
$$\operatorname{Ind} W_\lambda = V_\lambda \oplus V_{\lambda'}, \qquad \operatorname{Res} V_\lambda = \operatorname{Res} V_{\lambda'} = W_\lambda \quad \text{when } \lambda' \neq \lambda,$$

$$\operatorname{Ind} W_\lambda' = \operatorname{Ind} W_\lambda'' = V_\lambda, \qquad \operatorname{Res} V_\lambda = W_\lambda' \oplus W_\lambda'' \quad \text{when } \lambda' = \lambda.$$

Note that

$$\begin{aligned} &\#\{\text{self-conjugate representations of } \mathfrak{S}_d\} \\ &= \#\{\text{symmetric Young diagrams}\} \\ &= \#\{\text{split pairs of conjugacy classes in } \mathfrak{A}_d\} \\ &= \#\{\text{conjugacy classes in } \mathfrak{S}_d \text{ breaking into two classes in } \mathfrak{A}_d\}. \end{aligned}$$

Now a conjugacy class of an element written as a product of disjoint cycles is split if and only if there is no odd permutation commuting with it, which is equivalent to all the cycles having odd length, and no two cycles having the same length. So the number of self-conjugate representations is the number of partitions of d as a sum of distinct odd numbers. In fact, there is a natural correspondence between these two sets: any such partition corresponds to a symmetric Young diagram, assembling hooks as indicated:



If λ is the partition, the lengths of the cycles in the corresponding split conjugacy classes are $q_1 = 2\lambda_1 - 1$, $q_2 = 2\lambda_2 - 3$, $q_3 = 2\lambda_3 - 5, \ldots$.

For a self-conjugate partition λ, let χ'_λ and χ''_λ denote the characters of W'_λ and W''_λ, and let c and c' be a pair of split conjugacy classes, consisting of cycles of odd lengths $q_1 > q_2 > \cdots > q_r$. The following proposition of Frobenius completes the description of the character table of $\mathfrak{A}_d$.

Proposition 5.3. (1) *If c and c′ do not correspond to the partition λ, then*

$$\chi'_\lambda(c) = \chi'_\lambda(c') = \chi''_\lambda(c) = \chi''_\lambda(c') = \tfrac{1}{2}\chi_\lambda(c \cup c').$$

(2) *If c and c′ correspond to λ, then*

$$\chi'_\lambda(c) = \chi''_\lambda(c') = x, \qquad \chi'_\lambda(c') = \chi''_\lambda(c) = y,$$

with x and y the two numbers

$$\tfrac{1}{2}\left((-1)^m \pm \sqrt{(-1)^m q_1 \cdot \ldots \cdot q_r}\right),$$

and $m = \frac{1}{2}(d - r) = \frac{1}{2}\sum(q_i - 1) \equiv \frac{1}{2}(\prod q_i - 1) \pmod 2$.

For example, if $d = 4$ and $\lambda = (2, 2)$, we have $r = 2$, $q_1 = 3$, $q_2 = 1$, and x and y are the cube roots of unity; the representations W'_λ and W''_λ are the representations labeled U' and U'' in the table in §2.3. For $d = 5$, $\lambda = (3, 1, 1)$, $r = 1$, $q_1 = 5$, and we find the representations called Y and Z in §3.1. For $d \le 7$, there is at most one split pair, so the character table can be derived from orthogonality alone.

Note that since only one pair of character values is not taken care of by the first case of Frobenius's formula, the choice of which representation is W'_λ and which W''_λ is equivalent to choosing the plus and minus sign in (2). Note also that the integer m occurring in (2) is the number of squares above the diagonal in the Young diagram of λ.

We outline a proof of the proposition as an exercise:

Exercise 5.4*. *Step 1.* Let $q = (q_1 > \cdots > q_r)$ be a sequence of positive odd integers adding to d, and let $c' = c'(q)$ and $c'' = c''(q)$ be the corresponding conjugacy classes in $\mathfrak{A}_d$. Let λ be a self-conjugate partition of d, and let χ'_λ and χ''_λ be the corresponding characters of $\mathfrak{A}_d$. Assume that χ'_λ and χ''_λ take on the same values on each element of $\mathfrak{A}_d$ that is not in c' or c''. Let $u = \chi'_\lambda(c') = \chi''_\lambda(c'')$ and $v = \chi'_\lambda(c'') = \chi''_\lambda(c')$.

(i) Show that u and v are real when $m = \frac{1}{2}\Sigma(q_i - 1)$ is even, and $\bar{u} = v$ when m is odd.

(ii) Let $\vartheta = \chi'_\lambda - \chi''_\lambda$. Deduce from the equation $(\vartheta, \vartheta) = 2$ that $|u - v|^2 = q_1 \cdot \ldots \cdot q_r$.

(iii) Show that λ is the partition that corresponds to q and that $u + v = (-1)^m$, and deduce that u and v are the numbers specified in (2) of the proposition.

Step 2. Prove the proposition by induction on d, and for fixed d, look at that q which has smallest q_1, and for which some character has values on the classes $c'(q)$ and $c''(q)$ other than those prescribed by the proposition.

(i) If $r = 1$, so $q_1 = d = 2m + 1$, the corresponding self-conjugate partition is $\lambda = (m + 1, 1, \ldots, 1)$. By induction, Step 1 applies to χ'_λ and χ''_λ.

(ii) If $r > 1$, consider the imbedding $H = \mathfrak{A}_{q_1} \times \mathfrak{A}_{d-q_1} \subset G = \mathfrak{A}_d$, and let X' and X'' be the representations of G induced from the representations $W'_1 \boxtimes W'_2$ and $W''_1 \boxtimes W'_2$, where W'_1 and W''_1 are the representations of $\mathfrak{A}_{q_1}$ corresponding to q_1, i.e., to the self-conjugate partition $(\frac{1}{2}(q_1 - 1), 1, \ldots, 1)$ of q_1; W'_2 is one of the representations of $\mathfrak{A}_{d-q_1}$ corresponding to $(q_2, \ldots, q_r)$; and $\boxtimes$ denotes the external tensor product (see Exercise 2.36). Show that X' and X'' are conjugate representations of $\mathfrak{A}_d$, and their characters χ' and χ'' take equal values on each pair of split conjugacy classes, with the exception of $c'(q)$ and $c''(q)$, and compute the values of these characters on $c'(q)$ and $c''(q)$.

(iii) Let $\vartheta = \chi' - \chi''$, and show that $(\vartheta, \vartheta) = 2$. Decomposing X' and X'' into their irreducible pieces, deduce that $X' = Y \oplus W'_\lambda$ and $X'' = Y \oplus W''_\lambda$ for some self-conjugate representation Y and some self-conjugate partition λ of d.

(iv) Apply Step 1 to the characters χ'_λ and χ''_λ, and conclude the proof.

Exercise 5.5*. Show that if $d > 6$, the only irreducible representations of $\mathfrak{A}_d$ of dimension less than d are the trivial representation and the $(n - 1)$-dimensional restriction of the standard representation of $\mathfrak{S}_d$. Find the exceptions for $d \leq 6$.

We have worked out the character tables for all $\mathfrak{S}_d$ and $\mathfrak{A}_d$ for $d \leq 5$. With the formulas of Frobenius, an interested reader can construct the tables for a few more d—until the number of partitions of d becomes large.

§5.2. Representations of $GL_2(\mathbb{F}_q)$ and $SL_2(\mathbb{F}_q)$

The groups $GL_2(\mathbb{F}_q)$ of invertible 2×2 matrices with entries in the finite field $\mathbb{F}_q$ with q elements, where q is a prime power, form another important series of finite groups, as do their subgroups $SL_2(\mathbb{F}_q)$ consisting of matrices of determinant one. The quotient $PGL_2(\mathbb{F}_q) = GL_2(\mathbb{F}_q)/\mathbb{F}_q^*$ is the automorphism group of the finite projective line $\mathbb{P}^1(\mathbb{F}_q)$. The quotients $PSL_2(\mathbb{F}_q) = SL_2(\mathbb{F}_q)/\{\pm 1\}$ are simple groups if $q \neq 2, 3$ (Exercise 5.9). In this section we sketch the character theory of these groups.

We begin with $G = GL_2(\mathbb{F}_q)$. There are several key subgroups:

$$G \supset B = \left\{\begin{pmatrix} a & b \\ 0 & d \end{pmatrix}\right\} \supset N = \left\{\begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}\right\}.$$

(This "Borel subgroup" B and the group of upper triangular unipotent matrices N will reappear when we look at Lie groups.) Since G acts transitively on the projective line $\mathbb{P}^1(\mathbb{F}_q)$, with B the isotropy group of the point (1:0), we have

$$|G| = |B| \cdot |\mathbb{P}^1(\mathbb{F}_q)| = (q - 1)^2 q(q + 1).$$

We will also need the diagonal subgroup

$$D = \left\{\begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix}\right\} = \mathbb{F}^* \times \mathbb{F}^*,$$

where we write $\mathbb{F}$ for $\mathbb{F}_q$. Let $\mathbb{F}' = \mathbb{F}_{q^2}$ be the extension of $\mathbb{F}$ of degree two, unique up to isomorphism. We can identify $GL_2(\mathbb{F}_q)$ as the group of all $\mathbb{F}$-linear invertible endomorphisms of $\mathbb{F}'$. This makes evident a large cyclic subgroup $K = (\mathbb{F}')^*$ of G. At least if q is odd, we may make this isomorphism explicit by choosing a generator ε for the cyclic group $\mathbb{F}^*$ and choosing a square root $\sqrt{\varepsilon}$ in $\mathbb{F}'$. Then 1 and $\sqrt{\varepsilon}$ form a basis for $\mathbb{F}'$ as a vector space over $\mathbb{F}$, so we can make the identification:

$$K = \left\{\begin{pmatrix} x & \varepsilon y \\ y & x \end{pmatrix}\right\} \cong (\mathbb{F}')^*, \qquad \begin{pmatrix} x & \varepsilon y \\ y & x \end{pmatrix} \leftrightarrow \zeta = x + y\sqrt{\varepsilon};$$

K is a cyclic subgroup of G of order $q^2 - 1$. We often make this identification, leaving it as an exercise to make the necessary modifications in case q is even.

The conjugacy classes in G are easily found:

Representative	No. Elements in Class	No. Classes
$a_x = \begin{pmatrix} x & 0 \\ 0 & x \end{pmatrix}$	1	$q - 1$
$b_x = \begin{pmatrix} x & 1 \\ 0 & x \end{pmatrix}$	$q^2 - 1$	$q - 1$
$c_{x,y} = \begin{pmatrix} x & 0 \\ 0 & y \end{pmatrix}, x \neq y$	$q^2 + q$	$\frac{(q-1)(q-2)}{2}$
$d_{x,y} = \begin{pmatrix} x & \varepsilon y \\ y & x \end{pmatrix}, y \neq 0$	$q^2 - q$	$\frac{q(q-1)}{2}$

Here $c_{x,y}$ and $c_{y,x}$ are conjugate by $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, and $d_{x,y}$ and $d_{x,-y}$ are conjugate by any $\begin{pmatrix} a & -\varepsilon c \\ c & -a \end{pmatrix}$. To count the number of elements in the conjugacy class of b_x, look at the action of G on this class by conjugation; the isotropy group is $\left\{\begin{pmatrix} a & b \\ 0 & a \end{pmatrix}\right\}$, so the number of elements in the class is the index of this group in G, which is $q^2 - 1$. Similarly the isotropy group for $c_{x,y}$ is D, and the isotropy group for $d_{x,y}$ is K. To see that the classes are disjoint, consider the eigenvalues and the Jordan canonical forms. Since they account for $|G|$ elements, the list is complete.

There are $q^2 - 1$ conjugacy classes, so we must find the same number of irreducible representations. Consider first the permutation representation of G on $\mathbb{P}^1(\mathbb{F})$, which has dimension $q + 1$. It contains the trivial representation;

let V be the complementary q-dimensional representation. The values of the character χ of V on the four types of conjugacy classes are $\chi(a_x) = q$, $\chi(b_x) = 0$, $\chi(c_{x,y}) = 1$, $\chi(d_{x,y}) = -1$, which we display as the table:

$$V: \quad q \quad 0 \quad 1 \quad -1$$

Since $(\chi, \chi) = 1$, V is irreducible.

For each of the $q - 1$ characters $\alpha\colon \mathbb{F}^* \to \mathbb{C}^*$ of $\mathbb{F}^*$, we have a one-dimensional representation U_α of G defined by $U_\alpha(g) = \alpha(\det(g))$. We also have the representations $V_\alpha = V \otimes U_\alpha$. The values of the characters of these representations are

$$\begin{array}{lllll} U_\alpha\colon & \alpha(x)^2 & \alpha(x)^2 & \alpha(x)\alpha(y) & \alpha(x^2 - \varepsilon y^2) \\ V_\alpha\colon & q\alpha(x)^2 & 0 & \alpha(x)\alpha(y) & -\alpha(x^2 - \varepsilon y^2) \end{array}$$

Note that if we identify $\begin{pmatrix} x & \varepsilon y \\ y & x \end{pmatrix}$ with $\zeta = x + y\sqrt{\varepsilon}$ in $\mathbb{F}'$, then

$$x^2 - \varepsilon y^2 = \det\begin{pmatrix} x & \varepsilon y \\ y & x \end{pmatrix} = \mathrm{Norm}_{\mathbb{F}'/\mathbb{F}}(\zeta) = \zeta \cdot \zeta^q = \zeta^{q+1}.$$

The next place to look for representations is at those that are induced from large subgroups. For each pair α, β of characters of $\mathbb{F}^*$, there is a character of the subgroup B:

$$B \to B/N = D = \mathbb{F}^* \times \mathbb{F}^* \to \mathbb{C}^* \times \mathbb{C}^* \xrightarrow{\cdot} \mathbb{C}^*,$$

which takes $\begin{pmatrix} a & b \\ 0 & d \end{pmatrix}$ to $\alpha(a)\beta(d)$. Let $W_{\alpha,\beta}$ be the representation induced from B to G by this representation; this is a representation of dimension $[G : B] = q + 1$. By Exercise 3.19 its character values are found to be:

$$W_{\alpha,\beta}\colon \quad (q + 1)\alpha(x)\beta(x) \qquad \alpha(x)\beta(x) \qquad \alpha(x)\beta(y) + \alpha(y)\beta(x) \qquad 0$$

We see from this that $W_{\alpha,\beta} \cong W_{\beta,\alpha}$, that $W_{\alpha,\alpha} \cong U_\alpha \oplus V_\alpha$, and that for $\alpha \neq \beta$ the representation is irreducible. This gives $\frac{1}{2}(q - 1)(q - 2)$ more irreducible representations, of dimension $q + 1$.

Comparing with the list of conjugacy classes, we see that there are $\frac{1}{2}q(q - 1)$ irreducible characters left to be found. A natural way to find new characters is to induce characters from the cyclic subgroup K. For a representation

$$\varphi\colon K = (\mathbb{F}')^* \to \mathbb{C}^*,$$

the character values of the induced representation of dimension $[G : K] = q^2 - 1$ are

$$\mathrm{Ind}(\varphi)\colon \quad q(q - 1)\varphi(x) \qquad 0 \qquad 0 \qquad \varphi(\zeta) + \varphi(\zeta)^q$$

Here again $\zeta = x + y\sqrt{\varepsilon} \in K = (\mathbb{F}')^*$. Note that $\mathrm{Ind}(\varphi^q) \cong \mathrm{Ind}(\varphi)$, so the representations $\mathrm{Ind}(\varphi)$ for $\varphi^q \neq \varphi$ give $\frac{1}{2}q(q - 1)$ different representations.

However, these represenations are not irreducible: the character χ of $\mathrm{Ind}(\varphi)$ satisfies $(\chi, \chi) = q - 1$ if $\varphi^q \neq \varphi$, and otherwise $(\chi, \chi) = q$. We will have to work a little harder to get irreducible representations from these $\mathrm{Ind}(\varphi)$.

Another attempt to find more representations is to look inside tensor products of representations we know. We have $V_\alpha \otimes U_\gamma = V_{\alpha\gamma}$, and $W_{\alpha,\beta} \otimes U_\gamma \cong W_{\alpha\gamma,\beta\gamma}$, so there are no new ones to be found this way. But tensor products of the V_α's and $W_{\alpha,\beta}$'s are more promising. For example, $V \otimes W_{\alpha,1}$ has character values:

$$V \otimes W_{\alpha,1}: \qquad q(q+1)\alpha(x) \qquad 0 \qquad \alpha(x) + \alpha(y) \qquad 0$$

We can calculate some inner products of these characters with each other to estimate how many irreducible representations each contains, and how many they have in common. For example,

$$\begin{aligned}
(\chi_{V \otimes W_{\alpha,1}}, \chi_{W_{\alpha,1}}) &= 2, \\
(\chi_{\mathrm{Ind}(\varphi)}, \chi_{W_{\alpha,1}}) &= 1 \quad \text{if } \varphi|_{\mathbb{F}^*} = \alpha, \\
(\chi_{V \otimes W_{\alpha,1}}, \chi_{V \otimes W_{\alpha,1}}) &= q + 3, \\
(\chi_{V \otimes W_{\alpha,1}}, \chi_{\mathrm{Ind}(\varphi)}) &= q \quad \text{if } \varphi|_{\mathbb{F}^*} = \alpha,
\end{aligned}$$

Comparing with the formula $(\chi_{\mathrm{Ind}(\varphi)}, \chi_{\mathrm{Ind}(\varphi)}) = q - 1$, one deduces that $V \otimes W_{\alpha,1}$ and $\mathrm{Ind}(\varphi)$ contain many of the same representations. With any luck, $\mathrm{Ind}(\varphi)$ and $W_{\alpha,1}$ should both be contained in $V \otimes W_{\alpha,1}$. This guess is easily confirmed; the virtual character

$$\chi_\varphi = \chi_{V \otimes W_{\alpha,1}} - \chi_{W_{\alpha,1}} - \chi_{\mathrm{Ind}(\varphi)}$$

takes values $(q-1)\alpha(x)$, $-\alpha(x)$, 0, and $-(\varphi(\zeta) + \varphi(\zeta)^q)$ on the four types of conjugacy classes. Therefore, $(\chi_\varphi, \chi_\varphi) = 1$, and $\chi_\varphi(1) = q - 1 > 0$, so χ_φ is, in fact, the character of an irreducible subrepresentation of $V \otimes W_{\alpha,1}$ of dimension $q - 1$. We denote this representation by X_φ. These $\frac{1}{2}q(q-1)$ representations, for $\varphi \neq \varphi^q$, and with $X_\varphi = X_{\varphi^q}$, therefore complete the list of irreducible representations for $GL_2(\mathbb{F})$. The character table is

	1	q^2-1	q^2+q	q^2-q
$GL_2(\mathbb{F}_q)$	$a_x = \begin{pmatrix} x & 0 \\ 0 & x \end{pmatrix}$	$b_x = \begin{pmatrix} x & 1 \\ 0 & x \end{pmatrix}$	$c_{x,y} = \begin{pmatrix} x & 0 \\ 0 & y \end{pmatrix}$	$d_{x,y} = \begin{pmatrix} x & \varepsilon y \\ y & x \end{pmatrix} = \zeta$
U_α	$\alpha(x^2)$	$\alpha(x^2)$	$\alpha(xy)$	$\alpha(\zeta^q)$
V_α	$q\alpha(x^2)$	0	$\alpha(xy)$	$-\alpha(\zeta^q)$
$W_{\alpha,\beta}$	$(q+1)\alpha(x)\beta(x)$	$\alpha(x)\beta(x)$	$\alpha(x)\beta(y) + \alpha(y)\beta(x)$	0
X_φ	$(q-1)\varphi(x)$	$-\varphi(x)$	0	$-(\varphi(\zeta) + \varphi(\zeta^q))$

Exercise 5.6. Find the multiplicity of each irreducible representation in the representations $V \otimes W_{\alpha,1}$ and $\mathrm{Ind}(\varphi)$.

Exercise 5.7. Find the character table of $PGL_2(\mathbb{F}) = GL_2(\mathbb{F})/\mathbb{F}^*$. Note that its characters are just the characters of $GL_2(\mathbb{F})$ that take the same values on elements equivalent mod $\mathbb{F}^*$.

We turn next to the subgroup $SL_2(\mathbb{F}_q)$ of 2×2 matrices of determinant one, with q odd. The conjugacy classes, together with the number of elements in each conjugacy class, and the number of conjugacy classes of each type, are

	Representative	No. Elements in Class	No. Classes
(1)	$e = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	1	1
(2)	$-e = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$	1	1
(3)	$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$	$\dfrac{q^2-1}{2}$	1
(4)	$\begin{pmatrix} 1 & \varepsilon \\ 0 & 1 \end{pmatrix}$	$\dfrac{q^2-1}{2}$	1
(5)	$\begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix}$	$\dfrac{q^2-1}{2}$	1
(6)	$\begin{pmatrix} -1 & \varepsilon \\ 0 & -1 \end{pmatrix}$	$\dfrac{q^2-1}{2}$	1
(7)	$\begin{pmatrix} x & 0 \\ 0 & x^{-1} \end{pmatrix}, x \neq \pm 1$	$q(q+1)$	$\dfrac{q-3}{2}$
(8)	$\begin{pmatrix} x & y \\ \varepsilon y & x \end{pmatrix}, x \neq \pm 1$	$q(q-1)$	$\dfrac{q-1}{2}$

The verifications are very much as we did for $GL_2(\mathbb{F}_q)$. In (7), the classes of $\begin{pmatrix} x & 0 \\ 0 & x^{-1} \end{pmatrix}$ and $\begin{pmatrix} x^{-1} & 0 \\ 0 & x \end{pmatrix}$ are the same. In (8), the classes for (x, y) and $(x, -y)$ are the same; as before, a better labeling is by the element ζ in the cyclic group

$$C = \{\zeta \in (\mathbb{F}')^*: \zeta^{q+1} = 1\};$$

the elements ± 1 are not used, and the classes of ζ and ζ^{-1} are the same.

The total number of conjugacy classes is $q + 4$, so we turn to the task of finding $q + 4$ irreducible representations. We first see what we get by restricting representations from $GL_2(\mathbb{F}_q)$. Since we know the characters, there is no problem working this out, and we simply state the results:

(1) The U_α all restrict to the trivial representation U. Hence, if we restrict any representation, we will get the same for all tensor products by U_α's.

(2) The restriction V of the V_α's is irreducible.
(3) The restriction W_α of $W_{\alpha,1}$ is irreducible if $\alpha^2 \neq 1$, and $W_\alpha \cong W_\beta$ when $\beta = \alpha$ or $\beta = \alpha^{-1}$. These give $\frac{1}{2}(q-3)$ irreducible representations of dimension $q+1$.
(3′) Let τ denote the character of $\mathbb{F}^*$ with $\tau^2 = 1, \tau \neq 1$. The restriction of $W_{\tau,1}$ is the sum of two distinct irreducible representations, which we denote W' and W''.
(4) The restriction of X_φ depends only on the restriction of φ to the subgroup C, and φ and φ^{-1} determine the same representation. The representation is irreducible if $\varphi^2 \neq 1$. This gives $\frac{1}{2}(q-1)$ irreducible representations of dimension $q-1$.
(4′) If ψ denotes the character of C with $\psi^2 = 1$, $\psi \neq 1$, the restriction of X_ψ is the sum of two distinct irreducible representations, which we denote X' and X''.

Altogether this list gives $q+4$ distinct irreducible representations, and it is therefore the complete list. To finish the character table, the problem is to describe the four representations W', W'', X', and X''. Since we know the sum of the squares of the dimensions of all representations, we can deduce that the sum of the squares of these four representations is q^2+1, which is only possible if the first two have dimension $\frac{1}{2}(q+1)$ and the other two $\frac{1}{2}(q-1)$. This is similar to what we saw happens for restrictions of representations to subgroups of index two. Although the index here is larger, we can use what we know about index two subgroups by finding a subgroup H of index two in $GL_2(\mathbb{F}_q)$ that contains $SL_2(\mathbb{F}_q)$, and analyzing the restrictions of these four representations to H.

For H we take the matrices in $GL_2(\mathbb{F}_q)$ whose determinant is a square. The representatives of the conjugacy classes are the same as those for $GL_2(\mathbb{F}_q)$, including, of course, only those representatives whose determinant is a square, but we must add classes represented by the elements $\begin{pmatrix} x & \varepsilon \\ 0 & x \end{pmatrix}$, $x \in \mathbb{F}^*$. These are conjugate to the elements $\begin{pmatrix} x & 1 \\ 0 & x \end{pmatrix}$ in $GL_2(\mathbb{F}_q)$, but not in H. These are the $q-1$ split conjugacy classes. The procedure of the preceding section can be used to work out all the representations of H, but we need only a little of this.

