Willoughby Seago

Notes from

Representation Theory

January 13th, 2024

University of Glasgow

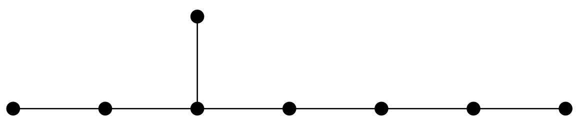
Representation Theory

Willoughby Seago

January 13th, 2024

These are my notes from the SMSTC