Note that the sign representation U' from G/H is U_τ, so that $W_{\tau,1} \cong W_{\tau,1} \otimes U'$ and $X_\psi \cong X_\psi \otimes U'$; their restrictions to H split into sums of conjugate irreducible representations of half their dimensions. This shows these representations stay irreducible on restriction from H to $SL_2(\mathbb{F}_q)$, so that W' and W'' are conjugate representations of dimension $\frac{1}{2}(q+1)$, and X' and X'' are conjugate representations of dimension $\frac{1}{2}(q-1)$. In addition, we know that their character values on all nonsplit conjugacy classes are the same as half the characters of the representations $W_{\tau,1}$ and X_ψ, respectively. This is all the information we need to finish the character table. Indeed, the only values not covered by this discussion are

	$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & \varepsilon \\ 0 & 1 \end{pmatrix}$	$\begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix}$	$\begin{pmatrix} -1 & \varepsilon \\ 0 & -1 \end{pmatrix}$
W'	s	t	s'	t'
W''	t	s	t'	s'
X'	u	v	u'	v'
X''	v	u	v'	u'

The first two rows are determined as follows. We know that $s + t = \chi_{W_{\tau,1}}\left(\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\right) = 1$. In addition, since $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$ is conjugate to $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ if q is congruent to 1 modulo 4, and to $\begin{pmatrix} 1 & \varepsilon \\ 0 & 1 \end{pmatrix}$ otherwise, and since $\chi(g^{-1}) = \overline{\chi(g)}$ for any character, we conclude that s and t are real if $q \equiv 1 \bmod(4)$, and $s = \bar{t}$ if $q \equiv 3 \bmod(4)$. In addition, since $-e$ acts as the identity or minus the identity for any irreducible representation (Schur's lemma),

$$\chi(-g) = \chi(g) \cdot \chi(1)/\chi(-e)$$

for any irreducible character χ. This gives the relations $s' = \tau(-1)s$ and $t' = \tau(-1)t$. Finally, applying the equation $(\chi, \chi) = 1$ to the character of W' gives a formula for $s\bar{t} + t\bar{s}$. Solving these equations gives $s, t = \frac{1}{2} \pm \frac{1}{2}\sqrt{\omega q}$, where $\omega = \tau(-1)$ is 1 or -1 according as $q \equiv 1$ or $3 \bmod(4)$. Similarly one computes that u and v are $-\frac{1}{2} \pm \frac{1}{2}\sqrt{\omega q}$. This concludes the computations needed to write out the character table.

Exercise 5.8. By considering the action of $SL_2(\mathbb{F}_q)$ on the set $\mathbb{P}^1(\mathbb{F}_q)$, show that $SL_2(\mathbb{F}_2) \cong \mathfrak{S}_3$, $PSL_2(\mathbb{F}_3) \cong \mathfrak{A}_4$, and $SL_2(\mathbb{F}_4) \cong \mathfrak{A}_5$.

Exercise 5.9*. Use the character table for $SL_2(\mathbb{F}_q)$ to show that $PSL_2(\mathbb{F}_q)$ is a simple group if q is odd and greater than 3.

Exercise 5.10. Compute the character table of $PSL_2(\mathbb{F}_q)$, either by regarding it as a quotient of $SL_2(\mathbb{F}_q)$, or as a subgroup of index two in $PGL_2(\mathbb{F}_q)$.

Exercise 5.11*. Find the conjugacy classes of $GL_3(\mathbb{F}_q)$, and compute the characters of the permutation representations obtained by the action of $GL_3(\mathbb{F}_q)$ on (i) the projective plane $\mathbb{P}^2(\mathbb{F}_q)$ and (ii) the "flag variety" consisting of a point on a line in $\mathbb{P}^2(\mathbb{F}_q)$. Show that the first is irreducible and that the second is a sum of the trivial representation, two copies of the first representation, and an irreducible representation.

Although the characters of the above groups were found by the early pioneers in representation theory, actually producing the representations in a natural way is more difficult. There has been a great deal of work extending this story to $GL_n(\mathbb{F}_q)$ and $SL_n(\mathbb{F}_q)$ for $n > 2$ (cf. [Gr]), and for corresponding groups, called finite Chevalley groups, related to other Lie groups. For some hints in this direction see [Hu3], as well as [Ti2]. Since all but a finite number of finite simple groups are now known to arise this way (or are cyclic or alternating groups, whose characters we already know), such representations play a fundamental role in group theory. In recent work their Lie-theoretic origins have been exploited to produce their representations, but to tell this story would go far beyond the scope of these lecture(r)s.

LECTURE 6

Weyl's Construction

In this lecture we introduce and study an important collection of functors generalizing the symmetric powers and exterior powers. These are defined simply in terms of the Young symmetrizers c_λ introduced in §4: given a representation V of an arbitrary group G, we consider the dth tensor power of V, on which both G and the symmetric group on d letters act. We then take the image of the action of c_λ on $V^{\otimes d}$; this is again a representation of G, denoted $\mathbb{S}_\lambda(V)$. This gives us a way of generating new representations, whose main application will be to Lie groups: for example, we will generate all representations of $\mathrm{SL}_n\mathbb{C}$ by applying these to the standard representation $\mathbb{C}^n$ of $\mathrm{SL}_n\mathbb{C}$. While it may be easiest to read this material while the definitions of the Young symmetrizers are still fresh in the mind, the construction will not be used again until §15, so that this lecture can be deferred until then.

§6.1. Schur Functors and Their Characters

For any finite-dimensional complex vector space V, we have the canonical decomposition

$$V \otimes V = \mathrm{Sym}^2 V \oplus \wedge^2 V.$$

The group $\mathrm{GL}(V)$ acts on $V \otimes V$, and this is, as we shall soon see, the decomposition of $V \otimes V$ into a direct sum of irreducible $\mathrm{GL}(V)$-representations. For the next tensor power,

$$V \otimes V \otimes V = \mathrm{Sym}^3 V \oplus \wedge^3 V \oplus \text{ another space.}$$

We shall see that this other space is a sum of two copies of an irreducible

$GL(V)$-representation. Just as $\mathrm{Sym}^d V$ and $\wedge^d V$ are images of symmetrizing operators from $V^{\otimes d} = V \otimes V \otimes \cdots \otimes V$ to itself, so are the other factors. The symmetric group $\mathfrak{S}_d$ acts on $V^{\otimes d}$, say on the right, by permuting the factors

$$(v_1 \otimes \cdots \otimes v_d) \cdot \sigma = v_{\sigma(1)} \otimes \cdots \otimes v_{\sigma(d)}.$$

This action commutes with the left action of $GL(V)$. For any partition λ of d we have from the last lecture a Young symmetrizer c_λ in $\mathbb{C}\mathfrak{S}_d$. We denote the image of c_λ on $V^{\otimes d}$ by $\mathbb{S}_\lambda V$:

$$\mathbb{S}_\lambda V = \mathrm{Im}(c_\lambda|_{V^{\otimes d}})$$

which is again a representation of $GL(V)$. We call the functor[1] $V \rightsquigarrow \mathbb{S}_\lambda V$ the *Schur functor* or *Weyl module*, or simply *Weyl's construction*, corresponding to λ. It was Schur who made the correspondence between representations of symmetric groups and representations of general linear groups, and Weyl who made the construction we give here.[2] We will give other descriptions later, cf. Exercise 6.14 and §15.5.

For example, the partition $d = d$ corresponds to the functor $V \rightsquigarrow \mathrm{Sym}^d V$, and the partition $d = 1 + \cdots + 1$ to the functor $V \rightsquigarrow \wedge^d V$.

We find something new for the partition $3 = 2 + 1$. The corresponding symmetrizer c_λ is

$$c_{(2,1)} = 1 + e_{(12)} - e_{(13)} - e_{(132)},$$

so the image of c_λ is the subspace of $V^{\otimes 3}$ spanned by all vectors

$$v_1 \otimes v_2 \otimes v_3 + v_2 \otimes v_1 \otimes v_3 - v_3 \otimes v_2 \otimes v_1 - v_3 \otimes v_1 \otimes v_2.$$

If $\wedge^2 V \otimes V$ is embedded in $V^{\otimes 3}$ by mapping

$$(v_1 \wedge v_3) \otimes v_2 \mapsto v_1 \otimes v_2 \otimes v_3 - v_3 \otimes v_2 \otimes v_1,$$

then the image of c_λ is the subspace of $\wedge^2 V \otimes V$ spanned by all vectors

$$(v_1 \wedge v_3) \otimes v_2 + (v_2 \wedge v_3) \otimes v_1.$$

It is not hard to verify that these vectors span the kernel of the canonical map from $\wedge^2 V \otimes V$ to $\wedge^3 V$, so we have

$$\mathbb{S}_{(2,1)} V = \mathrm{Ker}(\wedge^2 V \otimes V \to \wedge^3 V).$$

(This gives the missing factor in the decomposition of $V^{\otimes 3}$.)

Note that some of the $\mathbb{S}_\lambda V$ can be zero if V has small dimension. We will see that this is the case precisely when the number of rows in the Young diagram of λ is greater than the dimension of V.

[1] The functoriality means simply that a linear map $\varphi\colon V \to W$ of vector spaces determines a linear map $\mathbb{S}_\lambda(\varphi)\colon \mathbb{S}_\lambda V \to \mathbb{S}_\lambda W$, with $\mathbb{S}_\lambda(\varphi \circ \psi) = \mathbb{S}_\lambda(\varphi) \circ \mathbb{S}_\lambda(\psi)$ and $\mathbb{S}_\lambda(\mathrm{Id}_V) = \mathrm{Id}_{\mathbb{S}_\lambda V}$

[2] The notion goes by a variety of names and notations in the literature, depending on the context. Constructions differ markedly when not over a field of characteristic zero; and many authors now parametrize them by the conjugate partitions. Our choice of notation is guided by the correspondence between these functors and Schur polynomials, which we will see are their characters.

When $G = \mathrm{GL}(V)$, and for important subgroups $G \subset \mathrm{GL}(V)$, these $\mathbb{S}_\lambda V$ give many of the irreducible representations of G; we will come back to this later in the book. For now we can use our knowledge of symmetric group representations to prove a few facts about them—in particular, we show that they decompose the tensor powers $V^{\otimes d}$, and that they are irreducible representations of $\mathrm{GL}(V)$. We will also compute their characters; this will eventually be seen to be a special case of the Weyl character formula.

Any endomorphism g of V gives rise to an endomorphism of $\mathbb{S}_\lambda V$. In order to tell what representations we get, we will need to compute the trace of this endomorphism on $\mathbb{S}_\lambda V$; we denote this trace by $\chi_{\mathbb{S}_\lambda V}(g)$. For the computation, let $x_1, \ldots, x_k$ be the eigenvalues of g on V, $k = \dim V$. Two cases are easy. For $\lambda = (d)$,

$$\mathbb{S}_{(d)} V = \mathrm{Sym}^d V, \qquad \chi_{\mathbb{S}_{(d)} V}(g) = H_d(x_1, \ldots, x_k), \tag{6.1}$$

where $H_d(x_1, \ldots, x_k)$ is the complete symmetric polynomial of degree d. The definition of these symmetric polynomials is given in (A.1) of Appendix A. The truth of (6.1) is evident when g is a diagonal matrix, and its truth for the dense set of diagonalizable endomorphisms implies it for all endomorphisms; or one can see it directly by using the Jordan canonical form of g. For $\lambda = (1, \ldots, 1)$, we have similarly

$$\mathbb{S}_{(1,\ldots,1)} V = \textstyle\bigwedge^d V, \qquad \chi_{\mathbb{S}_{(1,\ldots,1)} V}(g) = E_d(x_1, \ldots, x_k), \tag{6.2}$$

with $E_d(x_1, \ldots, x_k)$ the elementary symmetric polynomial [see (A.3)]. The polynomials H_d and E_d are special cases of the *Schur polynomials*, which we denote by $S_\lambda = S_\lambda(x_1, \ldots, x_k)$. As λ varies over the partitions of d into at most k parts, these polynomials S_λ form a basis for the symmetric polynomials of degree d in these k variables. Schur polynomials are defined and discussed in Appendix A, especially (A.4)–(A.6). The above two formulas can be written

$$\chi_{\mathbb{S}_\lambda V}(g) = S_\lambda(x_1, \ldots, x_k) \quad \text{for } \lambda = (d) \text{ and } \lambda = (1, \ldots, 1).$$

We will show that this equation is valid for all λ:

Theorem 6.3. (1) *Let $k = \dim V$. Then $\mathbb{S}_\lambda V$ is zero if $\lambda_{k+1} \neq 0$. If $\lambda = (\lambda_1 \geq \cdots \geq \lambda_k \geq 0)$, then*

$$\dim \mathbb{S}_\lambda V = S_\lambda(1, \ldots, 1) = \prod_{1 \leq i < j \leq k} \frac{\lambda_i - \lambda_j + j - i}{j - i}.$$

(2) *Let m_λ be the dimension of the irreducible representation V_λ of $\mathfrak{S}_d$ corresponding to λ. Then*

$$V^{\otimes d} \cong \bigoplus_\lambda \mathbb{S}_\lambda V^{\otimes m_\lambda}.$$

(3) *For any $g \in \mathrm{GL}(V)$, the trace of g on $\mathbb{S}_\lambda V$ is the value of the Schur polynomial on the eigenvalues $x_1, \ldots, x_k$ of g on V:*

$$\chi_{\mathbb{S}_\lambda V}(g) = S_\lambda(x_1, \ldots, x_k).$$

(4) *Each* $\mathbb{S}_\lambda V$ *is an irreducible representation of* $\mathrm{GL}(V)$.

This theorem will be proved in the next section. Other formulas for the dimension of $\mathbb{S}_\lambda V$ are given in Exercises A.30 and A.31. The following is another:

Exercise 6.4*. Show that

$$\dim \mathbb{S}_\lambda V = \frac{m_\lambda}{d!} \prod (k - i + j) = \prod \frac{(k - i + j)}{h_{ij}},$$

where the products are over the d pairs (i, j) that number the row and column of boxes for λ, and h_{ij} is the hook number of the corresponding box.

Exercise 6.5. Show that $V^{\otimes 3} \cong \mathrm{Sym}^3 V \oplus \wedge^3 V \oplus (\mathbb{S}_{(2,1)} V)^{\oplus 2}$, and

$$V^{\otimes 4} \cong \mathrm{Sym}^4 V \oplus \wedge^4 V \oplus (\mathbb{S}_{(3,1)} V)^{\oplus 3} \oplus (\mathbb{S}_{(2,2)} V)^{\oplus 2} \oplus (\mathbb{S}_{(2,1,1)} V)^{\oplus 3}.$$

Compute the dimensions of each of the irreducible factors.

The proof of the theorem actually gives the following corollary:

Corollary 6.6. *If* $c \in \mathbb{C}\mathfrak{S}_d$, *and* $(\mathbb{C}\mathfrak{S}_d)\cdot c = \bigoplus_\lambda V_\lambda^{\oplus r_\lambda}$ *as representations of* $\mathfrak{S}_d$, *then there is a corresponding decomposition of* $\mathrm{GL}(V)$*-spaces*:

$$V^{\otimes d}\cdot c = \bigoplus_\lambda \mathbb{S}_\lambda V^{\oplus r_\lambda}.$$

If $x_1, \ldots, x_k$ *are the eigenvalues of an endomorphism of* V, *the trace of the induced endomorphism of* $V^{\otimes d}\cdot c$ *is* $\sum r_\lambda S_\lambda(x_1, \ldots, x_k)$.

If λ and μ are different partitions, each with at most $k = \dim V$ parts, the irreducible $\mathrm{GL}(V)$-spaces $\mathbb{S}_\lambda V$ and $\mathbb{S}_\mu V$ are not isomorphic. Indeed, their characters are the Schur polynomials S_λ and S_μ, which are different. More generally, at least for those representations of $\mathrm{GL}(V)$ which can be decomposed into a direct sum of copies of the represenations $\mathbb{S}_\lambda V$'s, *the representations are completely determined by their characters.* This follows immediately from the fact that the Schur polynomials are linearly independent.

Note, however, that we cannot hope to get *all* finite-dimensional irreducible representations of $\mathrm{GL}(V)$ this way, since the duals of these representations are not included. We will see in Lecture 15 that this is essentially the only omission. Note also that although the operation that takes representations of $\mathfrak{S}_d$ to representations of $\mathrm{GL}(V)$ preserves direct sums, the situation with respect to other linear algebra constructions such as tensor products is more complicated.

One important application of Corollary 6.6 is to the decomposition of a tensor product $\mathbb{S}_\lambda V \otimes \mathbb{S}_\mu V$ of two Weyl modules, with, say, λ a partition of

d and μ a partition of m. The result is

$$\mathbb{S}_\lambda V \otimes \mathbb{S}_\mu V \cong \bigoplus_\nu N_{\lambda\mu\nu} \mathbb{S}_\nu V; \tag{6.7}$$

here the sum is over partitions ν of $d + m$, and $N_{\lambda\mu\nu}$ are numbers determined by the *Littlewood–Richardson rule.* This is a rule that gives $N_{\lambda\mu\nu}$ as the number of ways to expand the Young diagram of λ, using μ in an appropriate way, to achieve the Young diagram for ν; see (A.8) for the precise formula. Two important special cases are easier to use and prove since they involve only the simpler Pieri formula (A.7). For $\mu = (m)$, we have

$$\mathbb{S}_\lambda V \otimes \mathrm{Sym}^m V \cong \bigoplus_\nu \mathbb{S}_\nu V, \tag{6.8}$$

the sum over all ν whose Young diagram is obtained by adding m boxes to the Young diagram of λ, with no two in the same column. Similarly for $\mu = (1, \ldots, 1)$,

$$\mathbb{S}_\lambda V \otimes \Lambda^m V = \bigoplus \mathbb{S}_\pi V, \tag{6.9}$$

the sum over all partitions π whose Young diagram is obtained from that of λ by adding m boxes, with no two in the same row.

To prove these formulas, we need only observe that

$$\begin{aligned}\mathbb{S}_\lambda V \otimes \mathbb{S}_\mu V &= V^{\otimes n} \cdot c_\lambda \otimes V^{\otimes m} \cdot c_\mu \\ &= V^{\otimes n} \otimes V^{\otimes m} \cdot (c_\lambda \otimes c_\mu) = V^{\otimes (n+m)} \cdot c,\end{aligned}$$

with $c = c_\lambda \otimes c_\mu \in \mathbb{C}\mathfrak{S}_d \otimes \mathbb{C}\mathfrak{S}_m = \mathbb{C}(\mathfrak{S}_d \times \mathfrak{S}_m) \subset \mathbb{C}\mathfrak{S}_{d+m}$. This proves that $\mathbb{S}_\lambda V \otimes \mathbb{S}_\mu V$ has a decomposition as in Corollary 6.6, and the coefficients are given by knowing the decomposition of the corresponding character. The character of a tensor product is the product of the characters of the factors; so this amounts to writing the product $S_\lambda S_\mu$ of Schur polynomials as a linear combination of Schur polynomials. This is done in Appendix A, and formulas (6.7), (6.8), and (6.9) follow from (A.8), (A.7), and Exercise A.32 (v), respectively.

For example, from $\mathrm{Sym}^d V \otimes V = \mathrm{Sym}^{d+1} V \oplus \mathbb{S}_{(d,1)} V$, it follows that

$$\mathbb{S}_{(d,1)} V = \mathrm{Ker}(\mathrm{Sym}^d V \otimes V \to \mathrm{Sym}^{d+1} V),$$

and similarly for the conjugate partition,

$$\mathbb{S}_{(2,1,\ldots,1)} V = \mathrm{Ker}(\Lambda^d V \otimes V \to \Lambda^{d+1} V).$$

Exercise 6.10*. One can also derive the preceding decompositions of tensor products directly from corresponding decompositions of representations of symmetric groups. Show that, in fact, $\mathbb{S}_\lambda V \otimes \mathbb{S}_\mu V$ corresponds to the "inner product" representation $V_\lambda \circ V_\mu$ of $\mathfrak{S}_{d+m}$ described in (4.41).

Exercise 6.11*. (a) The Littlewood–Richardson rule also comes into the decomposition of a Schur functor of a direct sum of vector spaces V and W. This generalizes the well-known identities

$$\operatorname{Sym}^n(V \oplus W) = \bigoplus_{a+b=n} (\operatorname{Sym}^a V \otimes \operatorname{Sym}^b W),$$

$$\Lambda^n(V \oplus W) = \bigoplus_{a+b=n} (\Lambda^a V \otimes \Lambda^b W).$$

Prove the general decomposition over $\mathrm{GL}(V) \times \mathrm{GL}(W)$:

$$\mathbb{S}_\nu(V \oplus W) = \bigoplus N_{\lambda\mu\nu}(\mathbb{S}_\lambda V \otimes \mathbb{S}_\mu W),$$

the sum over all partitions λ, μ such that the sum of the numbers partitioned by λ and μ is the number partitioned by ν. (To be consistent with Exercise 2.36 one should use the notation $\boxtimes$ for these "external" tensor products.)

(b) Similarly prove the formula for the Schur functor of a tensor product:

$$\mathbb{S}_\nu(V \otimes W) = \bigoplus C_{\lambda\mu\nu}(\mathbb{S}_\lambda V \otimes \mathbb{S}_\mu W),$$

where the coefficients $C_{\lambda\mu\nu}$ are defined in Exercise 4.51. In particular show that

$$\operatorname{Sym}^d(V \otimes W) = \bigoplus \mathbb{S}_\lambda V \otimes \mathbb{S}_\lambda W,$$

the sum over all partitions λ of d with at most $\dim V$ or $\dim W$ rows. Replacing W by W^*, this gives the decomposition for the space of polynomial functions of degree d on the space $\operatorname{Hom}(V, W)$ over $\mathrm{GL}(V) \times \mathrm{GL}(W)$. For variations on this theme, see [Ho3]. Similarly,

$$\Lambda^d(V \otimes W) = \bigoplus \mathbb{S}_\lambda V \otimes \mathbb{S}_{\lambda'} W,$$

the sum over partitions λ of d with at most $\dim V$ rows and at most $\dim W$ columns.

Exercise 6.12. Regarding

$$\mathrm{GL}_n\,\mathbb{C} = \mathrm{GL}_n\,\mathbb{C} \times \{1\} \subset \mathrm{GL}_n\,\mathbb{C} \times \mathrm{GL}_m\,\mathbb{C} \subset \mathrm{GL}_{n+m}\,\mathbb{C},$$

the preceding exercise shows how the restriction of a representation decomposes:

$$\operatorname{Res}(\mathbb{S}_\nu(\mathbb{C}^{n+m})) = \sum (N_{\lambda\mu\nu} \dim \mathbb{S}_\mu(\mathbb{C}^m))\mathbb{S}_\lambda(\mathbb{C}^n).$$

In particular, for $m = 1$, Pieri's formula gives

$$\operatorname{Res}(\mathbb{S}_\nu(\mathbb{C}^{n+1})) = \bigoplus \mathbb{S}_\lambda(\mathbb{C}^n),$$

the sum over all λ obtained from ν by removing any number of boxes from its Young diagram, with no two in any column.

Exercise 6.13*. Show that for any partition $\mu = (\mu_1, \ldots, \mu_r)$ of d,

$$\Lambda^{\mu_1} V \otimes \Lambda^{\mu_2} V \otimes \cdots \otimes \Lambda^{\mu_r} V \cong \bigoplus_\lambda K_{\lambda\mu} \mathbb{S}_{\lambda'} V,$$

where $K_{\lambda\mu}$ is the Kostka number and λ' the conjugate of λ.

Exercise 6.14*. Let $\mu = \lambda'$ be the conjugate partition. Put the factors of the dth tensor power $V^{\otimes d}$ in one-to-one correspondence with the squares of the Young diagram of λ. Show that $\mathbb{S}_\lambda V$ is the image of this composite map:

$$\bigotimes_i (\wedge^{\mu_i} V) \to \bigotimes_i (\otimes^{\mu_i} V) \to V^{\otimes d} \to \bigotimes_j (\otimes^{\lambda_j} V) \to \bigotimes_j (\mathrm{Sym}^{\lambda_j} V),$$

the first map being the tensor product of the obvious inclusions, the second grouping the factors of $V^{\otimes d}$ according to the columns of the Young diagram, the third grouping the factors according to the rows of the Young diagram, and the fourth the obvious quotient map. Alternatively, $\mathbb{S}_\lambda V$ is the image of a composite map

$$\bigotimes_i (\mathrm{Sym}^{\lambda_i} V) \to \bigotimes_i (\otimes^{\lambda_i} V) \to V^{\otimes d} \to \bigotimes_j (\otimes^{\mu_j} V) \to \bigotimes_j (\wedge^{\mu_j} V).$$

In particular, $\mathbb{S}_\lambda V$ can be realized as a subspace of tensors in $V^{\otimes d}$ that are invariant by automorphisms that preserve the rows of a Young tableau of λ, *or* a subspace that is anti-invariant under those that preserve the columns, but not both, cf. Exercise 4.48.

Problem 6.15*. The preceding exercise can be used to describe a basis for the space $\mathbb{S}_\lambda V$. Let $v_1, \ldots, v_k$ be a basis for V. For each semistandard tableau T on λ, one can use it to write down an element v_T in $\bigotimes_i (\wedge^{\mu_i} V)$; v_T is a tensor product of wedge products of basis elements, the ith factor in $\wedge^{\mu_i} V$ being the wedge product (in order) of those basis vectors whose indices occur in the ith column of T. The fact to be proved is that the images of these elements v_T under the first composite map of the preceding exercise form a basis for $\mathbb{S}_\lambda V$.

At the end of Lecture 15, using more representation theory than we have at the moment, we will work out a simple variation of the construction of $\mathbb{S}_\lambda V$ which will give quick proofs of refinements of the preceding exercise and problem.

Exercise 6.16*. The Pieri formula gives a decomposition

$$\mathrm{Sym}^d V \otimes \mathrm{Sym}^d V = \bigoplus \mathbb{S}_{(d+a, d-a)} V,$$

the sum over $0 \le a \le d$. The left-hand side decomposes into a direct sum of $\mathrm{Sym}^2(\mathrm{Sym}^d V)$ and $\wedge^2(\mathrm{Sym}^d V)$. Show that, in fact,

$$\mathrm{Sym}^2(\mathrm{Sym}^d V) = \mathbb{S}_{(2d,0)} V \oplus \mathbb{S}_{(2d-2,2)} V \oplus \mathbb{S}_{(2d-4,4)} V \oplus \cdots,$$

$$\wedge^2(\mathrm{Sym}^d V) = \mathbb{S}_{(2d-1,1)} V \oplus \mathbb{S}_{(2d-3,3)} V \oplus \mathbb{S}_{(2d-5,5)} V \oplus \cdots.$$

Similarly using the dual form of Pieri to decompose $\wedge^d V \otimes \wedge^d V$ into the sum $\bigoplus \mathbb{S}_\lambda V$, the sum over all $\lambda = (2, \ldots, 2, 1, \ldots, 1)$ consisting of $d - a$ 2's and 2a 1's, $0 \le a \le d$, show that $\mathrm{Sym}^2(\wedge^d V)$ is the sum of those factors with a even, and $\wedge^2(\wedge^d V)$ is the sum of those with a odd.

Exercise 6.17*. If λ and μ are any partitions, we can form the composite functor $\mathbb{S}_\mu(\mathbb{S}_\lambda V)$. The original "plethysm" problem—which remains very difficult in general—is to decompose these composites:

$$\mathbb{S}_\mu(\mathbb{S}_\lambda V) = \bigoplus_\nu M_{\lambda\mu\nu} \mathbb{S}_\nu V,$$

the sum over all partitions ν of dm, where λ is a partition of d and μ is a partition of m. The preceding exercise carried out four special cases of this.

(a) Show that there always *exists* such a decomposition for some non-negative integers $M_{\lambda\mu\nu}$ by constructing an element c in $\mathbb{C}\mathfrak{S}_{dm}$, depending on λ and μ, such that $\mathbb{S}_\mu(\mathbb{S}_\lambda V)$ is $V^{\otimes dm} \cdot c$.

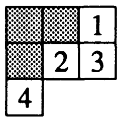
(b) Compute $\mathrm{Sym}^2(\mathbb{S}_{(2,2)} V)$ and $\bigwedge^2(\mathbb{S}_{(2,2)} V)$.

Exercise 6.18* "Hermite reciprocity." Show that if $\dim V = 2$ there are isomorphisms

$$\mathrm{Sym}^p(\mathrm{Sym}^q V) \cong \mathrm{Sym}^q(\mathrm{Sym}^p V)$$

of $\mathrm{GL}(V)$-representations, for all p and q.

Exercise 6.19*. Much of the story about Young diagrams and representations of symmetric and general linear groups can be generalized to *skew Young diagrams*, which are the differences of two Young diagrams. If λ and μ are partitions with $\mu_i \leq \lambda_i$ for all i, λ/μ denotes the complement of the Young diagram for μ in that of λ. For example, if $\lambda = (3, 3, 1)$ and $\mu = (2, 1)$, λ/μ is the numbered part of



To each λ/μ we have a *skew Schur function* $S_{\lambda/\mu}$, which can be defined by any of several generalizations of constructions of ordinary Schur functions. Using the notation of Appendix A, the following definitions are equivalent:

(i) $$S_{\lambda/\mu} = |H_{\lambda_i - \mu_j - i + j}|,$$

(ii) $$S_{\lambda/\mu} = |E_{\lambda'_i - \mu'_j - i + j}|,$$

(iii) $$S_{\lambda/\mu} = \sum m_a x_1^{a_1} \cdot \ldots \cdot x_k^{a_k},$$

where m_a is the number of ways to number the boxes of λ/μ with a_1 1's, a_2 2's, $\ldots$, a_k k's, with nondecreasing rows and strictly increasing columns.

In terms of ordinary Schur polynomials, we have

(iv) $$S_{\lambda/\mu} = \sum N_{\mu\nu\lambda} S_\nu,$$

where $N_{\mu\nu\lambda}$ is the Littlewood–Richardson number.

Each λ/μ determines elements $a_{\lambda/\mu}$, $b_{\lambda/\mu}$, and Young symmetrizers $c_{\lambda/\mu} = a_{\lambda/\mu}b_{\lambda/\mu}$ in $A = \mathbb{C}\mathfrak{S}_d$, $d = \sum \lambda_i - \mu_i$, exactly as in §4.1, and hence a representation denoted $V_{\lambda/\mu} = Ac_{\lambda/\mu}$ of $\mathfrak{S}_d$. Equivalently, $V_{\lambda/\mu}$ is the image of the map $Ab_{\lambda/\mu} \to Aa_{\lambda/\mu}$ given by right multiplication by $a_{\lambda/\mu}$, or the image of the map $Aa_{\lambda/\mu} \to Ab_{\lambda/\mu}$ given by right multiplication by $b_{\lambda/\mu}$. The decomposition of $V_{\lambda/\mu}$ into irreducible representations is

(v) $$V_{\lambda/\mu} = \sum N_{\mu\nu\lambda} V_\nu.$$

Similarly there are *skew Schur functors* $\mathbb{S}_{\lambda/\mu}$, which take a vector space V to the image of $c_{\lambda/\mu}$ on $V^{\otimes d}$; equivalently, $\mathbb{S}_{\lambda/\mu} V$ is the image of a natural map (generalizing that in the Exercise 6.14)

(vi) $$\bigotimes_i (\textstyle\bigwedge^{\lambda_i' - \mu_i'} V) \to V^{\otimes d} \to \bigotimes_j (\mathrm{Sym}^{\lambda_j - \mu_j} V),$$

or

(vii) $$\bigotimes_i (\mathrm{Sym}^{\lambda_i - \mu_i} V) \to V^{\otimes d} \to \bigotimes_j (\textstyle\bigwedge^{\lambda_j' - \mu_j'} V).$$

Given a basis $v_1, \ldots, v_k$ for V and a standard tableau T on λ/μ, one can write down an element v_T in $\bigotimes_i (\bigwedge^{\lambda_j' - \mu_j'} V)$; for example, corresponding to the displayed tableau, $v_T = v_4 \otimes v_2 \otimes (v_1 \wedge v_3)$. A key fact, generalizing the result of Exercise 6.15, is that the images of these elements under the map (vi) form a basis for $\mathbb{S}_{\lambda/\mu} V$.

The character of $\mathbb{S}_{\lambda/\mu} V$ is given by the Schur function $S_{\lambda/\mu}$: if g is an endomorphism of V with eigenvalues $x_1, \ldots, x_k$, then

(viii) $$\chi_{\mathbb{S}_{\lambda/\mu} V}(g) = S_{\lambda/\mu}(x_1, \ldots, x_k).$$

In terms of basic Schur functors,

(ix) $$\mathbb{S}_{\lambda/\mu} V \cong \sum N_{\mu\nu\lambda} \mathbb{S}_\nu V.$$

Exercise 6.20*. (a) Show that if $\lambda = (p, q)$, $\mathbb{S}_{(p,q)} V$ is the kernel of the contraction map

$$c_{p,q}\colon \mathrm{Sym}^p V \otimes \mathrm{Sym}^q V \to \mathrm{Sym}^{p+1} V \otimes \mathrm{Sym}^{q-1} V.$$

(b) If $\lambda = (p, q, r)$, show that $\mathbb{S}_{(p,q,r)} V$ is the intersection of the kernels of two contraction maps $c_{p,q} \otimes 1_r$ and $1_p \otimes c_{p,r}$, where 1_i denotes the identity map on $\mathrm{Sym}^i V$.

In general, for $\lambda = (\lambda_1, \ldots, \lambda_k)$, $\mathbb{S}_\lambda V \subset \mathrm{Sym}^{\lambda_1} V \otimes \cdots \otimes \mathrm{Sym}^{\lambda_k} V$ is the intersection of the kernels of the $k - 1$ maps

$$\psi_i = 1_{\lambda_1} \otimes \cdots \otimes 1_{\lambda_{i-1}} \otimes c_{\lambda_i, \lambda_{i+1}} \otimes 1_{\lambda_{i+2}} \otimes \cdots \otimes 1_{\lambda_k}, \quad 1 \le i \le k - 1.$$

(c) For $\lambda = (p, 1, \ldots., 1)$, show that $\mathbb{S}_\lambda V$ is the kernel of the contraction map:

$$\mathbb{S}_{(p,1,\ldots,1)} V = \mathrm{Ker}(\mathrm{Sym}^p V \otimes \textstyle\bigwedge^{d-p} V \to \mathrm{Sym}^{p+1} V \otimes \textstyle\bigwedge^{d-p-1} V).$$

In general, for any choice of a between 1 and $k - 1$, the intersection of the kernels of all ψ_i except ψ_a is $\mathbb{S}_\sigma V \otimes \mathbb{S}_\tau V$, where $\sigma = (\lambda_1, \ldots, \lambda_a)$ and $\tau = (\lambda_{a+1}, \ldots, \lambda_k)$; so $\mathbb{S}_\lambda V$ is the kernel of a contraction map defined on $\mathbb{S}_\sigma V \otimes \mathbb{S}_\tau V$. For example, if a is $k - 1$, and we set $r = \lambda_k$, Pieri's formula writes $\mathbb{S}_\sigma V \otimes \mathrm{Sym}^r V$ as a direct sum of $\mathbb{S}_\lambda V$ and other factors $\mathbb{S}_\nu V$; the general assertion in (b) is equivalent to the claim that $\mathbb{S}_\lambda V$ is the only factor that is in the kernel of the contraction, ie.,

$$\mathbb{S}_\lambda V = \mathrm{Ker}(\mathbb{S}_{(\lambda_1, \ldots, \lambda_{k-1})} V \otimes \mathrm{Sym}^r V \to V^{\otimes(d-r+1)} \otimes \mathrm{Sym}^{r-1} V).$$

These results correspond to writing the representations $V_\lambda \subset U_\lambda$ of the symmetric group as the intersection of kernels of maps to $U_{\lambda_1, \ldots, \lambda_i + 1, \lambda_{i+1} - 1, \ldots, \lambda_k}$.

Exercise 6.21. The functorial nature of Weyl's construction has many consequences, which are not explored in this book. For example, if E_* is a complex of vector spaces, the tensor product $E_*^{\otimes d}$ is also a complex, and the symmetric group $\mathfrak{S}_d$ acts on it; when factors in E_p and E_q are transposed past each other, the usual sign $(-1)^{pq}$ is inserted. The image of the Young symmetrizer c_λ is a complex $\mathbb{S}_\lambda(E_*)$, sometimes called a *Schur complex*. Show that if E_* is the complex $E_{-1} = V \to E_0 = V$, with the boundary map the identity map, and $\lambda = (d)$, then $\mathbb{S}_\lambda(E_*)$ is the Koszul complex

$$0 \to \Lambda^d \to \Lambda^{d-1} \otimes S^1 \to \Lambda^{d-2} \otimes S^2 \to \cdots \to \Lambda^1 \otimes S^{d-1} \to S^d \to 0,$$

where $\Lambda^i = \Lambda^i V$, and $S^j = \mathrm{Sym}^j V$.

§6.2. The Proofs

We need first a small piece of the general story about semisimple algebras, which we work out by hand. For the moment G can be any finite group, although our application is for the symmetric group. If U is a right module over $A = \mathbb{C}G$, let

$$B = \mathrm{Hom}_G(U, U) = \{\varphi\colon U \to U\colon \varphi(v \cdot g) = \varphi(v) \cdot g, \forall v \in U, g \in G\}.$$

Note that B acts on U on the left, commuting with the right action of A; B is called the *commutator* algebra. If $U = \bigoplus U_i^{\oplus n_i}$ is an irreducible decomposition with U_i nonisomorphic irreducible right A-modules, then by Schur's Lemma 1.7

$$B = \bigoplus_i \mathrm{Hom}_G(U_i^{\oplus n_i}, U_i^{\oplus n_i}) = \bigoplus_i M_{n_i}(\mathbb{C}),$$

where $M_{n_i}(\mathbb{C})$ is the ring of $n_i \times n_i$ complex matrices.

If W is any left A-module, the tensor product

$$U \otimes_A W = U \otimes_{\mathbb{C}} W/\text{subspace generated by } \{va \otimes w - v \otimes aw\}$$

is a left B-module by acting on the first factor: $b \cdot (v \otimes w) = (b \cdot v) \otimes w$.

Lemma 6.22. *Let U be a finite-dimensional right A-module.*

(i) *For any $c \in A$, the canonical map $U \otimes_A Ac \to Uc$ is an isomorphism of left B-modules.*

(ii) *If $W = Ac$ is an irreducible left A-module, then $U \otimes_A W = Uc$ is an irreducible left B-module.*

(iii) *If $W_i = Ac_i$ are the distinct irreducible left A-modules, with m_i the dimension of W_i, then*

$$U \cong \bigoplus_i (U \otimes_A W_i)^{\oplus m_i} \cong \bigoplus_i (Uc_i)^{\oplus m_i}$$

is the decomposition of U into irreducible left B-modules.

Proof. Note first that Ac is a direct summand of A as a left A-module; this is a consequence of the semisimplicity of all representations of G (Proposition 1.5). To prove (i), consider the commutative diagram

$$\begin{tikzcd} U \otimes_A A \arrow[r, "\cdot c"] \arrow[d] & U \otimes_A Ac \arrow[r, hook] \arrow[d] & U \otimes_A A \arrow[d] \\ U \arrow[r, "\cdot c"] & U \cdot c \arrow[r, hook] & U \end{tikzcd},$$

where the vertical maps are the maps $v \otimes a \mapsto v \cdot a$; since the left horizontal maps are surjective, the right ones injective, and the outside vertical maps are isomorphisms, the middle vertical map must be an isomorphism.

For (ii), consider first the case where U is an irreducible A-module, so $B = \mathbb{C}$. It suffices to show that $\dim U \otimes_A W \leq 1$. For this we use Proposition 3.29 to identify A with a direct sum $\bigoplus_{i=1}^r M_{m_i}\mathbb{C}$ of r matrix algebras. We can identify W with a minimal left ideal of A. Any minimal ideal in the sum of matrix algebras is isomorphic to one which consists of r-tuples of matrices which are zero except in one factor, and in this factor are all zero except for one column. Similarly, U can be identified with the right ideal of r-tuples which are zero except in one factor, and in that factor all are zero except in one row. Then $U \otimes_A W$ will be zero unless the factor is the same for U and W, in which case $U \otimes_A W$ can be identified with the matrices which are zero except in one row and column of that factor. This completes the proof when U is irreducible. For the general case of (ii), decompose $U = \bigoplus_i U_i^{\oplus n_i}$ into a sum of irreducible right A-modules, so $U \otimes_A W = \bigoplus_i (U_i \otimes_A W)^{\oplus n_i} = \mathbb{C}^{\oplus n_k}$ for some k, which is visibly irreducible over $B = \bigoplus M_{n_j}(\mathbb{C})$.

Part (iii) follows, since the isomorphism $A \cong \bigoplus W_i^{\oplus m_i}$ determines an isomorphism

$$U \cong U \otimes_A A \cong U \otimes_A \left(\bigoplus_i W_i^{\oplus m_i}\right) \cong \bigoplus_i (U \otimes_A W_i)^{\oplus m_i}. \qquad \square$$

To prove Theorem 6.3, we will apply Lemma 6.22 to the right $\mathbb{C}\mathfrak{S}_d$-module $U = V^{\otimes d}$. That lemma shows how to decompose U as a B-module, where B

is the algebra of all endomorphisms of U that commute with all permutations of the factors. The endomorphisms of U induced by endomorphisms of V are certainly in this algebra B. Although B is generally much larger than $\operatorname{End}(V)$, we have

Lemma 6.23. *The algebra B is spanned as a linear subspace of* $\operatorname{End}(V^{\otimes d})$ *by* $\operatorname{End}(V)$. *A subspace of $V^{\otimes d}$ is a sub-B-module if and only if it is invariant by* $\operatorname{GL}(V)$.

PROOF. Note that if W is any finite-dimensional vector space, then $\operatorname{Sym}^d W$ is the subspace of $W^{\otimes d}$ spanned by all $w^d = d!w \otimes \cdots \otimes w$ as w runs through W. Applying this to $W = \operatorname{End}(V) = V^* \otimes V$ proves the first statement, since $\operatorname{End}(V^{\otimes d}) = (V^*)^{\otimes d} \otimes V^{\otimes d} = W^{\otimes d}$, with compatible actions of $\mathfrak{S}_d$. The second follows from the fact that $\operatorname{GL}(V)$ is dense in $\operatorname{End}(V)$. $\square$

We turn now to the proof of Theorem 6.3. Note that $\mathbb{S}_\lambda V$ is Uc_λ, so parts (2) and (4) follow from Lemmas 6.22 and 6.23. We use the same methods to give a rather indirect but short proof of part (3); for a direct approach see Exercise 6.28. From Lemma 6.22 we have an isomorphism of $\operatorname{GL}(V)$-modules:

$$\mathbb{S}_\lambda V \cong V^{\otimes d} \otimes_A V_\lambda \tag{6.24}$$

with $V_\lambda = A \cdot c_\lambda$. Similarly for $U_\lambda = A \cdot a_\lambda$, and since the image of right multiplication by a_λ on $V^{\otimes d}$ is the tensor product of symmetric powers, we have

$$\operatorname{Sym}^{\lambda_1} V \otimes \operatorname{Sym}^{\lambda_2} V \otimes \cdots \otimes \operatorname{Sym}^{\lambda_k} V \cong V^{\otimes d} \otimes_A U_\lambda. \tag{6.25}$$

But we have an isomorphism $U_\lambda \cong \bigoplus_\mu K_{\mu\lambda} V_\mu$ of A-modules by Young's rule (4.39), so we deduce an isomorphism of $\operatorname{GL}(V)$-modules

$$\operatorname{Sym}^{\lambda_1} V \otimes \operatorname{Sym}^{\lambda_2} V \otimes \cdots \otimes \operatorname{Sym}^{\lambda_k} V \cong \bigoplus_\mu K_{\mu\lambda} \mathbb{S}_\mu V. \tag{6.26}$$

By what we saw before the statement of the theorem, the trace of g on the left-hand side of (6.26) is the product $H_\lambda(x_1, \ldots, x_k)$ of the complete symmetric polynomials $H_{\lambda_i}(x_1, \ldots, x_k)$. Let $\mathbb{S}_\lambda(g)$ denote the endomorphism of $\mathbb{S}_\lambda V$ determined by an endomorphism g of V. We therefore have

$$H_\lambda(x_1, \ldots, x_k) = \Sigma_\mu K_{\mu\lambda} \operatorname{Trace}(\mathbb{S}_\mu(g)).$$

But these are precisely the relations between the functions H_λ and the Schur polynomials S_μ [see formula (A.9)], and these relations are invertible, since the matrix $(K_{\mu\lambda})$ of coefficients is triangular with 1's on the diagonal. It follows that $\operatorname{Trace}(\mathbb{S}_\lambda(g)) = S_\lambda(x_1, \ldots, x_k)$, which proves part (3).

Note that if $\lambda = (\lambda_1, \ldots, \lambda_d)$ with $d > k$ and $\lambda_{k+1} \neq 0$, this same argument shows that the trace is $S_\lambda(x_1, \ldots, x_k, 0, \ldots, 0)$, which is zero, for example by (A.6). For g the identity, this shows that $\mathbb{S}_\lambda V = 0$ in this case. From part (3) we also get

$$\dim \mathbb{S}_\lambda V = S_\lambda(1, \ldots, 1), \tag{6.27}$$

and computing $S_\lambda(1, \ldots, 1)$ via Exercise A.30(ii) yields part (1). □

Exercise 6.28. If you have given an independent proof of Problem 6.15, part (3) of Theorem 6.3 can be seen directly. The basis elements v_T for $\mathbb{S}_\lambda V$ specified in Problem 6.15 are eigenvectors for a diagonal matrix with entries $x_1, \ldots, x_k$, with eigenvalue $X^a = x_1^{a_1} \cdot \ldots \cdot x_k^{a_k}$, where the tableau T has a_1 1's, a_2 2's, ..., a_k k's. The trace is therefore $\sum K_{\lambda a} X^a$, where $K_{\lambda a}$ is the number of ways to number the boxes of the Young diagram of λ with a_1 1's, a_2 2's, ..., a_k k's. This is just the expression for S_λ obtained in Exercise A.31(a).

We conclude this lecture with a few of the standard elaborations of these ideas, in exercise form; they are not needed in these lectures.

Exercise 6.29*. Show that, in the context of Lemma 6.22, if U is a faithful A-module, then A is the commutator of its commutator B:

$$A = \{\psi\colon U \to U\colon \psi(bv) = b\psi(v), \forall v \in U, b \in B\}.$$

If U is not faithful, the canonical map from A to its bicommutator is surjective. Conclude that, in Theorem 6.3, the algebra of endomorphisms of $V^{\otimes d}$ that commute with $\mathrm{GL}(V)$ is spanned by the permutations in $\mathfrak{S}_d$.

Exercise 6.30. Show that, in Lemma 6.22, there is a natural one-to-one correspondence between the irreducible right A-modules U_i that occur in U and the irreducible left B-modules V_i. Show that there is a canonical decomposition

$$U = \bigoplus_i (V_i \otimes_{\mathbb{C}} U_i)$$

as a left B-module and as a right A-module. This shows again that the number of times V_i occurs in U is the dimension of U_i, and dually that the number of times U_i occurs is the dimension of V_i. Deduce the canonical decomposition

$$V^{\otimes d} = \bigoplus \mathbb{S}_\lambda V \otimes_{\mathbb{C}} V_\lambda,$$

the sum over partitions λ of d into at most $k = \dim V$ parts; this decomposition is compatible with the actions of $\mathrm{GL}(V)$ and $\mathfrak{S}_d$. In particular, the number of times V_λ occurs in the representation $V^{\otimes d}$ of $\mathfrak{S}_d$ is the dimension of $\mathbb{S}_\lambda V$.

Exercise 6.31. Let e be an idempotent in the group algebra $A = \mathbb{C}G$, and let $U = eA$ be the corresponding right A-module. Let $E = eAe$, a subalgebra of A. The algebra structure in A makes eA a left E-module. Show that this defines an isomorphism of $\mathbb{C}$-algebras

$$E = eAe \cong \mathrm{Hom}_A(eA, eA) = \mathrm{Hom}_G(U, U) = B.$$

Exercise 6.32. If H is a subgroup of G, and $e \in \mathbb{C}H$ is an idempotent, corresponding to a representation $W = \mathbb{C}H \cdot e$ of H, show that $\mathbb{C}G \cdot e$ is the induced representation $\mathrm{Ind}_H^G(W)$. For example, if $\vartheta \colon H \to \mathbb{C}^*$ is a one-dimensional representation, then

$$\mathrm{Ind}_H^G(\vartheta) = \mathbb{C}G \cdot e_\vartheta, \quad \text{where } e_\vartheta = \frac{1}{|G|} \sum_{g \in G} \overline{\vartheta(g)} e_g.$$

PART II
LIE GROUPS AND LIE ALGEBRAS

From a naive point of view, Lie groups seem to stand at the opposite end of the spectrum of groups from finite ones.[1] On the one hand, as abstract groups they seem enormously complicated: for example, being of uncountable order, there is no question of giving generators and relations. On the other hand, they do come with the additional data of a topology and a manifold structure; this makes it possible—and, given the apparent hopelessness of approaching them purely as algebraic objects, necessary—to use geometric concepts to study them.

Lie groups thus represent a confluence of algebra, topology, and geometry, which perhaps accounts in part for their ubiquity in modern mathematics. It also makes the subject a potentially intimidating one: to have to understand, both individually and collectively, all these aspects of a single object may be somewhat daunting.

Happily, just because the algebra and the geometry/topology of a Lie group are so closely entwined, there is an object we can use to approach the study of Lie groups that extracts much of the structure of a Lie group (primarily its algebraic structure) while seemingly getting rid of the topological complexity. This is, of course, the *Lie algebra*. The Lie algebra is, at least according to its definition, a purely algebraic object, consisting simply of a vector space with bilinear operation; and so it might appear that in associating to a Lie group its Lie algebra we are necessarily giving up a lot of information about the group. This is, in fact, not the case: as we shall see in many cases (and perhaps all of the most important ones), encoded in the algebraic structure of a Lie algebra is almost all of the geometry of the group. In particular, we will

[1] In spite of this there are deep, if only partially understood, relations between finite and Lie groups, extending even to their simple group classifications.

see by the end of Lecture 8 that there is a very close relationship between representations of the Lie group we start with and representations of the Lie algebra we associate to it; and by the end of the book we will make that correspondence exact.

We said that passing from the Lie group to its Lie algebra represents a simplification because it eliminates whatever nontrivial topological structure the group may have had; it "flattens out," or "linearizes," the group. This, in turn, allows for a further simplification: since a Lie algebra is just a vector space with bilinear operation, it makes perfect sense, if we are asked to study a real Lie algebra (or one over any subfield of $\mathbb{C}$) to tensor with the complex numbers. Thus, we may investigate first the structure and representations of *complex Lie algebras*, and then go back to apply this knowledge to the study of real ones. In fact, this turns out to be a feasible approach, in every respect: the structure of complex Lie algebras tends to be substantially simpler than that of real Lie algebras; and knowing the representations of the complex Lie algebra will solve the problem of classifying the representations of the real one.

There is one further reduction to be made: some very elementary Lie algebra theory allows us to narrow our focus further to the study of *semisimple Lie algebras*. This is a subset of Lie algebras analogous to simple groups in that they are in some sense atomic objects, but better behaved in a number of ways: a semisimple Lie algebra is a direct sum of simple ones; there are easy criteria for the semisimplicity of a given Lie algebra; and, most of all, their representation theory can be approached in a completely uniform manner. Moreover, as in the case of finite groups, there is a complete classification theorem for simple Lie algebras.

We may thus describe our approach to the representation theory of Lie groups by the sequence of objects

Lie group

$\rightsquigarrow$ Lie algebra

$\rightsquigarrow$ complex Lie algebra

$\rightsquigarrow$ semisimple complex Lie algebra.

We describe this progression in Lectures 7–9. In Lectures 7 and 8 we introduce the definitions of and some basic facts about Lie groups and Lie algebras. Lecture 8 ends with a description of the exponential map, which allows us to establish the close connection between the first two objects above. We then do, in Lecture 9, the very elementary classification theory of Lie algebras that motivates our focus on semisimple complex Lie algebras, and at least state the classification theorem for these. This establishes the fact that the second, third, and fourth objects above have essentially the same irreducible representations. (This lecture may also serve to give a brief taste of some general theory, which is mostly postponed to later lectures or appendices.) In Lecture 10 we discuss examples of Lie algebras in low dimensions.

From that point on we will proceed to devote ourselves almost exclusively to the study of semisimple complex Lie algebras and their representations. We do this, we have to say, in an extremely inefficient manner: we start with a couple of very special cases, which occupy us for three lectures (11–13); enunciate the general paradigm in Lecture 14; carry this out for the classical Lie algebras in Lectures 15–20; and (finally) finish off the general theory in Lectures 21–26. Thus, it will not be until the end that we go back and use the knowledge we have gained to say something about the original problem. In view of this long interlude, it is perhaps a good idea to enunciate one more time our basic

Point of View: The primary objects of interest are Lie groups and their representations; these are what actually occur in real life and these are what we want to understand. The notion of a complex Lie algebra is introduced primarily as a tool in this study; it is an essential tool[2] and we should consider ourselves incredibly lucky to have such a wonderfully effective one; but in the end it is for us a means to an end.

The special cases worked out in Lectures 11–13 are the Lie algebras of SL_2 and SL_3. Remarkably, most of the structure shared by all semisimple Lie algebras can be seen in these examples. We should probably point out that much of what we do by hand in these cases could be deduced from the Weyl construction we saw in Lecture 6 (as we will do generally in Lecture 15), but we mainly ignore this, in order to work from a "Lie algebra" point of view and motivate the general story.

[2] Perhaps not logically so; cf. Adams' book [Ad].

LECTURE 7

Lie Groups

In this lecture we introduce the definitions and basic examples of Lie groups and Lie algebras. We assume here familiarity with the definition of differentiable manifolds and maps between them, but no more; in particular, we do not mention vector fields, differential forms, Riemannian metrics, or any other tensors. Section 7.3, which discusses maps of Lie groups that are covering space maps of the underlying manifolds, may be skimmed and referred back to as needed, though working through it will help promote familiarity with basic examples of Lie groups.

§7.1. Lie Groups: Definitions

You probably already know what a Lie group is; it is just a set endowed simultaneously with the compatible structures of a group and a $\mathscr{C}^\infty$ manifold. "Compatible" here means that the multiplication and inverse operations in the group structure

$$\times : G \times G \to G$$

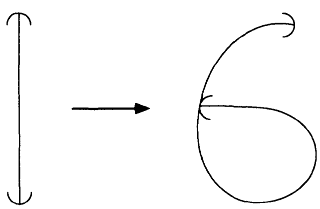
and

$$\iota : G \to G$$

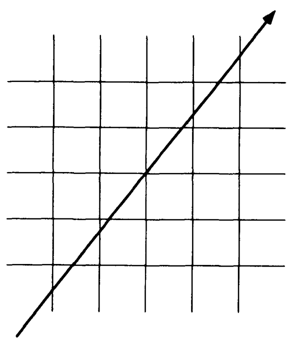
are actually differentiable maps (logically, this is equivalent to the single requirement that the map $G \times G \to G$ sending (x, y) to $x \cdot y^{-1}$ is $\mathscr{C}^\infty$).

A *map*, or *morphism*, between two Lie groups G and H is just a map $\rho: G \to H$ that is both differentiable and a group homomorphism. In general, qualifiers applied to Lie groups refer to one or another of the two structures, usually without much ambiguity; thus, *abelian* refers to the group structure, *n-dimensional* or *connected* refers to the manifold structure. Sometimes a condition on one structure turns out to be equivalent to a condition on the other; for example, we will see below that to say that a map of connected Lie groups $\varphi: G \to H$ is a surjective map of groups is equivalent to saying that the differential $d\varphi$ is surjective at every point.

One area where there is some potential confusion is in the definition of a Lie subgroup. This is essentially a difficulty inherited directly from manifold theory, where we have to make a distinction between a *closed submanifold* of a manifold M, by which we mean a subset $X \subset M$ that inherits a manifold structure from M (i.e., that may be given, locally in M, by setting a subset of the local coordinates equal to zero), and an *immersed submanifold*, by which we mean the image of a manifold X under a one-to-one map with injective differential everywhere—that is, a map that is an embedding *locally in X*. The distinction is necessary simply because the underlying topological space structure of an immersed submanifold may not agree with the topological structure induced by the inclusion of X in M. For example, the map from X to M could be the immersion of an open interval in $\mathbb{R}$ into the plane $\mathbb{R}^2$ as a figure "6":



Another standard example of this, which is also an example in the category of groups, would be to take M to be the two-dimensional real torus $\mathbb{R}^2/\mathbb{Z}^2 = S^1 \times S^1$, and X the image in M of a line $V \subset \mathbb{R}^2$ having irrational slope:



The upshot of this is that we define a *Lie subgroup* (or *closed Lie subgroup*, if we want to emphasize the point) of a Lie group G to be a subset that is simultaneously a subgroup and a *closed* submanifold; and we define an *immersed subgroup* to be the image of a Lie group H under an injective morphism to G. (That a one-to-one morphism of Lie groups has everywhere injective differential will follow from discussions later in this lecture.)

The definition of a *complex Lie group* is exactly analogous, the words "differentiable manifold" being replaced by "complex manifold" and all related notions revised accordingly. Similarly, to define an *algebraic group* one replaces "differentiable manifold" by "algebraic variety" and "differentiable map" by "regular morphism." As we will see, the category of complex Lie groups is in many ways markedly different from that of real Lie groups (for example, there are many fewer complex Lie groups than real ones). Of course, the study of algebraic groups in general is quite different from either of these since an algebraic group comes with a field of definition that may or may not be a subfield of $\mathbb{C}$ (it may, for that matter, have positive characteristic). In practice, though, while the two are not the same (we will see examples of this in Lecture 10, for example), the category of algebraic groups over $\mathbb{C}$ behaves very much like the category of complex Lie groups.

§7.2. Examples of Lie Groups

The basic example of a Lie group is of course the *general linear group* $\mathrm{GL}_n\mathbb{R}$ of invertible $n \times n$ real matrices; this is an open subset of the vector space of all $n \times n$ matrices, and gets its manifold structure accordingly (so that the entries of the matrix are coordinates on $\mathrm{GL}_n\mathbb{R}$). That the multiplication map $\mathrm{GL}_n\mathbb{R} \times \mathrm{GL}_n\mathbb{R} \to \mathrm{GL}_n\mathbb{R}$ is differentiable is clear; that the inverse map $\mathrm{GL}_n\mathbb{R} \to \mathrm{GL}_n\mathbb{R}$ is follows from Cramer's formula for the inverse. Occasionally $\mathrm{GL}_n\mathbb{R}$ will come to us as the group of automorphisms of an n-dimensional real vector space V; when we want to think of $\mathrm{GL}_n\mathbb{R}$ in this way (e.g., without choosing a basis for V and thereby identifying G with the group of matrices), we will write it as $\mathrm{GL}(V)$ or $\mathrm{Aut}(V)$. A *representation* of a Lie group G, of course, is a morphism from G to $\mathrm{GL}(V)$.

Most other Lie groups are defined initially as subgroups of GL_n (though they may appear in other contexts as subgroups of other general linear groups, which is, of course, the subject matter of these lectures). For the most part, such subgroups may be described either by equations on the entries of an $n \times n$ matrix, or as the subgroup of automorphisms of $V \cong \mathbb{R}^n$ preserving some structure on $\mathbb{R}^n$. For example, we have:

the *special linear group* $\mathrm{SL}_n\mathbb{R}$ of automorphisms of $\mathbb{R}^n$ preserving the volume element; equivalently, $n \times n$ matrices A with determinant 1.

the group B_n of *upper-triangular* matrices; equivalently, the subgroup of automorphisms of $\mathbb{R}^n$ preserving the flag[1]

[1] In general, a *flag* is a sequence of subspaces of a fixed vector space, each properly contained in the next; it is a *complete* flag if each has one dimension larger than the preceding, and *partial* otherwise.

$$0 = V_0 \subset V_1 \subset V_2 \subset \cdots V_{n-1} \subset V_n = \mathbb{R}^n,$$

where V_i is the span of the standard basis vectors $e_1, \ldots, e_i$. Note that choosing a different basis and correspondingly a different flag yields a different subgroup of $\mathrm{GL}_n\mathbb{R}$, but one isomorphic to (indeed, conjugate to) B_n. Somewhat more generally, for any sequence of positive integers $a_1, \ldots, a_k$ with sum n we can look at the group of block-upper-triangular matrices; this is the subgroup of automorphisms of $\mathbb{R}^n$ preserving a partial flag

$$0 = V_0 \subset V_1 \subset V_2 \subset \cdots \subset V_{k-1} \subset V_k = \mathbb{R}^n,$$

where the dimension of V_i is $a_1 + \cdots + a_i$. If the subspace V_i is spanned by the first $a_1 + \cdots + a_i$ basis vectors, the group will be the set of matrices of the form

$$\left(\begin{array}{c|c|c|c} * & * & * & * \\ \hline 0 & * & * & * \\ \hline 0 & 0 & * & * \\ \hline 0 & 0 & 0 & * \end{array}\right) \begin{array}{l} \}a_1 \\ \}a_2 \\ \\ \}a_k \end{array}$$

The group N_n of *upper-triangular unipotent* matrices (that is, upper triangular with 1's on the diagonal); equivalently, the subgroup of automorphisms of $\mathbb{R}^n$ preserving the complete flag $\{V_i\}$ where V_i is the span of the standard basis vectors $e_1, \ldots, e_i$, and acting as the identity on the successive quotients V_{i+1}/V_i. As before, we can, for any sequence of positive integers $a_1, \ldots, a_k$ with sum n, look at the group of block-upper-triangular unipotent matrices; this is the subgroup of automorphisms of $\mathbb{R}^n$ preserving a partial flag and acting as the identity on successive quotients, i.e., matrices of the form

$$\left(\begin{array}{c|c|c|c} I & * & * & * \\ \hline 0 & I & * & * \\ \hline 0 & 0 & I & * \\ \hline 0 & 0 & 0 & I \end{array}\right) \begin{array}{l} \}a_1 \\ \}a_2 \\ \\ \}a_k \end{array}$$

Next, there are the subgroups of $\mathrm{GL}_n\mathbb{R}$ defined as the group of transformations of $V = \mathbb{R}^n$ of determinant 1 preserving some bilinear form $Q\colon V \times V \to \mathbb{R}$. If the bilinear form Q is symmetric and positive definite, the group we get is called the (*special*) *orthogonal group* $\mathrm{SO}_n\mathbb{R}$ (sometimes written $\mathrm{SO}(n)$; see p. 100). If Q is symmetric and nondegenerate but not definite—e.g., if it has k positive eigenvalues and l negative—the group is denoted $\mathrm{SO}_{k,l}\mathbb{R}$ or $\mathrm{SO}(k, l)$; note that $\mathrm{SO}(k, l) \cong \mathrm{SO}(l, k)$. If Q is skew-symmetric and nondegenerate, the group is called the *symplectic group* and denoted $\mathrm{Sp}_n\mathbb{R}$; note that in this case n must be even.

The equations that define the subgroup of $\mathrm{GL}_n\mathbb{R}$ preserving a bilinear form Q are easy to write down. If we represent Q by a matrix M—that is, we write

$$Q(v, w) = {}^t v \cdot M \cdot w$$

for all $v, w \in \mathbb{R}^n$—then the condition

$$Q(Av, Aw) = Q(v, w)$$

translates into the condition that

$${}^t v \cdot {}^t A \cdot M \cdot A \cdot w = {}^t v \cdot M \cdot w$$

for all v and w; this is equivalent to saying that

$${}^t A \cdot M \cdot A = M.$$

Thus, for example, if Q is the symmetric form $Q(v, w) = {}^t v \cdot w$ given by the identity matrix $M = I_n$, the group $\mathrm{SO}_n\mathbb{R}$ is just the group of $n \times n$ real matrices A of determinant 1 such that ${}^t A = A^{-1}$.

Exercise 7.1*. Show that in the case of $\mathrm{Sp}_{2n}\mathbb{R}$ the requirement that the transformations have determinant 1 is redundant; whereas in the case of $\mathrm{SO}_n\mathbb{R}$ if we do not require the transformations to have determinant 1 the group we get (denoted $\mathrm{O}_n\mathbb{R}$, or sometimes $\mathrm{O}(n)$) is disconnected.

Exercise 7.2*. Show that $\mathrm{SO}(k, l)$ has two connected components if k and l are both positive. The connected component containing the identity is often denoted $\mathrm{SO}^+(k, l)$. (Composing with a projection onto $\mathbb{R}^k$ or $\mathbb{R}^l$, we may associate to an automorphism $A \in \mathrm{SO}(k, l)$ automorphisms of $\mathbb{R}^k$ and $\mathbb{R}^l$; $\mathrm{SO}^+(k, l)$ will consist of those $A \in \mathrm{SO}(k, l)$ whose associated automorphisms preserve the orientations of $\mathbb{R}^k$ and $\mathbb{R}^l$.)

Note that if the form Q is degenerate, a transformation preserving Q will carry its kernel

$$\mathrm{Ker}(Q) = \{v \in V \colon Q(v, w) = 0 \ \forall w \in V\}$$

into itself; so that the group we get is simply the group of matrices preserving the subspace $\mathrm{Ker}(Q)$ and preserving the induced nondegenerate form $\tilde{Q}$ on the quotient $V/\mathrm{Ker}(Q)$. Likewise, if Q is a general bilinear form, that is, neither symmetric nor skew-symmetric, a linear transformation preserving Q will preserve the symmetric and skew-symmetric parts of Q individually, so we just get an intersection of the subgroups encountered already. At any rate, we usually limit our attention to nondegenerate forms that are either symmetric or skew-symmetric.

Of course, the group $\mathrm{GL}_n\mathbb{C}$ of complex linear automorphisms of a complex vector space $V = \mathbb{C}^n$ can be viewed as subgroup of the general linear group $\mathrm{GL}_{2n}\mathbb{R}$; it is, thus, a real Lie group as well, as is the subgroup $\mathrm{SL}_n\mathbb{C}$ of $n \times n$ complex matrices of determinant 1. Similarly, the subgroups $\mathrm{SO}_n\mathbb{C} \subset \mathrm{SL}_n\mathbb{C}$ and $\mathrm{Sp}_{2n}\mathbb{C} \subset \mathrm{SL}_{2n}\mathbb{C}$ of transformations of a complex vector space preserving a symmetric and skew-symmetric nondegenerate *bilinear* form, respectively, are real as well as complex Lie subgroups. Note that since all nondegenerate bilinear symmetric forms on a complex vector space are isomorphic (in particular, there is no such thing as a signature), there is only one complex orthogonal subgroup $\mathrm{SO}_n\mathbb{C} \subset \mathrm{SL}_n\mathbb{C}$ up to conjugation; there are no analogs of the groups $\mathrm{SO}_{k,l}\mathbb{R}$.

Another example we can come up with here is the *unitary group* U_n or $\mathrm{U}(n)$, defined to be the group of complex linear automorphisms of an n-dimensional complex vector space V preserving a positive definite *Hermitian* inner product H on V. (A *Hermitian* form H is required to be conjugate linear in the first[2] factor, and linear in the second: $H(\lambda v, \mu w) = \bar{\lambda} H(v, w)\mu$, and $H(w, v) = \overline{H(v, w)}$; it is *positive definite* if $H(v, v) > 0$ for $v \neq 0$.)

Just as in the case of the subgroups SO and Sp, it is easy to write down the equations for $\mathrm{U}(n)$: for some $n \times n$ matrix M we can write the form H as

$$H(v, w) = {}^t\bar{v} \cdot M \cdot w, \quad \forall v, w \in \mathbb{C}^n$$

(note that for H to be conjugate symmetric, M must be conjugate symmetric, i.e., ${}^tM = \overline{M}$); then the group $\mathrm{U}(n)$ is just the group of $n \times n$ complex matrices A satisfying

$${}^t\bar{A} \cdot M \cdot A = M.$$

In particular, if H is the "standard" Hermitian inner product $H(v, w) = {}^t\bar{v} \cdot w$ given by the identity matrix, $\mathrm{U}(n)$ will be the group of $n \times n$ complex matrices A such that ${}^t\bar{A} = A^{-1}$.

Exercise 7.3. Show that if H is a Hermitian form on a complex vector space V, then the real part $R = \mathrm{Re}(H)$ of H is a symmetric form on the underlying real space, and the imaginary part $C = \mathrm{Im}(H)$ is a skew-symmetric real form; these are related by $C(v, w) = R(iv, w)$. Both R and C are invariant by multiplication by i: $R(iv, iw) = R(v, w)$. Show conversely that any such real symmetric R is the real part of a unique Hermitian H. Show that if H is standard, so is R, and C corresponds to the matrix $J = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}$. Deduce that

$$\mathrm{U}(n) = \mathrm{O}(2n) \cap \mathrm{Sp}_{2n}\mathbb{R}.$$

Note that the determinant of a unitary matrix can be any complex number of modulus 1; the *special unitary group*, $\mathrm{SU}(n)$, is the subgroup of $\mathrm{U}(n)$ of automorphisms with determinant 1. The subgroup of $\mathrm{GL}_n\mathbb{C}$ preserving an indefinite Hermitian inner product with k positive eigenvalues and l negative ones is denoted $\mathrm{U}_{k,l}$ or $\mathrm{U}(k, l)$; the subgroup of those of determinant 1 is denoted $\mathrm{SU}_{k,l}$ or $\mathrm{SU}(k, l)$.

In a similar vein, the group $\mathrm{GL}_n\mathbb{H}$ of quaternionic linear automorphisms of an n-dimensional vector space V over the ring $\mathbb{H}$ of quaternions is a real

[2] This choice of which factor is linear and which conjugate linear is less common than the other. It makes little difference in what follows, but it does have the small advantage of being compatible with the natural choice for quaternions.

Lie subgroup of the group $GL_{4n}\mathbb{R}$, as are the further subgroups of $\mathbb{H}$-linear transformations of V preserving a bilinear form. Since $\mathbb{H}$ is not commutative, care must be taken with the conventions here, and it may be worth a little digression to go through this now. We take the vector spaces V to be right $\mathbb{H}$-modules; $\mathbb{H}^n$ is the space of column vectors with *right* multiplication by scalars. In this way the $n \times n$ matrices with entries in $\mathbb{H}$ act in the usual way on $\mathbb{H}^n$ on the left. Scalar multiplication on the left (only) is $\mathbb{H}$-linear.

View $\mathbb{H} = \mathbb{C} \oplus j\mathbb{C} \cong \mathbb{C}^2$. Then left multiplication by elements of $\mathbb{H}$ give $\mathbb{C}$-linear endomorphisms of $\mathbb{C}^2$, which determines a mapping $\mathbb{H} \to M_2\mathbb{C}$ to the 2×2 complex matrices. In particular, $\mathbb{H}^* = GL_1\mathbb{H} \hookrightarrow GL_2\mathbb{C}$. Similarly $\mathbb{H}^n = \mathbb{C}^n \oplus j\mathbb{C}^n = \mathbb{C}^{2n}$, so we have an embedding $GL_n\mathbb{H} \hookrightarrow GL_{2n}\mathbb{C}$. Note that a $\mathbb{C}$-linear mapping $\varphi\colon \mathbb{H}^n \to \mathbb{H}^n$ is $\mathbb{H}$-linear exactly when it commutes with j: $\varphi(vj) = \varphi(v)j$. If $v = v_1 + jv_2$, then $v \cdot j = -\bar{v}_2 + j\bar{v}_1$, so multiplication by j takes $\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ to $\begin{pmatrix} 0 & -I \\ I & 0 \end{pmatrix}\begin{pmatrix} \bar{v}_1 \\ \bar{v}_2 \end{pmatrix}$. It follows that if J is the matrix of the preceding exercise, then

$$GL_n\mathbb{H} = \{A \in GL_{2n}\mathbb{C}\colon AJ = J\bar{A}\}.$$

Those matrices with real determinant 1 form a subgroup $SL_n\mathbb{H}$.

A *Hermitian form* (or "symplectic scalar product") on a quaternionic vector space V is an $\mathbb{R}$-bilinear form $K\colon V \times V \to \mathbb{H}$ that is conjugate $\mathbb{H}$-linear in the first factor and $\mathbb{H}$-linear in the second: $K(v\lambda, w\mu) = \bar{\lambda}K(v, w)\mu$, and satisfies $K(w, v) = \overline{K(v, w)}$. It is *positive definite* if $K(v, v) > 0$ for $v \neq 0$. (The conjugate $\bar{\lambda}$ of a quaternion $\lambda = a + bi + cj + dk$ is defined to be $a - bi - cj - dk$.) The *standard* Hermitian form on $\mathbb{H}^n$ is $\Sigma \bar{v}_i w_i$. The group of automorphisms of an n-dimensional quaternionic space preserving such a form is called the *compact symplectic* group and denoted $Sp(n)$ or $U_{\mathbb{H}}(n)$. The *standard* Hermitian form on $\mathbb{H}^n$ is $\Sigma \bar{v}_i w_i$.

Exercise 7.4. Regarding V as a complex vector space, show that every quaternionic Hermitian form K has the form

$$K(v, w) = H(v, w) + jQ(v, w),$$

where H is a complex Hermitian form and Q is a skew-symmetric complex linear form on V, with H and Q related by $Q(v, w) = H(vj, w)$, and H satisfying the condition $H(vj, wj) = \overline{H(v, w)}$. Conversely, any such Hermitian H is the complex part of a unique K. If K is standard, so is H, and Q is given by the same matrix as in Exercise 7.3. Deduce that

$$Sp(n) = U(2n) \cap Sp_{2n}\mathbb{C}.$$

This shows that the two notions of "symplectic" are compatible.

More generally, if K is not positive definite, but has signature (p, q), say the standard $\sum_{i=1}^{p} \bar{v}_i w_i - \sum_{i=p+1}^{p+q} \bar{v}_i w_i$, the automorphisms preserving it form a group $U_{p,q}\mathbb{H}$. Or if the form is a skew Hermitian form (satisfying the same linearity conditions, but with $K(w, v) = -\overline{K(v, w)}$), the group is denoted $U_n^*\mathbb{H}$.

Exercise 7.5. Identify, among all the real Lie groups described above, which ones are compact.

Complex Lie Groups

So far, all of our examples have been examples of real Lie groups. As for *complex* Lie groups, these are fewer in number. The general linear group $GL_n\mathbb{C}$ is one, and again, all the elementary examples come to us as subgroups of the general linear group $GL_n\mathbb{C}$. There is, for example, the subgroup $SO_n\mathbb{C}$ of automorphisms of an n-dimensional complex vector space V having determinant 1 and preserving a nondegenerate symmetric bilinear form Q (note that Q no longer has a signature); and likewise the subgroup $Sp_n\mathbb{C}$ of transformations of determinant 1 preserving a skew-symmetric bilinear form.

Exercise 7.6. Show that the subgroup $SU(n) \subset SL_n\mathbb{C}$ is *not* a complex Lie subgroup. (It is not enough to observe that the defining equations given above are not holomorphic.)

Exercise 7.7. Show that none of the complex Lie groups described above is compact.

We should remark here that both of these exercises are immediate consequences of the general fact that *any compact complex Lie group is abelian*; we will prove this in the next lecture. A *representation* of a complex Lie group G is a map of complex Lie groups from G to $GL(V) = GL_n\mathbb{C}$ for an n-dimensional complex vector space V; note that such a map is required to be complex analytic.

Remarks on Notation

A common convention is to use a notation without subscripts or mention of ground field to denote the *real groups*:

$$O(n), \quad SO(n), \quad SO(p, q), \quad U(n), \quad SU(n), \quad SU(p, q), \quad Sp(n)$$

and to use subscripts for the algebraic groups GL_n, SL_n, SO_n, and Sp_n. This, of course, introduces some anomalies: for example, $SO_n\mathbb{R}$ is $SO(n)$, but $Sp_n\mathbb{R}$ is not $Sp(n)$; but some violation of symmetry seems inevitable in any notation. The notations $GL(n, \mathbb{R})$ or $GL(n, \mathbb{C})$ are often used in place of our $GL_n\mathbb{R}$ or $GL_n\mathbb{C}$, and similarly for SL, SO, and Sp.

Also, where we have written Sp_{2n}, some write Sp_n. In practice, it seems that those most interested in algebraic groups or Lie algebras use the former notation, and those interested in compact groups the latter. Other common notations are $U^*(2n)$ in place of our $GL_n\mathbb{H}$, $Sp(p, q)$ for our $U_{p,q}\mathbb{H}$, and $O^*(2n)$ for our $U_n^*\mathbb{H}$.

Exercise 7.8. Find the dimensions of the various real Lie groups $\mathrm{GL}_n\mathbb{R}$, $\mathrm{SL}_n\mathbb{R}$, B_n, N_n, $\mathrm{SO}_n\mathbb{R}$, $\mathrm{SO}_{k,l}\mathbb{R}$, $\mathrm{Sp}_{2n}\mathbb{R}$, $\mathrm{U}(n)$, $\mathrm{SU}(n)$, $\mathrm{GL}_n\mathbb{C}$, $\mathrm{SL}_n\mathbb{C}$, $\mathrm{GL}_n\mathbb{H}$, and $\mathrm{Sp}(n)$ introduced above.

§7.3. Two Constructions

There are two constructions, in some sense inverse to one another, that arise frequently in dealing with Lie groups (and that also provide us with further examples of Lie groups). They are expressed in the following two statements.

Proposition 7.9. *Let G be a Lie group, H a connected manifold, and $\varphi\colon H \to G$ a covering space map.*[3] *Let e' be an element lying over the identity e of G. Then there is a unique Lie group structure on H such that e' is the identity and φ is a map of Lie groups; and the kernel of φ is in the center of H.*

Proposition 7.10. *Let H be a Lie group, and $\Gamma \subset Z(H)$ a discrete subgroup of its center. Then there is a unique Lie group structure on the quotient group $G = H/\Gamma$ such that the quotient map $H \to G$ is a Lie group map.*

The proof of the second proposition is straightforward. To prove the first, one shows that the multiplication on G lifts uniquely to a map $H \times H \to H$ which takes (e', e') to e', and verifies that this product satisfies the group axioms. In fact, it suffices to do this when H is the universal covering of G, for one can then apply the second proposition to intermediate coverings. □

Exercise 7.11*. (a) Show that any discrete normal subgroup of a connected Lie group G is in the center $Z(G)$.

(b) If $Z(G)$ is discrete, show that $G/Z(G)$ has trivial center.

These two propositions motivate a definition: we say that a Lie group map between two Lie groups G and H is an *isogeny* if it is a covering space map of the underlying manifolds; and we say two Lie groups G and H are *isogenous* if there is an isogeny between them (in either direction). Isogeny is not an equivalence relation, but generates one; observe that every isogeny equivalence class has an initial member (that is, one that maps to every other one by an isogeny)—that is, just the universal covering space $\tilde{G}$ of any one—and, if the center of this universal cover is discrete, as will be the case for all our semisimple groups, a final object $\tilde{G}/Z(\tilde{G})$ as well. For any group G in such an equivalence class, we will call $\tilde{G}$ the *simply connected form* of the group G, and $\tilde{G}/Z(\tilde{G})$ (if it exists) the *adjoint form* (we will see later a more general definition of adjoint form).

[3] This means that φ is a continuous map with the property that every point of G has a neighborhood U such that $\varphi^{-1}(U)$ is a disjoint union of open sets each mapping homeomorphically to U.

Exercise 7.12. If $H \to G$ is a covering of connected Lie groups, show that $Z(G)$ is discrete if and only if $Z(H)$ is discrete, and then $H/Z(H) = G/Z(G)$. Therefore, if $Z(G)$ is discrete, the adjoint form of G exists and is $G/Z(G)$.

To apply these ideas to some of the examples discussed, note that the center of SL_n (over $\mathbb{R}$ or $\mathbb{C}$) is just the subgroup of multiples of the identity by an nth root of unity; the quotient may be denoted $\mathrm{PSL}_n\mathbb{R}$ or $\mathrm{PSL}_n\mathbb{C}$. In the complex case, $\mathrm{PSL}_n\mathbb{C}$ is isomorphic to the quotient of $\mathrm{GL}_n\mathbb{C}$ by its center $\mathbb{C}^*$ of scalar matrices, and so one often writes $\mathrm{PGL}_n\mathbb{C}$ instead of $\mathrm{PSL}_n\mathbb{C}$. The center of the group SO_n is the subgroup $\{\pm I\}$ when n is even, and trivial when n is odd; in the former case the quotient will be denoted $\mathrm{PSO}_n\mathbb{R}$ or $\mathrm{PSO}_n\mathbb{C}$. Finally the center of the group Sp_{2n} is similarly the subgroup $\{\pm I\}$, and the quotient is denoted $\mathrm{PSp}_{2n}\mathbb{R}$ or $\mathrm{PSp}_{2n}\mathbb{C}$.

Exercise 7.13*. Realize $\mathrm{PGL}_n\mathbb{C}$ as a matrix group, i.e., find an embedding (faithful representation) $\mathrm{PGL}_n\mathbb{C} \hookrightarrow \mathrm{GL}_N\mathbb{C}$ for some N. Do the same for the other quotients above.

In the other direction, whenever we have a Lie group that is not simply connected, we can ask what its universal covering space is. This is, for example, how the famous *spin groups* arise: as we will see, the orthogonal groups $\mathrm{SO}_n\mathbb{R}$ and $\mathrm{SO}_n\mathbb{C}$ have fundamental group $\mathbb{Z}/2$, and so by the above there exist connected, two-sheeted covers of these groups. These are denoted $\mathrm{Spin}_n\mathbb{R}$ and $\mathrm{Spin}_n\mathbb{C}$, and will be discussed in Lecture 20; for the time being, the reader may find it worthwhile (if frustrating) to try to realize these as matrix groups. The last exercises of this section sketch a few steps in this direction which can be done now by hand.

Exercise 7.14. Show that the universal covering of $U(n)$ can be identified with the subgroup of the product $U(n) \times \mathbb{R}$ consisting of pairs (g, t) with $\det(g) = e^{\pi i t}$.

Exercise 7.15. We have seen in Exercise 7.4 that

$$\mathrm{SU}(2) = \mathrm{Sp}(2) = \{q \in \mathbb{H} \colon q\bar{q} = 1\}.$$

Identifying $\mathbb{R}^3$ with the imaginary quaternions (with basis i, j, k), show that, for $q\bar{q} = 1$, the map $v \mapsto qv\bar{q}$ maps $\mathbb{R}^3$ to itself, and is an isometry. Verify that the resulting map

$$\mathrm{SU}(2) = \mathrm{Sp}(2) \to \mathrm{SO}(3)$$

is a $2:1$ covering map. Since the equation $q\bar{q} = 1$ describes a 3-sphere, $\mathrm{SU}(2)$ is the universal covering of $\mathrm{SO}(3)$; and $\mathrm{SO}(3)$ is the adjoint form of $\mathrm{SU}(2)$.

Exercise 7.16. Let $M_2\mathbb{C} = \mathbb{C}^4$ be the space of 2×2 matrices, with symmetric form $Q(A, B) = \frac{1}{2}\,\mathrm{Trace}(AB^\natural)$, where $B^\natural$ is the adjoint of the matrix B; the quadratic form associated to Q is the determinant. For g and h in $\mathrm{SL}_2\mathbb{C}$, the mapping $A \mapsto gAh^{-1}$ is in $\mathrm{SO}_4\mathbb{C}$. Show that this gives a $2:1$ covering

$$\mathrm{SL}_2\mathbb{C} \times \mathrm{SL}_2\mathbb{C} \to \mathrm{SO}_4\mathbb{C},$$

which, since $\mathrm{SL}_2\mathbb{C}$ is simply connected, realizes the universal covering of $\mathrm{SO}_4\mathbb{C}$.

Exercise 7.17. Identify $\mathbb{C}^3$ with the space of traceless matrices in $M_2\mathbb{C}$, so $g \in \mathrm{SL}_2\mathbb{C}$ acts by $A \mapsto gAg^{-1}$. Show that this gives a $2:1$ covering

$$\mathrm{SL}_2\mathbb{C} \to \mathrm{SO}_3\mathbb{C},$$

which realizes the universal covering of $\mathrm{SO}_3\mathbb{C}$.

LECTURE 8

Lie Algebras and Lie Groups

In this crucial lecture we introduce the definition of the Lie algebra associated to a Lie group and its relation to that group. All three sections are logically necessary for what follows; §8.1 is essential. We use here a little more manifold theory: specifically, the differential of a map of manifolds is used in a fundamental way in §8.1, the notion of the tangent vector to an arc in a manifold is used in §8.2 and §8.3, and the notion of a vector field is introduced in an auxiliary capacity in §8.3. The Campbell-Hausdorff formula is introduced only to establish the First and Second Principles of §8.1 below; if you are willing to take those on faith the formula (and exercises dealing with it) can be skimmed. Exercises 8.27–8.29 give alternative descriptions of the Lie algebra associated to a Lie group, but can be skipped for now.

§8.1. Lie Algebras: Motivation and Definition

Given that we want to study the representations of a Lie group, how do we go about it? As we have said, the notions of generators and relations is hardly relevant here. The answer, of course, is that we have to use the continuous structure of the group. The first step in doing this is

Exercise 8.1. Let G be a connected Lie group, and $U \subset G$ any neighborhood of the identity. Show that U generates G.

This statement implies that any map $\rho: G \to H$ between connected Lie groups will be determined by what it does on any open set containing the identity in G, i.e., ρ is determined by its germ at $e \in G$. In fact, we can extend this idea a good bit further: later in this lecture we will establish the

First Principle: Let G and H be Lie groups, with G connected. A map $\rho\colon G \to H$ is uniquely determined by its differential $d\rho_e\colon T_eG \to T_eH$ at the identity.

This is, of course, great news: we can completely describe a homomorphism of Lie groups by giving a linear map between two vector spaces. It is not really worth that much, however, unless we can give at least some answer to the next, obvious question: *which maps between these two vector spaces actually arise as differentials of group homomorphisms*? The answer to this is expressed in the *Second Principle* below, but it will take us a few pages to get there. To start, we have to ask ourselves what it means for a map to be a homomorphism, and in what ways this may be reflected in the differential.

To begin with, the definition of a homomorphism is simply a $\mathscr{C}^\infty$ map ρ such that

$$\rho(gh) = \rho(g)\cdot\rho(h)$$

for all g and h in G. To express this in a more confusing way, we can say that *a homomorphism respects the action of a group on itself by left or right multiplication*: that is, for any $g \in G$ we denote by $m_g\colon G \to G$ the differentiable map given by multiplication by g, and observe that a $\mathscr{C}^\infty$ map $\rho\colon G \to H$ of Lie groups will be a homomorphism if it carries m_g to $m_{\rho(g)}$ in the sense that the diagram

$$\begin{tikzcd} G \arrow[r, "\rho"] \arrow[d, "m_g"'] & H \arrow[d, "m_{\rho(g)}"] \\ G \arrow[r, "\rho"'] & H \end{tikzcd}$$

commutes.

The problem with this characterization is that, since the maps m_g have no fixed points, it is hard to associate to them any operation on the tangent space to G at one point. This suggests looking, not at the diffeomorphisms m_g, but at the automorphisms of G given by conjugation. Explicitly, for any $g \in G$ we define the map

$$\Psi_g\colon G \to G$$

by

$$\Psi_g(h) = g\cdot h\cdot g^{-1}.$$

(Ψ_g is actually a Lie group map, but that is not relevant for our present purposes.) It is now equally the case that *a homomorphism ρ respects the action of a group G on itself by conjugation*: that is, it will carry Ψ_g into $\Psi_{\rho(g)}$ in the sense that the diagram

$$\begin{tikzcd} G \arrow[r, "\rho"] \arrow[d, "\Psi_g"'] & H \arrow[d, "\Psi_{\rho(g)}"] \\ G \arrow[r, "\rho"'] & H \end{tikzcd}$$

commutes. We have, in other words, a natural map

$$\Psi: G \to \operatorname{Aut}(G).$$

The advantage of working with Ψ_g is that it fixes the identity element $e \in G$; we can therefore extract some of its structure by looking at its differential at e: we set

$$\operatorname{Ad}(g) = (d\Psi_g)_e: T_e G \to T_e G. \tag{8.2}$$

This is a representation

$$\operatorname{Ad}: G \to \operatorname{Aut}(T_e G) \tag{8.3}$$

of the group G on its own tangent space, called the *adjoint representation* of the group. This gives a third characterization[1]: *a homomorphism ρ respects the adjoint action of a group G on its tangent space $T_e G$ at the identity.* In other words, for any $g \in G$ the actions of $\operatorname{Ad}(g)$ on $T_e G$ and $\operatorname{Ad}(\rho(g))$ on $T_e H$ must commute with the differential $(d\rho)_e$: $T_e G \to T_e H$, i.e., the diagram

$$\begin{tikzcd} T_e G \arrow[r, "(d\rho)_e"] \arrow[d, "\operatorname{Ad}(g)"'] & T_e H \arrow[d, "\operatorname{Ad}(\rho(g))"] \\ T_e G \arrow[r, "(d\rho)_e"'] & T_e H \end{tikzcd}$$

commutes; equivalently, for any tangent vector $v \in T_e G$,

$$d\rho(\operatorname{Ad}(g)(v)) = \operatorname{Ad}(\rho(g))(d\rho(v)). \tag{8.4}$$

This is nice, but does not yet answer our question, for preservation of the adjoint representation $\operatorname{Ad}: G \to \operatorname{Aut}(T_e G)$ still involves the map ρ on the group G itself, and so is not purely a condition on the differential $(d\rho)_e$. We have instead to go one step further, and *take the differential of the map* Ad. The group $\operatorname{Aut}(T_e G)$ being just an open subset of the vector space of endomorphisms of $T_e G$, its tangent space at the identity is naturally identified with $\operatorname{End}(T_e G)$; taking the differential of the map Ad we arrive at a map

$$\operatorname{ad}: T_e G \to \operatorname{End}(T_e G). \tag{8.5}$$

This is essentially a trilinear gadget on the tangent space $T_e G$; that is, we can view the image $\operatorname{ad}(X)(Y)$ of a tangent vector Y under the map $\operatorname{ad}(X)$ as a

[1] "Characterization" is not the right word here (or in the preceding case), since we do not mean an equivalent condition, but rather something implied by the condition that ρ be a homomorphism.

function of the two variables X and Y, so that we get a bilinear map

$$T_eG \times T_eG \to T_eG.$$

We use the notation $[\ ,\ ]$ for this bilinear map; that is, for a pair of tangent vectors X and Y to G at e, we write

$$[X, Y] \stackrel{\text{def}}{=} \operatorname{ad}(X)(Y). \tag{8.6}$$

As desired, the map ad involves only the tangent space to the group G at e, and so gives us our final characterization: *the differential $(d\rho)_e$ of a homomorphism ρ on a Lie group G respects the adjoint action of the tangent space to G on itself.* Explicitly, the fact that ρ and $d\rho_e$ respect the adjoint representation implies in turn that the diagram

$$\begin{tikzcd} T_eG \arrow[r, "(d\rho)_e"] \arrow[d, "\operatorname{ad}(v)"'] & T_eH \arrow[d, "\operatorname{ad}(d\rho(v))"] \\ T_eG \arrow[r, "(d\rho)_e"'] & T_eH \end{tikzcd}$$

commutes; i.e., for any pair of tangent vectors X and Y to G at e,

$$d\rho_e(\operatorname{ad}(X)(Y)) = \operatorname{ad}(d\rho_e(X))(d\rho_e(Y)). \tag{8.7}$$

or, equivalently,

$$d\rho_e([X, Y]) = [d\rho_e(X), d\rho_e(Y)]. \tag{8.8}$$

All this may be fairly confusing (if it is not, you probably do not need to be reading this book). Two things, however, should be borne in mind. They are:

(i) *It is not so bad*, in the sense that we can make the bracket operation, as defined above, reasonably explicit. We do this first for the general linear group $G = \mathrm{GL}_n\mathbb{R}$. Note that in this case conjugation extends to the ambient linear space $E = \operatorname{End}(\mathbb{R}^n) = M_n\mathbb{R}$ of $\mathrm{GL}_n\mathbb{R}$ by the same formula: $\operatorname{Ad}(g)(M) = gMg^{-1}$, and this ambient space is identified with the tangent space T_eG; differentiation in E is usual differentiation of matrices. For any pair of tangent vectors X and Y to $\mathrm{GL}_n\mathbb{R}$ at e, let $\gamma\colon I \to G$ be an arc with $\gamma(0) = e$ and tangent vector $\gamma'(0) = X$. Then our definition of $[X, Y]$ is that

$$[X, Y] = \operatorname{ad}(X)(Y) = \left.\frac{d}{dt}\right|_{t=0} (\operatorname{Ad}(\gamma(t))(Y)).$$

Applying the product rule to $\operatorname{Ad}(\gamma(t))(Y) = \gamma(t)Y\gamma(t)^{-1}$, this is

$$\begin{aligned} &= \gamma'(0)\cdot Y\cdot\gamma(0) + \gamma(0)\cdot Y\cdot(-\gamma(0)^{-1}\cdot\gamma'(0)\cdot\gamma(0)^{-1}) \\ &= X\cdot Y - Y\cdot X, \end{aligned}$$

which, of course, explains the bracket notation. In general, any time a Lie group is given as a subgroup of a general linear group $\mathrm{GL}_n\mathbb{R}$, we can view its tangent space $T_e G$ at the identity as a subspace of the space of endomorphisms of $\mathbb{R}^n$; and since bracket is preserved by (differentials of) maps of Lie groups, the bracket operation on $T_e G$ will coincide with the commutator.

(ii) *Even if it were that bad, it would be worth it.* This is because it turns out that the bracket operation is exactly the answer to the question we raised before. Precisely, later in this lecture we will prove the

Second Principle: Let G and H be Lie groups, with G connected and simply connected. A linear map $T_e G \to T_e H$ is the differential of a homomorphism $\rho: G \to H$ if and only if it preserves the bracket operation, in the sense of (8.8) above.

We are now almost done: maps between Lie groups are classified by maps between vector spaces preserving the structure of a bilinear map from the vector space to itself. We have only one more question to answer: when does a vector space with this additional structure actually arise as the tangent space at the identity to a Lie group, with the adjoint or bracket product? Happily, we have the answer to this as well. First, though it is far from clear from our initial definition, it follows from our description of the bracket as a commutator that *the bracket is skew-symmetric*, i.e, $[X, Y] = -[Y, X]$. Second, it likewise follows from the description of $[X, Y]$ as a commutator that it satisfies the *Jacobi identity*: for any three tangent vectors X, Y, and Z,

$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0.$$

We thus make the

Definition 8.9. A *Lie algebra* $\mathfrak{g}$ is a vector space together with a skew-symmetric bilinear map

$$[\ ,\ ]: \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$$

satisfying the Jacobi identity.

We should take a moment out here to make one important point. Why, you might ask, do we define the bracket operation in terms of the relatively difficult operations Ad and ad, instead of just defining $[X, Y]$ to be the commutator $X \cdot Y - Y \cdot X$? The answer is that the "composition" $X \cdot Y$ of elements of a Lie algebra is not well defined. Specifically, any time we embed a Lie group G in a general linear group $\mathrm{GL}(V)$, we get a corresponding embedding of its Lie algebra $\mathfrak{g}$ in the space $\mathrm{End}(V)$, and can talk about the composition $X \cdot Y \in \mathrm{End}(V)$ of elements of $\mathfrak{g}$ in this context; but it must be borne in mind that this composition $X \cdot Y$ will depend on the embedding of $\mathfrak{g}$, and for that matter need not even be an element of $\mathfrak{g}$. *Only* the commutator $X \cdot Y - Y \cdot X$ is always an element of $\mathfrak{g}$, independent of the representation. The terminology sometimes heightens the confusion: for example, when we speak of embedding a Lie algebra in the algebra $\mathrm{End}(V)$ of endomorphisms of V, the word *algebra* may mean two very different things. In general, when we want to refer to the endomorphisms of a vector space V (resp. $\mathbb{R}^n$) as a Lie algebra, we will write $\mathfrak{gl}(V)$ (resp. $\mathfrak{gl}_n\mathbb{R}$) instead of $\mathrm{End}(V)$ (resp. $M_n\mathbb{R}$).

To return to our discussion of Lie algebras, a *map* of Lie algebras is a linear map of vector spaces preserving the bracket, in the sense of (8.8); notions like *Lie subalgebra* are defined accordingly. We note in passing one thing that will turn out to be significant: the definition of Lie algebra does not specify the field. Thus, we have real Lie algebras, complex Lie algebras, etc., all defined in the same way; and in addition, given a real Lie algebra $\mathfrak{g}$ we may associate to it a complex Lie algebra, whose underlying vector space is $\mathfrak{g} \otimes \mathbb{C}$ and whose bracket operation is just the bracket on $\mathfrak{g}$ extended by linearity.

Exercise 8.10*. The skew-commutativity and Jacobi identity also follow from the naturality of the bracket (8.8), without using an embedding in $\mathfrak{gl}(V)$:

(a) Deduce the skew-commutativity $[X, X] = 0$ from that fact that any X can be written the image of a vector by $d\rho_e$ for some homomorphism $\rho\colon \mathbb{R} \to G$. (See §8.3 for the existence of ρ.)
(b) Given that the bracket is skew-commutative, verify that the Jacobi identity is equivalent to the assertion that

$$\mathrm{ad} = d(\mathrm{Ad})_e\colon \mathfrak{g} \to \mathrm{End}(\mathfrak{g})$$

preserves the bracket. In particular, ad is a map of Lie algebras.

To sum up our progress so far: taking for the moment on faith the statements made, we have seen that

(i) the tangent space $\mathfrak{g}$ at the identity to a Lie group G is naturally endowed with the structure of a Lie algebra;
(ii) if G and H are Lie groups with G connected and simply connected, the maps from G to H are in one-to-one correspondence with maps of the associated Lie algebras, by associating to $\rho\colon G \to H$ its differential $(d\rho)_e\colon \mathfrak{g} \to \mathfrak{h}$.

Of course, we make the

Definition 8.11. A *representation* of a Lie algebra $\mathfrak{g}$ on a vector space V is simply a map of Lie algebras

$$\rho\colon \mathfrak{g} \to \mathfrak{gl}(V) = \mathrm{End}(V),$$

i.e., a linear map that preserves brackets, or an action of $\mathfrak{g}$ on V such that

$$[X, Y](v) = X(Y(v)) - Y(X(v)).$$

Statement (ii) above implies in particular that *representations of a connected and simply connected Lie group are in one-to-one correspondence with repre-*

sentations of its Lie algebra. This is, then, the first step of the series of reductions outlined in the introduction to Part II.

At this point, a few words are in order about the relation between representations of a Lie group and the corresponding representations of its Lie algebra. The first remark to make is about tensors. Recall that if V and W are representations of a Lie group G, then we define the representation $V \otimes W$ to be the vector space $V \otimes W$ with the action of G described by

$$g(v \otimes w) = g(v) \otimes g(w).$$

The definition for representations of a Lie algebra, however, is quite different. For one thing, if $\mathfrak{g}$ is the Lie algebra of G, so that the representation of G on the vector spaces V and W induces representations of $\mathfrak{g}$ on these spaces, we want the tensor product of the representations V and W of $\mathfrak{g}$ to be the representation induced by the action of G on $V \otimes W$ above. But now suppose that $\{\gamma_t\}$ is an arc in G with $\gamma_0 = e$ and tangent vector $\gamma_0' = X \in \mathfrak{g}$. Then by definition the action of X on V is given by

$$X(v) = \frac{d}{dt}\bigg|_{t=0} \gamma_t(v)$$

and similarly for $w \in W$; it follows that the action of X on the tensor product $v \otimes w$ is

$$\begin{aligned} X(v \otimes w) &= \frac{d}{dt}\bigg|_{t=0} (\gamma_t(v) \otimes \gamma_t(w)) \\ &= \left(\frac{d}{dt}\bigg|_{t=0} \gamma_t(v)\right) \otimes w + v \otimes \left(\frac{d}{dt}\bigg|_{t=0} \gamma_t(w)\right), \end{aligned}$$

so

$$X(v \otimes w) = X(v) \otimes w + v \otimes X(w). \tag{8.12}$$

This, then, is how we *define* the action of a Lie algebra $\mathfrak{g}$ on the tensor product of two representations of $\mathfrak{g}$. This describes as well other tensors: for example, if V is a representation of the group G, $v \in V$ is any vector and $v^2 \in \mathrm{Sym}^2 V$ its square, then for any $g \in G$,

$$g(v^2) = g(v)^2.$$

On the other hand, if V is a representation of the Lie algebra $\mathfrak{g}$ and $X \in \mathfrak{g}$ is any element, we have

$$X(v^2) = 2 \cdot v \cdot X(v). \tag{8.13}$$

One further example: if $\rho: G \to \mathrm{GL}(V)$ is a representation of the group G, the dual representation $\rho': G \to \mathrm{GL}(V^*)$ is defined by setting

$$\rho'(g) = {}^t\rho(g^{-1}): V^* \to V^*.$$

Differentiating this, we find that if $\rho: \mathfrak{g} \to \mathfrak{gl}(V)$ is a representation of a Lie algebra $\mathfrak{g}$, the dual representation of $\mathfrak{g}$ on V^* will be given by

$$\rho'(X) = {}^t\rho(-X) = -{}^t\rho(X)\colon V^* \to V^*. \tag{8.14}$$

A second and related point to be made concerns terminology. Obviously, when we speak of the action of a group G on a vector space V preserving some extra structure on V, we mean that literally: for example, if we have a quadratic form Q on V, to say that G preserves Q means just that

$$Q(g(v), g(w)) = Q(v, w), \quad \forall g \in G \text{ and } v, w \in V.$$

Equivalently, we mean that the associated action of G on the vector space $\mathrm{Sym}^2 V^*$ fixes the element $Q \in \mathrm{Sym}^2 V^*$. But by the above calculation, the action of the associated Lie algebra $\mathfrak{g}$ on V satisfies

$$Q(v, X(w)) + Q(X(v), w) = 0, \quad \forall X \in \mathfrak{g} \text{ and } v, w \in V \tag{8.15}$$

or, equivalently, $Q(v, X(v)) = 0$ for all $X \in \mathfrak{g}$ and $v \in V$; in other words, *the induced action on* $\mathrm{Sym}^2 V^*$ *kills the element* Q. By way of terminology, then, *we will in general say that the action of a Lie algebra on a vector space preserves some structure when a corresponding Lie group action does.*

The next section will be spent in giving examples. In §8.3 we will establish the basic relations between Lie groups and their Lie algebras, to the point where we can prove the First and Second Principles above. The further statement that any Lie algebra is the Lie algebra of some Lie group will follow from the statement (see Appendix E) that every Lie algebra may be embedded in $\mathfrak{gl}_n\mathbb{R}$.

Exercise 8.16*. Show that if G is connected the image of $\mathrm{Ad}\colon G \to \mathrm{GL}(\mathfrak{g})$ is the adjoint form of the group G when that exists.

Exercise 8.17*. Let V be a representation of a connected Lie group G and $\rho\colon \mathfrak{g} \to \mathrm{End}(V)$ the corresponding map of Lie algebras. Show that a subspace W of V is invariant by G if and only if it is carried into itself under the action of the Lie algebra $\mathfrak{g}$, i.e., $\rho(X)(W) \subset W$ for all X in $\mathfrak{g}$. Hence, V is irreducible over G if and only if it is irreducible over $\mathfrak{g}$.

§8.2. Examples of Lie Algebras

We start with the Lie algebras associated to each of the groups mentioned in Lecture 7. Each of these groups is given as a subgroup of $\mathrm{GL}(V) = \mathrm{GL}_n\mathbb{R}$, so their Lie algebras will be subspaces of $\mathrm{End}(V) = \mathfrak{gl}_n\mathbb{R}$.

Consider first the special linear group $\mathrm{SL}_n\mathbb{R}$. If $\{A_t\}$ is an arc in $\mathrm{SL}_n\mathbb{R}$ with $A_0 = I$ and tangent vector $A_0' = X$ at $t = 0$, then by definition we have for any basis $e_1, \ldots, e_n$ of $V = \mathbb{R}^n$,

$$A_t(e_1) \wedge \cdots \wedge A_t(e_n) \equiv e_1 \wedge \cdots \wedge e_n.$$

Taking the derivative and evaluating at $t = 0$ we have by the product rule

$$\begin{aligned}0 &= \frac{d}{dt}\bigg|_{t=0} (A_t(e_1) \wedge \cdots \wedge A_t(e_n)) \\ &= \sum e_1 \wedge \cdots \wedge X(e_i) \wedge \cdots \wedge e_n \\ &= \operatorname{Trace}(X)\cdot(e_1 \wedge \cdots \wedge e_n).\end{aligned}$$

The tangent vectors to $\mathrm{SL}_n\mathbb{R}$ are thus all endomorphisms of trace 0; comparing dimensions we can see that the Lie algebra $\mathfrak{sl}_n\mathbb{R}$ is exactly the vector space of traceless $n \times n$ matrices.

The orthogonal and symplectic cases are somewhat simpler. For example, the orthogonal group $\mathrm{O}_n\mathbb{R}$ is defined to be the automorphisms A of an n-dimensional vector space V preserving a quadratic form Q, so that if $\{A_t\}$ is an arc in $\mathrm{O}_n\mathbb{R}$ with $A_0 = I$ and $A'_0 = X$ we have for every pair of vectors v, $w \in V$

$$Q(A_t(v), A_t(w)) \equiv Q(v, w).$$

Taking derivatives, we see that

$$Q(X(v), w) + Q(v, X(w)) = 0 \tag{8.18}$$

for all $v, w \in V$; this is exactly the condition that describes the orthogonal Lie algebra $\mathfrak{so}_n\mathbb{R} = \mathfrak{o}_n\mathbb{R}$. In coordinates, if the quadratic form Q is given on $V = \mathbb{R}^n$ as

$$Q(v, w) = {}^t v \cdot M \cdot w \tag{8.19}$$

for some symmetric $n \times n$ matrix M, then as we have seen the condition on $A \in \mathrm{GL}_n\mathbb{R}$ to be in $\mathrm{O}_n\mathbb{R}$ is that

$${}^t A \cdot M \cdot A = M. \tag{8.20}$$

Differentiating, the condition on an $n \times n$ matrix X to be in the Lie algebra $\mathfrak{so}_n\mathbb{R}$ of the orthogonal group is that

$${}^t X \cdot M + M \cdot X = 0. \tag{8.21}$$

Note that if M is the identity matrix—i.e., Q is the "standard" quadratic form $Q(v, w) = {}^t v \cdot w$ on $\mathbb{R}^n$—then this says that $\mathfrak{so}_n\mathbb{R}$ *is the subspace of skew-symmetric* $n \times n$ matrices. To put it intrinsically, in terms of the identification of V with V^* given by the quadratic form Q, and the consequent identification $\operatorname{End}(V) = V \otimes V^* = V \otimes V$, the Lie algebra $\mathfrak{so}_n\mathbb{R} \subset \operatorname{End}(V)$ is just the subspace $\bigwedge^2 V \subset V \otimes V$ of skew-symmetric tensors:

$$\mathfrak{so}_n\mathbb{R} = \textstyle\bigwedge^2 V \subset \operatorname{End}(V) = V \otimes V. \tag{8.22}$$

All of the above, with the exception of the last paragraph, works equally well to describe the Lie algebra $\mathfrak{sp}_{2n}\mathbb{R}$ of the Lie group $\mathrm{Sp}_{2n}\mathbb{R}$ of transformations preserving a skew-symmetric bilinear form Q; that is, $\mathfrak{sp}_{2n}\mathbb{R}$ is the subspace of endomorphisms of V satisfying (8.18) for every pair of vectors v, $w \in V$, or, if Q is given by a skew-symmetric $2n \times 2n$ matrix M as in (8.19), the space of matrices satisfying (8.21). The one statement that has to be substantially modified is the last one of the last paragraph: because Q is skew-symmetric, condition (8.18) is equivalent to saying that

$$Q(X(v), w) = Q(X(w), v)$$

for all v, $w \in V$; thus, in terms of the identification of V with V^* given by Q, the Lie algebra $\mathfrak{sp}_{2n}\mathbb{R} \subset \mathrm{End}(V) = V \otimes V^* = V \otimes V$ is the subspace $\mathrm{Sym}^2 V \subset V \otimes V$:

$$\mathfrak{sp}_{2n}\mathbb{R} = \mathrm{Sym}^2 V \subset \mathrm{End}(V) = V \otimes V. \tag{8.23}$$

Exercise 8.24*. With Q a standard skew form, say of Exercise 7.3, describe $\mathrm{Sp}_{2n}\mathbb{R}$ and its Lie algebra $\mathfrak{sp}_{2n}\mathbb{R}$ (as subgroup of $\mathrm{GL}_{2n}\mathbb{R}$ and subalgebra of $\mathfrak{gl}_{2n}\mathbb{R}$). Do a corresponding calculation for $\mathrm{SO}_{k,l}\mathbb{R}$.

One more similar example is that of the Lie algebra $\mathfrak{u}_n$ of the unitary group $\mathrm{U}(n)$; by a similar calculation we find that the Lie algebra of complex linear endomorphisms of $\mathbb{C}^n$ preserving a Hermitian inner product H is just the space of matrices X satisfying

$$H(X(v), w) + H(v, X(w)) = 0, \quad \forall v, w \in V;$$

if H is given by $H(v, w) = {}^t\bar{v} \cdot w$, this amounts to saying that X is conjugate skew-symmetric, i.e., that ${}^t\bar{X} = -X$.

Exercise 8.25. Find the Lie algebras of the real Lie groups $\mathrm{SL}_n\mathbb{C}$ and $\mathrm{SL}_n\mathbb{H}$—the elements in $\mathrm{GL}_n\mathbb{H}$ whose real determinant is 1.

Exercise 8.26. Show that the Lie algebras of the Lie groups B_n and N_n introduced in §7.2 are the algebra $\mathfrak{b}_n\mathbb{R}$ of upper triangular $n \times n$ matrices and the algebra $\mathfrak{n}_n\mathbb{R}$ of strictly upper triangular $n \times n$ matrices, respectively.

If G is a complex Lie group, its Lie algebra is a complex Lie algebra. Just as in the real case, we have the complex Lie algebras $\mathfrak{gl}_n\mathbb{C}$, $\mathfrak{sl}_n\mathbb{C}$, $\mathfrak{so}_m\mathbb{C}$, and $\mathfrak{sp}_{2n}\mathbb{C}$ of the Lie groups $\mathrm{GL}_n\mathbb{C}$, $\mathrm{SL}_n\mathbb{C}$, $\mathrm{SO}_m\mathbb{C}$, and $\mathrm{Sp}_{2n}\mathbb{C}$.

Exercise 8.27. Let A be any (real or complex) algebra, not necessarily finite dimensional, or even associative. A *derivation* is a linear map $D: A \to A$ satisfying the Leibniz rule $D(ab) = aD(b) + D(a)b$.

(a) Show that the derivations $\mathrm{Der}(A)$ form a Lie algebra under the bracket $[D, E] = D \circ E - E \circ D$. If A is finite dimensional, so is $\mathrm{Der}(A)$.
(b) The group of automorphisms of A is a closed subgroup G of the group $\mathrm{GL}(A)$ of linear automorphisms of A. Show that the Lie algebra of G is $\mathrm{Der}(A)$.
(c) If the algebra A is a Lie algebra, the map $A \to \mathrm{Der}(A)$, $X \mapsto D_X$, where $D_X(Y) = [X, Y]$, is a map of Lie algebras.

Exercise 8.28*. If $\mathfrak{g}$ is a Lie algebra, the Lie algebra automorphisms of $\mathfrak{g}$ form a Lie subgroup $\mathrm{Aut}(\mathfrak{g})$ of the general linear group $GL(\mathfrak{g})$.

(a) Show that the Lie algebra of $\mathrm{Aut}(\mathfrak{g})$ is $\mathrm{Der}(\mathfrak{g})$. If G is a simply connected Lie group with Lie algebra $\mathfrak{g}$, the map $\mathrm{Aut}(G) \to \mathrm{Aut}(\mathfrak{g})$ by $\varphi \mapsto d\varphi$ is one-to-one and onto, giving $\mathrm{Aut}(G)$ the structure of a Lie group with Lie algebra $\mathrm{Der}(\mathfrak{g})$.

(b) Show that the automorphism group of any connected Lie group is a Lie subgroup of the automorphism group of its Lie algebra.

Exercise 8.29*. For any manifold M, the C^∞ vector fields on M form a Lie algebra $\mathfrak{v}(M)$, as follows: a vector field v can be identified with a derivation of the ring A of C^∞ functions on M, with $v(f)$ the function whose value at a point x of M is the value of the tangent vector v_x on f at x. Show that the vector fields on M form a Lie algebra, in fact a Lie subalgebra of the Lie algebra $\mathrm{Der}(A)$. If a Lie group G acts on M, the G-invariant vector fields form a Lie subalgebra $\mathfrak{v}_G M$ of $\mathfrak{v}(M)$. If the action is transitive, the invariant vector fields form a finite-dimensional Lie algebra.

If G is a Lie group, $\mathfrak{v}_G(G) = T_eG$ becomes a Lie algebra by the above process. Show that this bracket agrees with that defined using the adjoint map (8.6). This gives another proof that the bracket is skew-symmetric and satisfies Jacobi's identity.

§8.3. The Exponential Map

The essential ingredient in studying the relationship between a Lie group G and its Lie algebra $\mathfrak{g}$ is the exponential map. This may be defined in very straightforward fashion, using the notion of *one-parameter subgroups*, which we study next. Suppose that $X \in \mathfrak{g}$ is any element, viewed simply as a tangent vector to G at the identity. For any element $g \in G$, denote by $m_g\colon G \to G$ the map of manifolds given by multiplication on the left by g. Then we can define a vector field v_X on all of G simply by setting

$$v_X(g) = (m_g)_*(X).$$

This vector field is clearly invariant under left translation (i.e., it is carried into itself under the diffeomorphism m_g for all g); and it is not hard to see that this gives an identification of $\mathfrak{g}$ with the space of all left-invariant vector fields on G. Under this identification, the bracket operation on the Lie algebra $\mathfrak{g}$ corresponds to Lie bracket of vector fields; indeed, this may be adopted as the definition of the Lie algebra associated to a Lie group (cf. Exercise 8.29). For our present purposes, however, all we need to know is that v_X exists and is left-invariant.

Given any vector field v on a manifold M and a point $p \in M$, a basic theorem from differential equations allows us to integrate the vector field. This gives a differentiable map $\varphi: I \to M$, defined on some open interval I containing 0, with $\varphi(0) = p$, whose tangent vector at any point is the vector assigned to that point by v, i.e., such that

$$\varphi'(t) = v(\varphi(t))$$

for all t in I. The map φ is uniquely characterized by these properties. Now suppose the manifold in question is a Lie group G, the vector field the field v_X associated to an element $X \in \mathfrak{g}$, and p the identity. We arrive then at a map $\varphi: I \to G$; we claim that, at least where φ is defined, it is a *homomorphism*, i.e., $\varphi(s + t) = \varphi(s)\varphi(t)$ whenever s, t, and $s + t$ are in I. To prove this, fix s and let t vary; that is, consider the two arcs α and β given by $\alpha(t) = \varphi(s)\cdot\varphi(t)$ and $\beta(t) = \varphi(s + t)$. Of course, $\alpha(0) = \beta(0)$; and by the invariance of the vector field v_X, we see that the tangent vectors satisfy $\alpha'(t) = v_X(\alpha(t))$ and $\beta'(t) = v_X(\beta(t))$ for all t. By the uniqueness of the integral curve of a vector field on a manifold, we deduce that $\alpha(t) = \beta(t)$ for all t.

From the fact that $\varphi(s + t) = \varphi(s)\varphi(t)$ for all s and t near 0, it follows that φ extends uniquely to all of $\mathbb{R}$, defining a homomorphism

$$\varphi_X: \mathbb{R} \to G$$

with $\varphi_X'(t) = v_X(\varphi)(t) = (m_{\varphi(t)})_*(X)$ for all t.

Exercise 8.30. Establish the *product rule* for derivatives of arcs in a Lie group G: if α and β are arcs in G and $\gamma(t) = \alpha(t)\cdot\beta(t)$, then

$$\gamma'(t) = dm_{\alpha(t)}(\beta'(t)) + dn_{\beta(t)}(\alpha'(t)),$$

where for any $g \in G$, the map m_g (resp. n_g): $G \to G$ is given by left (resp. right) multiplication by g. Use this to give another proof that φ is a homomorphism.

Exercise 8.31. Show that φ_X is uniquely determined by the fact that it is a homomorphism of $\mathbb{R}$ to G with tangent vector $\varphi_X'(0)$ at the identity equal to X. Deduce that if $\psi: G \to H$ is a map of Lie groups, then $\varphi_{\psi_* X} = \psi \circ \varphi_X$.

The Lie group map $\varphi_X: \mathbb{R} \to G$ is called the *one-parameter subgroup of* G *with tangent vector* X *at the identity*. The construction of these one-parameter subgroups for each X amounts to the verification of the Second Principle of §8.1 for homomorphisms from $\mathbb{R}$ to G. The fact that there exists such a one-parameter subgroup of G with any given tangent vector at the identity is crucial. For example, it is not hard to see (we will do this in a moment) that these one-parameter subgroups fill up a neighborhood of the identity in G, which immediately implies the First Principle of §8.1. To carry this out, we define the *exponential map*

$$\exp: \mathfrak{g} \to G$$

by

$$\exp(X) = \varphi_X(1). \tag{8.32}$$

Note that by the uniqueness of φ_X, we have

$$\varphi_{(\lambda X)}(t) = \varphi_X(\lambda t);$$

so that the exponential map restricted to the lines through the origin in $\mathfrak{g}$ gives the one-parameter subgroups of G. Indeed, Exercise 8.31 implies the characterization:

Proposition 8.33. *The exponential map is the unique map from* $\mathfrak{g}$ *to* G *taking* 0 *to* e *whose differential at the origin*

$$(\exp_*)_0 \colon T_0\mathfrak{g} = \mathfrak{g} \to T_eG = \mathfrak{g}$$

is the identity, and whose restrictions to the lines through the origin in $\mathfrak{g}$ *are one-parameter subgroups of* G.

This in particular implies (cf. Exercise 8.31) that the exponential map is natural, in the sense that for any map $\psi \colon G \to H$ of Lie groups the diagram

$$\begin{tikzcd} \mathfrak{g} \arrow[r, "\psi_*"] \arrow[d, "\exp"'] & \mathfrak{h} \arrow[d, "\exp"] \\ G \arrow[r, "\psi"'] & H \end{tikzcd}$$

commutes.

Now, since the differential of the exponential map at the origin in $\mathfrak{g}$ is an isomorphism, the image of exp will contain a neighborhood of the identity in G. If G is connected, this will generate all of G; from this follows the First Principle: *if G is connected, then the map ψ is determined by its differential $(d\psi)_e$ at the identity.*

Using (8.32), we can write down the exponential map very explicitly in the case of $\mathrm{GL}_n\mathbb{R}$, and hence for any subgroup of $\mathrm{GL}_n\mathbb{R}$. We just use the standard power series for the function e^x, and set, for $X \in \mathrm{End}(V)$,

$$\exp(X) = 1 + X + \frac{X^2}{2} + \frac{X^3}{6} + \cdots. \tag{8.34}$$

Observe that this converges and is invertible, with inverse $\exp(-X)$. Clearly, the differential of this map from $\mathfrak{g}$ to G at the origin is the identity; and by the standard power series computation, the restriction of the map to any line through the origin in $\mathfrak{g}$ is a one-parameter subgroup of G. Thus, the map coincides with the exponential as defined originally; and by naturality the same is true for any subgroup of G. (Note that, as we have pointed out, the individual terms in the expression on the right of (8.34) are very much dependent of the particular embedding of G in a general linear group $\mathrm{GL}(V)$ and correspondingly of $\mathfrak{g}$ in $\mathrm{End}(V)$, even though the *sum* on the right in (8.34) is not.)

This explicit form of the exponential map allows us to give substance to the assertion that "the group structure of G is encoded in the Lie algebra." Explicitly, we claim that not only do the exponentials $\exp(X)$ generate G, but for X and Y in a sufficiently small neighborhood of the origin in $\mathfrak{g}$, we can write down the product $\exp(X)\cdot\exp(Y)$ as an exponential. To do this, we introduce first the "inverse" of the exponential map: for $g \in G \subset \mathrm{GL}_n\mathbb{R}$, we set

$$\log(g) = (g - I) - \frac{(g - I)^2}{2} + \frac{(g - I)^3}{3} - \cdots \in \mathfrak{gl}_n\mathbb{R}.$$

Of course, this will be defined only for g sufficiently close to the identity in G; but where it is defined it will be an inverse to the exponential map.

Now, we define a new bilinear operation on $\mathfrak{gl}_n\mathbb{R}$: we set

$$X * Y = \log(\exp(X)\cdot\exp(Y)).$$

We have to be careful what we mean by this, of course; we substitute for g in the expression above for $\log(g)$ the quantity

$$\begin{aligned}\exp(X)\cdot\exp(Y) &= \left(I + X + \frac{X^2}{2} + \cdots\right)\cdot\left(I + Y + \frac{Y^2}{2} + \cdots\right)\\ &= I + (X + Y) + \left(\frac{X^2}{2} + X\cdot Y + \frac{Y^2}{2}\right) + \cdots,\end{aligned}$$

being careful, of course, to preserve the order of the factors in each product. Doing this, we arrive at

$$\begin{aligned}X * Y &= (X + Y) + \left(-\frac{(X + Y)^2}{2} + \left(\frac{X^2}{2} + X\cdot Y + \frac{Y^2}{2}\right)\right) + \cdots\\ &= X + Y + \tfrac{1}{2}[X, Y] + \cdots.\end{aligned}$$

Observe in particular that the terms of degree 2 in X and Y do not involve the squares of X and Y or the product $X \cdot Y$ alone, but only the commutator. In fact, this is true of each term in the formula, i.e., the quantity $\log(\exp(X)\cdot\exp(Y))$ can be expressed purely in terms of X, Y, and the bracket operation; the resulting formula is called the *Campbell–Hausdorff formula* (although the actual formula in closed form was given by Dynkin). To degree three, it is

$$X * Y = X + Y + \tfrac{1}{2}[X, Y] \pm \tfrac{1}{12}[X, [X, Y]] \pm \tfrac{1}{12}[Y, [Y, X]] + \cdots.$$

Exercise 8.35*. Verify (and find the correct signs in) the cubic term of the Campbell–Hausdorff formula.

Exercise 8.36. Prove the assertion of the last paragraph that the power series $\log(\exp(X)\cdot\exp(Y))$ can be expressed purely in terms of X, Y, and the bracket operation.

Exercise 8.37. Show that for X and Y sufficiently small, the power series $\log(\exp(X)\cdot\exp(Y))$ converges.

Exercise 8.38*. (a) Show that there is a constant C such that for $X, Y \in \mathfrak{gl}_n$, $X * Y = X + Y + [X, Y] + E$, where $\|E\| \leq C(\|X\| + \|Y\|)^3$.

(b) Show that $\exp(X + Y) = \lim_{n\to\infty} (\exp(X/n)\cdot \exp(Y/n))^n$.

(c) Show that

$$\exp([X, Y]) = \lim_{n\to\infty} \left(\exp\left(\frac{X}{n}\right)\cdot\exp\left(\frac{Y}{n}\right)\cdot\exp\left(-\frac{X}{n}\right)\cdot\exp\left(-\frac{Y}{n}\right)\right)^{n^2}.$$

Exercise 8.39. Show that if G is a subgroup of $\mathrm{GL}_n\mathbb{R}$, the elements of its Lie algebra are the "infinitesimal transformations" of G in the sense of von Neumann, i.e., they are the matrices in $\mathfrak{gl}_n\mathbb{R}$ which can be realized as limits

$$\lim_{t\to 0} \frac{A_t - I}{\varepsilon_t}, \quad A_t \in G,\ \varepsilon_t > 0,\ \varepsilon_t \to 0.$$

Exercise 8.40. Show that exp is surjective for $G = \mathrm{GL}_n\mathbb{C}$ but not for $G = \mathrm{GL}_n^+\mathbb{R}$ if $n > 1$, or for $G = \mathrm{SL}_2\mathbb{C}$.

By the Campbell–Hausdorff formula, we can not only identify all the elements of G in a neighborhood of the identity, but we can also say what their pairwise products are, thus making precise the sense in which $\mathfrak{g}$ and its bracket operation determines G and its group law locally. Of course, we have not written a closed-form expression for the Campbell–Hausdorff formula; but, as we will see shortly, its very existence is significant. (For such a closed form, see [Se1, I§4.8].)

We now consider another very natural question, namely, when a vector subspace $\mathfrak{h} \subset \mathfrak{g}$ is the Lie algebra of (i.e., tangent space at the identity to) an immersed subgroup of G. Obviously, a necessary condition is that $\mathfrak{h}$ is closed under the bracket operation; we claim here that this is sufficient as well:

Proposition 8.41. *Let G be a Lie group, $\mathfrak{g}$ its Lie algebra, and $\mathfrak{h} \subset \mathfrak{g}$ a Lie subalgebra. Then the subgroup of the group G generated by* $\exp(\mathfrak{h})$ *is an immersed subgroup H with tangent space $T_eH = \mathfrak{h}$.*

Proof. Note that the subgroup generated by $\exp(\mathfrak{h})$ is the same as the subgroup generated by $\exp(U)$, where U is any neighborhood of the origin in $\mathfrak{h}$. It will suffice, then (see Exercise 8.42), to show that the image of $\mathfrak{h}$ under the exponential map is "locally" closed under multiplication, i.e., that for a sufficiently small disc $\Delta \subset \mathfrak{h}$, the product $\exp(\Delta)\cdot\exp(\Delta)$ (that is, the set of pairwise products $\exp(X)\cdot\exp(Y)$ for $X, Y \in \Delta$) is contained in the image of $\mathfrak{h}$ under the exponential map.

We will do this under the hypothesis that G may be realized as a subgroup of a general linear group $\mathrm{GL}_n\mathbb{R}$, so that we can use the formula (8.34) for the exponential map. This is a harmless assumption, given the statement (to be proved in Appendix E) that any finite-dimensional Lie algebra may be embedded in the Lie algebra $\mathfrak{gl}_n\mathbb{R}$: the subgroup of $\mathrm{GL}_n\mathbb{R}$ generated by $\exp(g)$ will be a group isogenous to G, and, as the reader can easily check, proving the proposition for a group isogenous to G is equivalent to proving it for G.

It thus suffices to prove the assertion in case the group G is $\mathrm{GL}_n\mathbb{R}$. But this is exactly the content of the Campbell–Hausdorff formula. □

When applied to an embedding of a Lie algebra $\mathfrak{g}$ into $\mathfrak{gl}_n$, we see, in particular, that *every finite-dimensional Lie algebra is the Lie algebra of a Lie group.* From what we have seen, this Lie group is unique if we require it to be simply connected, and then all others are obtained by dividing this simply connected model by a discrete subgroup of its center.

Exercise 8.42*. Let $G_0 = \exp(\Delta)$, where Δ is a disk centered at the origin in $\mathfrak{g}$, and let $H_0 = \exp(\Delta \cap \mathfrak{h})$. Show that $G_0^{-1} = G_0$, $H_0^{-1} = H_0$, and $H_0 \cdot H_0 \cap G_0 = H_0$. Use this to show that the subgroup H of G generated by H_0 is an immersed Lie subgroup of G.

As a fairly easy consequence of this proposition, we can finally give a proof of the Second Principle stated in §8.1, which we may restate as

Second Principle. *Let G and H be Lie groups with G simply connected, and let $\mathfrak{g}$ and $\mathfrak{h}$ be their Lie algebras. A linear map $\alpha\colon \mathfrak{g} \to \mathfrak{h}$ is the differential of a map $A\colon G \to H$ of Lie groups if and only if α is a map of Lie algebras.*

PROOF. To see this, consider the product $G \times H$. Its Lie algebra is just $\mathfrak{g} \oplus \mathfrak{h}$. Let $\mathfrak{j} \subset \mathfrak{g} \oplus \mathfrak{h}$ be the graph of the map α. Then the hypothesis that α is a map of Lie algebras is equivalent to the statement that $\mathfrak{j}$ is a Lie subalgebra of $\mathfrak{g} \oplus \mathfrak{h}$; and given this, by the proposition there exists an immersed Lie subgroup $J \subset G \times H$ with tangent space $T_eJ = \mathfrak{j}$.

Look now at the map $\pi\colon J \to G$ given by projection on the first factor. By hypothesis, the differential of this map $d\pi_e\colon \mathfrak{j} \to \mathfrak{g}$ is an isomorphism, so that the map $J \to G$ is an isogeny; but since G is simply connected *it follows that π is an isomorphism.* The projection $\eta\colon G \cong J \to H$ on the second factor is then a Lie group map whose differential at the identity is α. □

Exercise 8.43*. If $\mathfrak{g} \to \mathfrak{g}'$ is a homomorphism of Lie algebras with kernel $\mathfrak{h}$, show that the kernel H of the corresponding map of simply connected Lie groups $G \to G'$ is a closed subgroup of G with Lie group $\mathfrak{h}$. This does not extend to non-normal subgroups, i.e., to the situation when $\mathfrak{h}$ is not the kernel of a homomorphism: give an example of an immersed subgroup of a simply connected Lie group G whose image in G is not closed.

Exercise 8.44. Use the ideas of this lecture to prove the assertion that a compact complex connected Lie group G must be abelian:

(a) Verify that the map Ad: $G \to \mathrm{Aut}(T_eG) \subset \mathrm{End}(T_eG)$ is holomorphic, and, therefore (by the maximum principle), constant.
(b) Deduce that if Ψ_g is conjugation by g, then $d\Psi_g$ is the identity, so $\Psi_g(\exp(X)) = \exp(d\Psi_g(X)) = \exp(X)$ for all $X \in T_eG$, which implies that G is abelian.
(c) Show that the exponential map from T_eG to G is surjective, with the kernel a lattice Λ, so $G = T_eG/\Lambda$ is a complex torus.

LECTURE 9

Initial Classification of Lie Algebras

In this lecture we define various subclasses of Lie algebras: nilpotent, solvable, semisimple, etc., and prove basic facts about their representations. The discussion is entirely elementary (largely because the hard theorems are stated without proof for now); there are no prerequisites beyond linear algebra. Apart from giving these basic definitions, the purpose of the lecture is largely to motivate the narrowing of our focus to semisimple algebras that will take place in the sequel. In particular, the first part of §9.3 is logically the most important for what follows.

§9.1. Rough Classification of Lie Algebras

We will give, in this section, a preliminary sort of classification of Lie algebras, reflecting the degree to which a given Lie algebra $\mathfrak{g}$ fails to be abelian. As we have indicated, the goal ultimately is to narrow our focus onto *semisimple* Lie algebras.

Before we begin, two definitions, both completely straightforward: First, we define the *center* $Z(\mathfrak{g})$ of a Lie algebra $\mathfrak{g}$ to be the subspace of $\mathfrak{g}$ of elements $X \in \mathfrak{g}$ such that $[X, Y] = 0$ for all $Y \in \mathfrak{g}$. Of course, we say $\mathfrak{g}$ is *abelian* if all brackets are zero.

Exercise 9.1. Let G be a Lie group, $\mathfrak{g}$ its Lie algebra. Show that the subgroup of G generated by exponentiating the Lie subalgebra $Z(\mathfrak{g})$ is the connected component of the identity in the center $Z(G)$ of G.

Next, we say that a Lie subalgebra $\mathfrak{h} \subset \mathfrak{g}$ of a Lie algebra $\mathfrak{g}$ is an *ideal* if it satisfies the condition

$$[X, Y] \in \mathfrak{h} \quad \text{for all } X \in \mathfrak{h},\ Y \in \mathfrak{g}.$$

Just as connected subgroups of a Lie group correspond to subalgebras of its Lie algebra, the notion of ideal in a Lie algebra corresponds to the notion of normal subgroup, in the following sense:

Exercise 9.2. Let G be a connected Lie group, $H \subset G$ a connected subgroup and $\mathfrak{g}$ and $\mathfrak{h}$ their Lie algebras. Show that H is a normal subgroup of G if and only if $\mathfrak{h}$ is an ideal of $\mathfrak{g}$.

Observe also that the bracket operation on $\mathfrak{g}$ induces a bracket on the quotient space $\mathfrak{g}/\mathfrak{h}$ if and only if $\mathfrak{h}$ is an ideal in $\mathfrak{g}$.

This, in turns, motivates the next bit of terminology: we say that a Lie algebra $\mathfrak{g}$ is *simple* if $\dim \mathfrak{g} > 1$ and it contains no nontrivial ideals. By the last exercise, this is equivalent to saying that the adjoint form G of the Lie algebra $\mathfrak{g}$ has no nontrivial normal Lie subgroups.

Now, to attempt to classify Lie algebras, we introduce two descending chains of subalgebras. The first is the *lower central series* of subalgebras $\mathscr{D}_k\mathfrak{g}$, defined inductively by

$$\mathscr{D}_1\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}]$$

and

$$\mathscr{D}_k\mathfrak{g} = [\mathfrak{g}, \mathscr{D}_{k-1}\mathfrak{g}].$$

Note that the subalgebras $\mathscr{D}_k\mathfrak{g}$ are in fact ideals in $\mathfrak{g}$. The other series is called the *derived series* $\{\mathscr{D}^k\mathfrak{g}\}$; it is defined by

$$\mathscr{D}^1\mathfrak{g} = [\mathfrak{g}, \mathfrak{g}]$$

and

$$\mathscr{D}^k\mathfrak{g} = [\mathscr{D}^{k-1}\mathfrak{g}, \mathscr{D}^{k-1}\mathfrak{g}].$$

Exercise 9.3. Use the Jacobi identity to show that $\mathscr{D}^k\mathfrak{g}$ is also an ideal in $\mathfrak{g}$. More generally, if $\mathfrak{h}$ is an ideal in a Lie algebra $\mathfrak{g}$, show that $[\mathfrak{h}, \mathfrak{h}]$ is also an ideal in $\mathfrak{g}$; hence all $\mathscr{D}^k\mathfrak{h}$ are ideals in $\mathfrak{g}$.

Observe that we have $\mathscr{D}^k\mathfrak{g} \subset \mathscr{D}_k\mathfrak{g}$ for all k, with equality when $k = 1$; we often write simply $\mathscr{D}\mathfrak{g}$ for $\mathscr{D}_1\mathfrak{g} = \mathscr{D}^1\mathfrak{g}$ and call this the *commutator subalgebra.* We now make the

Definitions

(i) We say that $\mathfrak{g}$ is *nilpotent* if $\mathscr{D}_k\mathfrak{g} = 0$ for some k.
(ii) We say that $\mathfrak{g}$ is *solvable* if $\mathscr{D}^k\mathfrak{g} = 0$ for some k.

(iii) We say that $\mathfrak{g}$ is *perfect* if $\mathscr{D}\mathfrak{g} = \mathfrak{g}$ (this is not a concept we will use much).
(iv) We say that $\mathfrak{g}$ is *semisimple* if $\mathfrak{g}$ has no nonzero solvable ideals.

The standard example of a nilpotent Lie algebra is the algebra $\mathfrak{n}_n\mathbb{R}$ of strictly upper-triangular $n \times n$ matrices; in this case the kth subalgebra $\mathscr{D}_k\mathfrak{g}$ in the lower central series will be the subspace $\mathfrak{n}_{k+1,n}\mathbb{R}$ of matrices $A = (a_{i,j})$ such that $a_{i,j} = 0$ whenever $j \leq i + k$, i.e., that are zero below the diagonal and within a distance k of it in each column or row. (In terms of a complete flag $\{V_i\}$ as in §7.2, these are just the endomorphisms that carry V_i into V_{i-k-1}.) It follows also that any subalgebra of the Lie algebra $\mathfrak{n}_n\mathbb{R}$ is likewise nilpotent; we will show later that any nilpotent Lie algebra is isomorphic to such a subalgebra. We will also see that if a Lie algebra $\mathfrak{g}$ is represented on a vector space V, such that each element acts as a nilpotent endomorphism, there is a basis for V such that, identifying $\mathfrak{gl}(V)$ with $\mathfrak{gl}_n\mathbb{R}$, $\mathfrak{g}$ maps to the subalgebra $\mathfrak{n}_n\mathbb{R} \subset \mathfrak{gl}_n\mathbb{R}$.

Similarly, a standard example of a solvable Lie algebra is the space $\mathfrak{b}_n\mathbb{R}$ of upper-triangular $n \times n$ matrices; in this Lie algebra the commutator $\mathscr{D}\mathfrak{b}_n\mathbb{R}$ is the algebra $\mathfrak{n}_n\mathbb{R}$ and the derived series is, thus, $\mathscr{D}^k\mathfrak{b}_n\mathbb{R} = \mathfrak{n}_{2^{k-1},n}\mathbb{R}$. Again, it follows that any subalgebra of the algebra $\mathfrak{b}_n\mathbb{R}$ is likewise solvable; and we will prove later that, conversely, *any* representation of a solvable Lie algebra on a vector space V consists, in terms of a suitable basis, entirely of upper-triangular matrices (i.e., given a solvable Lie subalgebra $\mathfrak{g}$ of $\mathfrak{gl}(V)$, there exists a basis for V such that under the corresponding identification of $\mathfrak{gl}(V)$ with $\mathfrak{gl}_n\mathbb{R}$, the subalgebra $\mathfrak{g}$ is contained in $\mathfrak{b}_n\mathbb{R} \subset \mathfrak{gl}_n\mathbb{R}$).

It is clear from the definitions that the properties of being nilpotent or solvable are inherited by subalgebras or homomorphic images. We will see that the same is true for semisimplicity in the case of homomorphic images, though not for subalgebras.

Note that $\mathfrak{g}$ is solvable if and only if $\mathfrak{g}$ has a sequence of Lie *subalgebras* $\mathfrak{g} = \mathfrak{g}_0 \supset \mathfrak{g}_1 \supset \cdots \supset \mathfrak{g}_k = 0$, such that $\mathfrak{g}_{i+1}$ is an ideal in $\mathfrak{g}_i$ and $\mathfrak{g}_i/\mathfrak{g}_{i+1}$ is abelian. Indeed, if this is the case, one sees by induction that $\mathscr{D}^i\mathfrak{g} \subset \mathfrak{g}_i$ for all i. (One may also refine such a sequence to one where each quotient $\mathfrak{g}_i/\mathfrak{g}_{i+1}$ is one dimensional.) It follows from this description that if $\mathfrak{h}$ is an ideal in a Lie algebra $\mathfrak{g}$, then $\mathfrak{g}$ *is solvable if and only if* $\mathfrak{h}$ *and* $\mathfrak{g}/\mathfrak{h}$ *are solvable Lie algebras.* (The analogous assertion for nilpotent Lie algebras is false: the ideal $\mathfrak{n}_n$ is nilpotent in the Lie algebra $\mathfrak{b}_n$ of upper-triangular matrices, and the quotient is the nilpotent algebra $\mathfrak{d}_n$ of diagonal matrices, but $\mathfrak{d}_n$ is not nilpotent.) If $\mathfrak{g}$ is the Lie algebra of a connected Lie group G, then $\mathfrak{g}$ is solvable if and only if there is a sequence of connected subgroups, each normal in G (or in the next in the sequence), such that the quotients are abelian.

In particular, the sum of two solvable ideals in a Lie algebra $\mathfrak{g}$ is again solvable [note that $(\mathfrak{a} + \mathfrak{b})/\mathfrak{b} \cong \mathfrak{a}/(\mathfrak{a} \cap \mathfrak{b})$]. It follows that the sum of all solvable ideals in $\mathfrak{g}$ is a maximal solvable ideal, called the *radical* of $\mathfrak{g}$ and denoted $\mathrm{Rad}(\mathfrak{g})$. The quotient $\mathfrak{g}/\mathrm{Rad}(\mathfrak{g})$ is semisimple. Any Lie algebra $\mathfrak{g}$ thus fits into an exact sequence

$$0 \to \mathrm{Rad}(\mathfrak{g}) \to \mathfrak{g} \to \mathfrak{g}/\mathrm{Rad}(\mathfrak{g}) \to 0 \tag{9.4}$$

where the first algebra is solvable and the last is semisimple. With this somewhat shaky justification (but see Proposition 9.17), we may say that to study the representation theory of an arbitrary Lie algebra, we have to understand individually the representation theories of solvable and semisimple Lie algebras. Of these, the former is relatively easy, at least as regards irreducible representations. The basic fact about them—that any irreducible representation of a solvable Lie algebra is one dimensional—will be proved later in this lecture. The representation theory of semisimple Lie algebras, on the other hand, is extraordinarily rich, and it is this subject that will occupy us for most of the remainder of the book.

Another easy consequence of the definitions is the fact that *a Lie algebra is semisimple if and only if it has no nonzero abelian ideals.* Indeed, the last nonzero term in the derived sequence of ideals $\mathscr{D}^k\mathrm{Rad}(\mathfrak{g})$ would be an abelian ideal in $\mathfrak{g}$ (cf. Exercise 9.3). A semisimple Lie algebra can have no center, so *the adjoint representation of a semisimple Lie algebra is faithful.*

It is a fact that the sequence (9.4) splits, in the sense that there are subalgebras of $\mathfrak{g}$ that map isomorphically onto $\mathfrak{g}/\mathrm{Rad}(\mathfrak{g})$. The existence of such a *Levi decomposition* is part of the general theory we are postponing. To show that an arbitrary Lie algebra has a faithful representation (*Ado's theorem*), one starts with a faithful representation of the center, and then builds a representation of the radical step by step, inserting a string of ideals between the center and the radical. Then one uses a splitting to get from a faithful representation on the radical to some representation on all of $\mathfrak{g}$; the sum of this representation and the adjoint representation is then a faithful representation. See Appendix E for details.

One reason for the terminology simple/semisimple will become clear later in this lecture, when we show that a semisimple Lie algebra is a direct sum of simple ones.

Exercise 9.5. Every semisimple Lie algebra is perfect. Show that the Lie group of Euclidean motions of $\mathbb{R}^3$ has a Lie algebra $\mathfrak{g}$ which is perfect, i.e., $\mathscr{D}\mathfrak{g} = \mathfrak{g}$, but $\mathfrak{g}$ is not semisimple. More generally, if $\mathfrak{h}$ is semisimple, and V is an irreducible representation of $\mathfrak{h}$, the twisted product

$$\mathfrak{g} = \{(v, X) | v \in V, X \in \mathfrak{h}\} \quad \text{with } [(v, X), (w, Y)] = (Xw - Yv, [X, Y])$$

is a Lie algebra with $\mathscr{D}\mathfrak{g} = \mathfrak{g}$, $\mathrm{Rad}(\mathfrak{g}) = V$ abelian, and $\mathfrak{g}/\mathrm{Rad}(\mathfrak{g}) = \mathfrak{h}$.

Exercise 9.6. (a) Show that the following are equivalent for a Lie algebra $\mathfrak{g}$: (i) $\mathfrak{g}$ is nilpotent. (ii) There is a chain of ideals $\mathfrak{g} = \mathfrak{g}_0 \supset \mathfrak{g}_1 \supset \cdots \supset \mathfrak{g}_n = 0$ with $\mathfrak{g}_i/\mathfrak{g}_{i+1}$ contained in the center of $\mathfrak{g}/\mathfrak{g}_{i+1}$. (iii) There is an integer n such that

$$\mathrm{ad}(X_1) \circ \mathrm{ad}(X_2) \circ \cdots \circ \mathrm{ad}(X_n)(Y) = [X_1, [X_2, \ldots, [X_n, Y]\ldots]] = 0$$

for all $X_1, \ldots, X_n, Y$ in $\mathfrak{g}$.

(b) Conclude that a connected Lie group G is nilpotent if and only if it can be realized as a succession of central extensions of abelian Lie groups.

Exercise 9.7*. If G is connected and nilpotent, show that the exponential map $\exp: \mathfrak{g} \to G$ is surjective, making $\mathfrak{g}$ the universal covering space of G.

Exercise 9.8. Show that the following are equivalent for a Lie algebra $\mathfrak{g}$: (i) $\mathfrak{g}$ is solvable. (ii) There is a chain of ideals $\mathfrak{g} = \mathfrak{g}_0 \supset \mathfrak{g}_1 \supset \cdots \supset \mathfrak{g}_n = 0$ with $\mathfrak{g}_i/\mathfrak{g}_{i+1}$ abelian. (iii) There is a chain of subalgebras $\mathfrak{g} = \mathfrak{g}_0 \supset \mathfrak{g}_1 \supset \cdots \supset \mathfrak{g}_n = 0$ such that $\mathfrak{g}_{i+1}$ is an ideal in $\mathfrak{g}_i$, and $\mathfrak{g}_i/\mathfrak{g}_{i+1}$ is abelian.

§9.2. Engel's Theorem and Lie's Theorem

We will now prove the statement made above about representations of solvable Lie algebras always being upper triangular. This may give the reader an idea of how the general theory proceeds, before we go back to the concrete examples that are our main concern. The starting point is

Theorem 9.9 (Engel's Theorem). *Let $\mathfrak{g} \subset \mathfrak{gl}(V)$ be any Lie subalgebra such that every $X \in \mathfrak{g}$ is a nilpotent endomorphism of V. Then there exists a nonzero vector $v \in V$ such that $X(v) = 0$ for all $X \in \mathfrak{g}$.*

Note this implies that there exists a basis for V in terms of which the matrix representative of each $X \in \mathfrak{g}$ is strictly upper triangular: since $\mathfrak{g}$ kills v, it will act on the quotient $\bar{V}$ of V by the span of v, and by induction we can find a basis $\bar{v}_2, \ldots, \bar{v}_n$ for $\bar{V}$ in terms of which this action is strictly upper triangular. Lifting $\bar{v}_i$ to any $v_i \in V$ and setting $v_1 = v$ then gives a basis for V as desired.

Proof of Theorem 9.9. One observation before we start is that if $X \in \mathfrak{gl}(V)$ is any nilpotent element, then the adjoint action $\mathrm{ad}(X): \mathfrak{gl}(V) \to \mathfrak{gl}(V)$ is nilpotent. This is straightforward: to say that X is nilpotent is to say that there exists a flag of subspaces $0 \subset V_1 \subset V_2 \subset \cdots \subset V_k \subset V_{k+1} = V$ such that $X(V_i) \subset V_{i-1}$; we can then check that for any endomorphism Y of V the endomorphism $\mathrm{ad}(X)^m(Y)$ carries V_i into V_{i+k-m}.

We now proceed by induction on the dimension of $\mathfrak{g}$. The first step is to show that, under the hypotheses of the problem, $\mathfrak{g}$ *contains an ideal* $\mathfrak{h}$ *of codimension one*. In fact, let $\mathfrak{h} \subset \mathfrak{g}$ be any maximal proper subalgebra; we claim that $\mathfrak{h}$ has codimension one and is an ideal. To see this, we look at the adjoint representation of $\mathfrak{g}$; since $\mathfrak{h}$ is a subalgebra the adjoint action $\mathrm{ad}(\mathfrak{h})$ of $\mathfrak{h}$ on $\mathfrak{g}$ preserves the subspace $\mathfrak{h} \subset \mathfrak{g}$ and so acts on $\mathfrak{g}/\mathfrak{h}$. Moreover, by our observation above, for any $X \in \mathfrak{h}$ $\mathrm{ad}(X)$ acts nilpotently on $\mathfrak{gl}(V)$, hence on $\mathfrak{g}$, hence on $\mathfrak{g}/\mathfrak{h}$. Thus, by induction, there exists a nonzero element $\bar{Y} \in \mathfrak{g}/\mathfrak{h}$ killed by $\mathrm{ad}(X)$ for all $X \in \mathfrak{h}$; equivalently, there exists an element $Y \in \mathfrak{g}$ not in $\mathfrak{h}$ such that $\mathrm{ad}(X)(Y) \in \mathfrak{h}$ for all $X \in \mathfrak{h}$. But this is to say that the subspace $\mathfrak{h}'$ of $\mathfrak{g}$ spanned by $\mathfrak{h}$ and Y is a Lie subalgebra of $\mathfrak{g}$, in which $\mathfrak{h}$ sits as an ideal of codimension one; by the maximality of $\mathfrak{h}$ we have $\mathfrak{h}' = \mathfrak{g}$ and we are done.

We return now to the representation of $\mathfrak{g}$ on V. We may apply the induction hypothesis to the subalgebra $\mathfrak{h}$ of $\mathfrak{g}$ found in the preceding paragraph to conclude that there exists a nonzero vector $v \in V$ such that $X(v) = 0$ for all $X \in \mathfrak{h}$; let $W \subset V$ be the subspace of all such vectors $v \in V$. Let Y be any element of $\mathfrak{g}$ not in $\mathfrak{h}$; since $\mathfrak{h}$ and Y span $\mathfrak{g}$, it will suffice to show that there exists a (nonzero) vector $v \in W$ such that $Y(v) = 0$. Now for any vector $w \in W$ and any $X \in \mathfrak{h}$, we have

$$X(Y(w)) = Y(X(w)) + [X, Y](w).$$

The first term on the right is zero because by hypothesis $w \in W$, $X \in \mathfrak{h}$ and so $X(w) = 0$; likewise, the second term is zero because $[X, Y] = \mathrm{ad}(X)(Y) \in \mathfrak{h}$. Thus, $X(Y(w)) = 0$ for all $X \in \mathfrak{h}$; we deduce that $Y(w) \in W$. But this means that the action of Y on V carries the subspace W into itself; since Y acts nilpotently on V, it follows that there exists a vector $v \in W$ such that $Y(v) = 0$. □

Exercise 9.10*. Show that a Lie algebra $\mathfrak{g}$ is nilpotent if and only if $\mathrm{ad}(X)$ is a nilpotent endomorphism of $\mathfrak{g}$ for every $X \in \mathfrak{g}$.

Engel's theorem, in turn, allows us to prove the basic statement made above that every representation of a solvable Lie group can be put in upper-triangular form. This is implied by

Theorem 9.11 (Lie's Theorem). *Let $\mathfrak{g} \subset \mathfrak{gl}(V)$ be a complex solvable Lie algebra. Then there exists a nonzero vector $v \in V$ that is an eigenvector of X for all $X \in \mathfrak{g}$.*

Exercise 9.12. Show that this implies the existence of a basis for V in terms of which the matrix representative of each $X \in \mathfrak{g}$ is upper triangular.

Proof of Theorem 9.11. Once more, the first step in the argument is to assert that $\mathfrak{g}$ contains an ideal $\mathfrak{h}$ of codimension one. This time, since $\mathfrak{g}$ is solvable we know that $\mathscr{D}\mathfrak{g} \neq \mathfrak{g}$, so that the quotient $\mathfrak{a} = \mathfrak{g}/\mathscr{D}\mathfrak{g}$ is a nonzero abelian Lie algebra; the inverse image in $\mathfrak{g}$ of any codimension one subspace of $\mathfrak{a}$ will then be a codimension one ideal in $\mathfrak{g}$.

Still following the lines of the previous argument, we may by induction assume that there is a vector $v_0 \in V$ that is an eigenvector for all $X \in \mathfrak{h}$. Denote the eigenvalue of X corresponding to v_0 by $\lambda(X)$. We then consider the subspace $W \subset V$ of all vectors satisfying the same relation, i.e., we set

$$W = \{v \in V: X(v) = \lambda(X) \cdot v \ \forall X \in \mathfrak{h}\}.$$

Let Y now be any element of $\mathfrak{g}$ not in $\mathfrak{h}$. As before, it will suffice to show that Y carries some vector $v \in W$ into a multiple of itself, and for this it is enough to show that Y carries W into itself. We prove this in a general context in the following lemma.

Lemma 9.13. *Let $\mathfrak{h}$ be an ideal in a Lie algebra $\mathfrak{g}$. Let V be a representation of $\mathfrak{g}$, and $\lambda: \mathfrak{h} \to \mathbb{C}$ a linear function. Set*

$$W = \{v \in V: X(v) = \lambda(X)\cdot v \ \forall X \in \mathfrak{h}\}.$$

Then $Y(W) \subset W$ for all $Y \in \mathfrak{g}$.

PROOF. Let w be any nonzero element of W; to test whether $Y(w) \in W$ we let X be any element of $\mathfrak{h}$ and write

$$\begin{aligned} X(Y(w)) &= Y(X(w)) + [X, Y](w) \\ &= \lambda(X)\cdot Y(w) + \lambda([X, Y])\cdot w \end{aligned} \tag{9.14}$$

since $[X, Y] \in \mathfrak{h}$. This differs from our previous calculation in that the second term on the right is not immediately seen to be zero; indeed, $Y(w)$ will lie in W if and only if $\lambda([X, Y]) = 0$ for all $X \in \mathfrak{h}$.

To verify this, we introduce another subspace of V, namely, the span U of the images w, $Y(w)$, $Y^2(w)$, ... of w under successive applications of Y. This subspace is clearly preserved by Y; we claim that any $X \in \mathfrak{h}$ carries U into itself as well. It is certainly the case that $\mathfrak{h}$ carries w into a multiple of itself, and hence into U, and (9.14) says that $\mathfrak{h}$ carries $Y(w)$ into a linear combination of $Y(w)$ and w, and so into U. In general, we can see that $\mathfrak{h}$ carries $Y^k(w)$ into U by induction: for any $X \in \mathfrak{h}$ we write

$$X(Y^k(w)) = Y(X(Y^{k-1}(w))) + [X, Y](Y^{k-1}(w)). \tag{9.15}$$

Since $X(Y^{k-1}(w)) \in U$ by induction the first term on the right is in U, and since $[X, Y] \in \mathfrak{h}$ the second term is in U as well.

In fact, we see something more from (9.14) and (9.15): it follows that, in terms of the basis w, $Y(w)$, $Y^2(w)$, ... for U, the action of any $X \in \mathfrak{h}$ is upper triangular, with diagonal entries all equal to $\lambda(X)$. In particular, for any $X \in \mathfrak{h}$ the trace of the restriction of X to U is just the dimension of U times $\lambda(X)$. On the other hand, for any element $X \in \mathfrak{h}$ the commutator $[X, Y]$ acts on U, and being the commutator of two endomorphisms of U the trace of this action is zero. It follows then that $\lambda([X, Y]) = 0$, and we are done. □

Exercise 9.16. Show that any irreducible representation of a solvable Lie algebra $\mathfrak{g}$ is one dimensional, and $\mathscr{D}\mathfrak{g}$ acts trivially.

At least for *irreducible* representations, Lie's theorem implies they will all be known for an arbitrary Lie algebra when they are known for the semisimple case. In fact, we have:

Proposition 9.17. *Let $\mathfrak{g}$ be a complex Lie algebra, $\mathfrak{g}_{ss} = \mathfrak{g}/\mathrm{Rad}(\mathfrak{g})$. Every irreducible representation of $\mathfrak{g}$ is of the form $V = V_0 \otimes L$, where V_0 is an irreducible* *representation of* $\mathfrak{g}_{ss}$ [*i.e., a representation of* $\mathfrak{g}$ *that is trivial on* $\mathrm{Rad}(\mathfrak{g})$], *and* L *is a one-dimensional representation.*

PROOF. By Lie's theorem there is a $\lambda \in (\mathrm{Rad}(\mathfrak{g}))^*$ such that

$$W = \{v \in V \colon X(v) = \lambda(X) \cdot v \ \forall X \in \mathrm{Rad}(\mathfrak{g})\}$$

is not zero. Apply the preceding lemma, with $\mathfrak{h} = \mathrm{Rad}(\mathfrak{g})$. Since V is irreducible, we must have $W = V$. In particular, $\mathrm{Tr}(X) = \dim(V) \cdot \lambda(X)$ for $X \in \mathrm{Rad}(\mathfrak{g})$, so λ vanishes on $\mathrm{Rad}(\mathfrak{g}) \cap [\mathfrak{g}, \mathfrak{g}]$. Extend λ to a linear function on $\mathfrak{g}$ that vanishes on $[\mathfrak{g}, \mathfrak{g}]$, and let L be the one-dimensional representation of $\mathfrak{g}$ determined by λ; in other words, $Y(z) = \lambda(Y) \cdot z$ for all $Y \in \mathfrak{g}$ and $z \in L$. Then $V \otimes L^*$ is a representation that is trivial on $\mathrm{Rad}(\mathfrak{g})$, so it comes from a representation of $\mathfrak{g}_{ss}$, as required. □

Exercise 9.18. Show that if $\mathfrak{g}'$ is a subalgebra of $\mathfrak{g}$ that maps isomorphically onto $\mathfrak{g}/\mathrm{Rad}(\mathfrak{g})$, then any irreducible representation of $\mathfrak{g}$ restricts to an irreducible representation of $\mathfrak{g}'$, and any irreducible representation of $\mathfrak{g}'$ extends to a representation of $\mathfrak{g}$.

§9.3. Semisimple Lie Algebras

As is clear from the above, many of the aspects of the representation theory of finite groups that were essential to our approach are no longer valid in the context of general Lie algebras and Lie groups. Most obvious of these is complete reducibility, which we have seen fails for Lie groups; another is the fact that not only can the action of elements of a Lie group or algebra on a vector space be nondiagonalizable, the action of some element of a Lie algebra may be diagonalizable under one representation and not under another.

That is the bad news. The good news is that, if we just restrict ourselves to semisimple Lie algebras, everything is once more as well behaved as possible. For one thing, we have complete reducibility again:

Theorem 9.19 (Complete Reducibility). *Let* V *be a representation of the semisimple Lie algebra* $\mathfrak{g}$ *and* $W \subset V$ *a subspace invariant under the action of* $\mathfrak{g}$. *Then there exists a subspace* $W' \subset V$ *complementary to* W *and invariant under* $\mathfrak{g}$.

The proof of this basic result will be deferred to Appendix C.

The other question, the diagonalizability of elements of a Lie algebra under a representation, requires a little more discussion. Recall first the statement of *Jordan decomposition*: any endomorphism X of a complex vector space V can be uniquely written in the form

$$X = X_s + X_n$$

where X_s is diagonalizable, X_n is nilpotent, and the two commute. Moreover, X_s and X_n may be expressed as polynomials in X.

Now, suppose that $\mathfrak{g}$ is an arbitrary Lie algebra, $X \in \mathfrak{g}$ any element, and $\rho: \mathfrak{g} \to \mathfrak{gl}_n\mathbb{C}$ any representation. We have seen that the image $\rho(X)$ need not be diagonalizable; we may still ask how $\rho(X)$ behaves with respect to the Jordan decomposition. The answer is that, in general, absolutely nothing need be true. For example, just taking $\mathfrak{g} = \mathbb{C}$, we see that under the representation

$$\rho_1: t \mapsto (t)$$

every element is diagonalizable, i.e., $\rho(X)_s = \rho(X)$; under the representation

$$\rho_2: t \mapsto \begin{pmatrix} 0 & t \\ 0 & 0 \end{pmatrix}$$

every element is nilpotent [i.e., $\rho(X)_s = 0$]; whereas under the representation

$$\rho_3: t \mapsto \begin{pmatrix} t & t \\ 0 & 0 \end{pmatrix}$$

not only are the images $\rho(X)$ neither diagonalizable nor nilpotent, the diagonalizable and nilpotent parts of $\rho(X)$ are not even in the image $\rho(\mathfrak{g})$ of the representation.

If we assume the Lie algebra $\mathfrak{g}$ is semisimple, however, the situation is radically different. Specifically, we have

Theorem 9.20 (Preservation of Jordan Decomposition). *Let $\mathfrak{g}$ be a semisimple Lie algebra. For any element $X \in \mathfrak{g}$, there exist X_s and $X_n \in \mathfrak{g}$ such that for any representation $\rho: \mathfrak{g} \to \mathfrak{gl}(V)$ we have*

$$\rho(X)_s = \rho(X_s) \quad \textit{and} \quad \rho(X)_n = \rho(X_n).$$

In other words, if we think of ρ as injective and $\mathfrak{g}$ accordingly as a Lie subalgebra of $\mathfrak{gl}(V)$, *the diagonalizable and nilpotent parts of any element X of $\mathfrak{g}$ are again in $\mathfrak{g}$ and are independent of the particular representation ρ.*

The proofs we will give of the last two theorems both involve introducing objects that are not essential for the rest of this book, and we therefore relegate them to Appendix C. It is worth remarking, however, that another approach was used classically by Hermann Weyl; this is the famous *unitary trick*, which we will describe briefly.

A Digression on "The Unitary Trick"

Basically, the idea is that the statements above (complete reducibility, preservation of Jordan decomposition) can be proved readily for the representations of a compact Lie group. To prove complete reducibility, for example, we can proceed more or less just as in the case of a finite group: if the compact group G acts on a vector space, we see that there is a Hermitian metric on V invariant under the action of G by taking an arbitrary metric on V and averaging its images under the action of G. If G fixes a subspace $W \subset V$, it will then fix as well its orthogonal complement $W^\perp$ with respect to this metric. (Alternatively, we can choose an arbitrary complement W' to W, not necessarily fixed by G, and average over G the projection map to $g(W')$ with kernel W; this average will have image invariant under G.)

How does this help us analyze the representation of a semisimple Lie algebra? The key fact here (to be proved in Lecture 26) is that *if* $\mathfrak{g}$ *is any complex semisimple Lie algebra, there exists a (unique) real Lie algebra* $\mathfrak{g}_0$ *with complexification* $\mathfrak{g}_0 \otimes \mathbb{C} = \mathfrak{g}$, *such that the simply connected form of the Lie algebra* $\mathfrak{g}_0$ *is a compact Lie group* G. Thus, restricting a given representation of $\mathfrak{g}$ to $\mathfrak{g}_0$, we can exponentiate to obtain a representation of G, for which complete reducibility holds; and we can deduce from this the complete reducibility of the original representation. For example, while it is certainly not true that any representation ρ of the Lie group $SL_n\mathbb{R}$ on a vector space V admits an invariant Hermitian metric (in fact, it cannot, unless it is the trivial representation), we can

(i) let ρ' be the corresponding (complex) representation of the Lie algebra $\mathfrak{sl}_n\mathbb{R}$;
(ii) by linearity extend the representation ρ' of $\mathfrak{sl}_n\mathbb{R}$ to a representation ρ'' of $\mathfrak{sl}_n\mathbb{C}$;
(iii) restrict to a representation ρ''' of the subalgebra $\mathfrak{su}_n \subset \mathfrak{sl}_n\mathbb{C}$;
(iv) exponentiate to obtain a representation ρ'''' of the unitary group SU_n.

We can now argue that

If a subspace $W \subset V$ is invariant under the action of $SL_n\mathbb{R}$,

it must be invariant under $\mathfrak{sl}_n\mathbb{R}$; and since $\mathfrak{sl}_n\mathbb{C} = \mathfrak{sl}_n\mathbb{R} \otimes \mathbb{C}$, it follows that
it will be invariant under $\mathfrak{sl}_n\mathbb{C}$; so of course
it will be invariant under $\mathfrak{su}_n$; and hence
it will be invariant under SU_n.

Now, since SU_n is compact, there will exist a complementary subspace W' preserved by SU_n; we argue that

W' will then be invariant under $\mathfrak{su}_n$; and since $\mathfrak{sl}_n\mathbb{C} = \mathfrak{su}_n \otimes \mathbb{C}$, it follows that
it will be invariant under $\mathfrak{sl}_n\mathbb{C}$. Restricting, we see that
it will be invariant under $\mathfrak{sl}_n\mathbb{R}$, and exponentiating,
it will be invariant under $SL_n\mathbb{R}$.

Similarly, if one wants to know that the diagonal elements of $SL_n\mathbb{R}$ act semisimply in any representation, or equivalently that the diagonal elements of $\mathfrak{sl}_n\mathbb{R}$ act semisimply, one goes through the same reasoning, coming down to the fact that the group of diagonal elements in $\mathfrak{su}_n$ is abelian and *compact*.

In general, most of the theorems about the finite-dimensional representation of semisimple Lie algebras admit proofs along two different lines: either algebraically, using just the structure of the Lie algebra; or by the unitary trick, that is, by associating to a representation of such a Lie algebra a representation of a compact Lie group and working with that. Which is preferable depends very much on taste and context; in this book we will for the most part go with the algebraic proofs, though in the case of the Weyl character formula in Part IV the proof via compact groups is so much more appealing it has to be mentioned.

The following exercises include a few applications of these two theorems.

Exercise 9.21*. Show that a Lie algebra $\mathfrak{g}$ is semisimple if and only if every finite-dimensional representation is semisimple, i.e., every invariant subspace has a complement.

Exercise 9.22. Use Weyl's unitary trick to show that, for $n > 2$, all representations of $\mathrm{SO}_n\mathbb{C}$ are semisimple, so that, in particular, the Lie algebras $\mathfrak{so}_n\mathbb{C}$ are semisimple. Do the same for $\mathrm{Sp}_{2s}\mathbb{C}$ and $\mathfrak{sp}_{2n}\mathbb{C}$, $n \geq 1$. Where does the argument break down for $\mathrm{SO}_2\mathbb{C}$?

Exercise 9.23. Show that a real Lie algebra $\mathfrak{g}$ is solvable if and only if the complex Lie algebra $\mathfrak{g} \otimes_{\mathbb{R}} \mathbb{C}$ is solvable. Similarly for nilpotent and semisimple.

Exercise 9.24*. If $\mathfrak{h}$ is an ideal in a Lie algebra $\mathfrak{g}$, show that $\mathfrak{g}$ is semisimple if and only if $\mathfrak{h}$ and $\mathfrak{g}/\mathfrak{h}$ are semisimple. Deduce that every semisimple Lie algebra is a direct sum of simple Lie algebras.

Exercise 9.25*. A Lie algebra is called *reductive* if its radical is equal to its center. A Lie group is reductive if its Lie algebra is reductive. For example, $\mathrm{GL}_n\mathbb{C}$ is reductive. Show that the following are true for a reductive Lie algebra $\mathfrak{g}$: (i) $\mathscr{D}\mathfrak{g}$ is semisimple; (ii) the adjoint representation of $\mathfrak{g}$ is semisimple; (iii) $\mathfrak{g}$ is a product of a semisimple and an abelian Lie algebra; (iv) $\mathfrak{g}$ has a finite-dimensional faithful semisimple representation. In fact, each of these conditions is equivalent to $\mathfrak{g}$ being reductive.

§9.4. Simple Lie Algebras

There is one more basic fact about Lie algebras to be stated here; though its proof will have to be considerably deferred, it informs our whole approach to the subject. This is the complete classification of simple Lie algebras:

Theorem 9.26. *With five exceptions, every simple complex Lie algebra is isomorphic to either $\mathfrak{sl}_n\mathbb{C}$, $\mathfrak{so}_n\mathbb{C}$, or $\mathfrak{sp}_{2n}\mathbb{C}$ for some n.*

The five exceptions can all be explicitly described, though none is particularly simple except in name; they are denoted $\mathfrak{g}_2$, $\mathfrak{f}_4$, $\mathfrak{e}_6$, $\mathfrak{e}_7$, and $\mathfrak{e}_8$. We will give a construction of each later in the book (§22.3). The algebras $\mathfrak{sl}_n\mathbb{C}$ (for $n > 1$), $\mathfrak{so}_n\mathbb{C}$ (for $n > 2$), and $\mathfrak{sp}_{2n}\mathbb{C}$ are commonly called the *classical Lie algebras* (and the corresponding groups the *classical Lie groups*); the other five algebras are called, naturally enough, the *exceptional Lie algebras.*

The nature of the classification theorem for simple Lie algebras creates a dilemma as to how we approach the subject: many of the theorems about simple Lie algebras can be proved either in the abstract, or by verifying them in turn for each of the particular algebras listed in the classification theorem. Another alternative is to declare that we are concerned with understanding only the representations of the classical algebras $\mathfrak{sl}_n\mathbb{C}$, $\mathfrak{so}_n\mathbb{C}$, and $\mathfrak{sp}_{2n}\mathbb{C}$, and verify any relevant theorems just in these cases.

Of these three approaches, the last is in many ways the least satisfactory; it is, however, the one that we shall for the most part take. Specifically, what we will do, starting in Lecture 11, is the following:

(i) Analyze in Lectures 11–13 a couple of examples, namely, $\mathfrak{sl}_2\mathbb{C}$ and $\mathfrak{sl}_3\mathbb{C}$, on what may appear to be an ad hoc basis.
(ii) On the basis of these examples, propose in Lecture 14 a general paradigm for the study of representations of a simple (or semisimple) Lie algebra.
(iii) Proceed in Lectures 15–20 to carry out this analysis for the classical algebras $\mathfrak{sl}_n\mathbb{C}$, $\mathfrak{so}_n\mathbb{C}$, and $\mathfrak{sp}_{2n}\mathbb{C}$.
(iv) Give in Part IV and the appendices proofs for general simple Lie algebras of the facts discovered in the preceding sections for the classical ones (as well as one further important result, the *Weyl character formula*).

We can at least partially justify this seemingly inefficient approach by saying that even if one makes a beeline for the general theorems about the structure and representation theory of a simple Lie algebra, to apply these results in practice we would still need to carry out the sort of explicit analysis of the individual algebras done in Lectures 11–20. This is, however, a fairly bald rationalization: the fact is, the reason we are doing it this way is that this is the only way we have ever been able to understand any of the general results.

LECTURE 10

Lie Algebras in Dimensions One, Two, and Three

Just to get a sense of what a Lie algebra is and what groups might be associated to it, we will classify here all Lie algebras of dimension three or less. We will work primarily with complex Lie algebras and Lie groups, but will mention the real case as well. Needless to say, this lecture is logically superfluous; but it is easy, fun, and serves a didactic purpose, so why not read it anyway. The analyses of both the Lie algebras and the Lie groups are completely elementary, with one exception: the classification of the complex Lie groups associated to abelian Lie algebras involves the theory of complex tori, and should probably be skipped by anyone not familiar with this subject.

§10.1. Dimensions One and Two

To begin with, any one-dimensional Lie algebra $\mathfrak{g}$ is clearly abelian, that is, $\mathbb{C}$ with all brackets zero.

The simply connected Lie group with this Lie algebra is just the group $\mathbb{C}$ under addition; and other connected Lie groups that have $\mathfrak{g}$ as their Lie algebra must all be quotients of $\mathbb{C}$ by discrete subgroups $\Lambda \subset \mathbb{C}$. If Λ has rank one, then the quotient is just $\mathbb{C}^*$ under multiplication. If Λ has rank two, however, G may be any one of a continuously varying family of *complex tori of dimension one* (or *Riemann surfaces of genus one*, or *elliptic curves over* $\mathbb{C}$). The set of isomorphism classes of such tori is parametrized by the complex plane with coordinate j, where the function j on the set of lattices $\Lambda \subset \mathbb{C}$ is as described in, e.g., [Ahl].

Over the real numbers, the situation is completely straightforward: the only real Lie algebra of dimension one is again $\mathbb{R}$ with trivial bracket; the simply connected Lie group associated to it is $\mathbb{R}$ under addition; and the only other connected real Lie group with this Lie algebra is $\mathbb{R}/\mathbb{Z} \cong S^1$.

Dimension Two

Here we have to consider two cases, depending on whether $\mathfrak{g}$ is abelian or not.

Case 1: $\mathfrak{g}$ *abelian.* This is very much like the previous case; the simply connected two-dimensional abelian complex Lie group is just $\mathbb{C}^2$ under addition, and the remaining connected Lie groups with Lie algebra $\mathfrak{g}$ are just quotients of $\mathbb{C}^2$ by discrete subgroups. Such a subgroup $\Lambda \subset \mathbb{C}^2$ can have rank 1, 2, 3, or 4, and we analyze these possibilities in turn (the reader who has seen enough complex tori in the preceding example may wish to skip directly to Case 2 at this point).

If the rank of Λ is 1, we can complete the generator of Λ to a basis for $\mathbb{C}^2$, so that $\Lambda = \mathbb{Z}e_1 \subset \mathbb{C}e_1 \oplus \mathbb{C}e_2$ and $G \cong \mathbb{C}^* \times \mathbb{C}$. If the rank of Λ is 2, there are two possibilities: either Λ lies in a one-dimensional complex subspace of $\mathbb{C}^2$ or it does not. If it does not, a pair of generators for Λ will also be a basis for $\mathbb{C}^2$ over $\mathbb{C}$, so that $\Lambda = \mathbb{Z}e_1 \oplus \mathbb{Z}e_2$, $\mathbb{C}^2 = \mathbb{C}e_1 \oplus \mathbb{C}e_2$, and $G \cong \mathbb{C}^* \times \mathbb{C}^*$. If on the other hand Λ does lie in a complex line in $\mathbb{C}^2$, so that we have $\Lambda = \mathbb{Z}e_1 \oplus \mathbb{Z}\tau e_1$ for some $\tau \in \mathbb{C} \setminus \mathbb{R}$, then $G = E \times \mathbb{C}$ will be the product of the torus $\mathbb{C}/(\mathbb{Z} \oplus \mathbb{Z}\tau)$ and $\mathbb{C}$; the remarks above apply to the classification of these (see Exercise 10.1).

The cases where Λ has rank 3 or 4 are a little less clear. To begin with, if the rank of Λ is 3, the main question to ask is whether any rank 2 sublattice Λ' of Λ lies in a complex line. If it does, then we can assume this sublattice is saturated (i.e., a pair of generators for Λ' can be completed to a set of generators for Λ) and write $\Lambda = \mathbb{Z}e_1 \oplus \mathbb{Z}\tau e_1 \oplus \mathbb{Z}e_2$, so that we will have $G = E \times C^*$, where E is a torus as above.

Exercise 10.1*. For two one-dimensional complex tori E and E', show that the complex Lie groups $G = E \times \mathbb{C}$ and $G' = E' \times \mathbb{C}$ are isomorphic if and only if $E \cong E'$. Similarly for $E \times \mathbb{C}^*$ and $E' \times \mathbb{C}^*$.

If, on the other hand, no such sublattice of Λ exists, the situation is much more mysterious. One way we can try to represent G is by choosing a generator for Λ and considering the projection of $\mathbb{C}^2$ onto the quotient of $\mathbb{C}^2$ by the line spanned by this generator; thus, if we write $\Lambda = \mathbb{Z}e_1 \oplus \mathbb{Z}e_2 \oplus \mathbb{Z}(\alpha e_1 + \beta e_2)$ then (assuming β is not real) we have maps

$$\begin{tikzcd} \mathbb{C}^2 \arrow[r] \arrow[d] & \mathbb{C}^2/\mathbb{C}e_1 = \mathbb{C} \arrow[d] \\ G = \mathbb{C}^2/\mathbb{Z}e_1 \oplus \mathbb{Z}e_2 \oplus \mathbb{Z}(\alpha e_1 + \beta e_2) \arrow[r] & \mathbb{C}/(\mathbb{Z} \oplus \mathbb{Z}\beta) \end{tikzcd}$$

expressing G as a bundle over a torus $E = \mathbb{C}/(\mathbb{Z} \oplus \mathbb{Z}\beta)$, with fibers isomorphic to $\mathbb{C}^*$. This expression of G does not, however, help us very much to describe the family of all such groups. For one thing, the elliptic curve E is surely not determined by the data of G: if we just exchange e_1 and e_2, for example, we replace E by $\mathbb{C}/(\mathbb{Z} \oplus \mathbb{Z}\alpha)$, which, of course, need not even be isogenous to E. Indeed, this yields an example of different algebraic groups isomorphic as complex Lie groups: expressing G as a $\mathbb{C}^*$ bundle in this way gives it the structure of an algebraic variety, which, in turn, determines the elliptic curve E (for example, the field of rational functions on G will be the field of rational functions on E with one variable adjoined). Thus, different expressions of the complex Lie group G as a $\mathbb{C}^*$ bundle yield nonisomorphic algebraic groups.

Finally, the case where Λ has rank 4 remains completely mysterious. Among such two-dimensional complex tori are the *abelian varieties*; these are just the tori that may be embedded in complex projective space (and hence may be realized as algebraic varieties). For polarized abelian varieties (that is, abelian varieties with equivalence class of embedding in projective space) there exists a reasonable moduli theory; but the set of abelian varieties forms only a countable dense union in the set of all complex tori (indeed, the general complex torus possesses no nonconstant meromorphic functions whatsoever). No satisfactory theory of moduli is known for these objects.

Needless to say, the foregoing discussion of the various abelian complex Lie groups in dimension two is completely orthogonal to our present purposes. We hope to make the point, however, that even in this seemingly trivial case there lurk some fairly mysterious phenomena. Of course, none of this occurs in the real case, where the two-dimensional abelian simply connected real Lie group is just $\mathbb{R} \times \mathbb{R}$ and any other connected two-dimensional abelian real Lie group is the quotient of this by a sublattice $\Lambda \subset \mathbb{R} \times \mathbb{R}$ of rank 1 or 2, which is to say either $\mathbb{R} \times S^1$ or $S^1 \times S^1$.

Case 2: $\mathfrak{g}$ *not abelian*. Viewing the Lie bracket as a linear map $[\ ,\ ]: \wedge^2\mathfrak{g} \to \mathfrak{g}$, we see that if it is not zero, it must have one-dimensional image. We can thus choose a basis $\{X, Y\}$ for $\mathfrak{g}$ as vector space with X spanning the image of $[\ ,\ ]$; after multiplying Y by an appropriate scalar we will have $[X, Y] = X$, which of course determines $\mathfrak{g}$ completely. There is thus a unique nonabelian two-dimensional Lie algebra $\mathfrak{g}$ over either $\mathbb{R}$ or $\mathbb{C}$.

What are the complex Lie groups with Lie algebra $\mathfrak{g}$? To find one, we start with the adjoint representation of $\mathfrak{g}$, which is faithful: we have

$$\begin{aligned} \operatorname{ad}(X)\colon X &\mapsto 0, & \qquad \operatorname{ad}(Y)\colon X &\mapsto -X, \\ Y &\mapsto X, & Y &\mapsto 0 \end{aligned}$$

or in matrix notation, in terms of the basis $\{X, Y\}$ for $\mathfrak{g}$,

$$\operatorname{ad}(X) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \qquad \operatorname{ad}(Y) = \begin{pmatrix} -1 & 0 \\ 0 & 0 \end{pmatrix}.$$

These generate the algebra $\mathfrak{g} = \begin{pmatrix} * & * \\ 0 & 0 \end{pmatrix} \subset \mathfrak{gl}_2\mathbb{C}$; we may exponentiate to arrive at the adjoint form

$$G_0 = \left\{ \begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} : a \neq 0 \right\} \subset \mathrm{GL}_2\mathbb{C}.$$

Topologically this group is homeomorphic to $\mathbb{C} \times \mathbb{C}^*$. To take its universal cover, we write a general member of G_0 as

$$\begin{pmatrix} e^t & s \\ 0 & 1 \end{pmatrix}.$$

The product of two such matrices is given by

$$\begin{pmatrix} e^t & s \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} e^{t'} & s' \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} e^{t+t'} & s + e^t s' \\ 0 & 1 \end{pmatrix},$$

so we may realize the universal cover G of G_0 as the group of pairs $(t, s) \in \mathbb{C} \times \mathbb{C}$ with group law

$$(t, s) \cdot (t', s') = (t + t', s + e^t s').$$

The center of G is just the subgroup

$$Z(G) = \{(2\pi i n, 0)\} \cong \mathbb{Z},$$

so that the connected groups with Lie algebra $\mathfrak{g}$ form a partially ordered tower

$$\begin{array}{l} G \\ \downarrow \\ \vdots \\ \downarrow \\ G_n = G/n\mathbb{Z} = \{(a, b) \in \mathbb{C}^* \times \mathbb{C}; (a, b) \cdot (a', b') = (aa', b + a^n b')\}. \\ \downarrow \\ \vdots \\ \downarrow \\ G_0 \end{array}$$

Exercise 10.2*. Show that for $n \neq m$ the two groups G_n and G_m are not isomorphic.

Finally, in the real case things are simpler: when we exponentiate the adjoint representation as above, the Lie group we arrive at is already simply connected, and so is the unique connected real Lie group with this Lie algebra.

§10.2. Dimension Three, Rank 1

As in the case of dimension two, we look at the Lie bracket as a linear map from $\wedge^2 \mathfrak{g}$ to $\mathfrak{g}$ and begin our classification by considering the rank of this map (that is, the dimension of $\mathscr{D}\mathfrak{g}$), which may be either 0, 1, 2, or 3. For the case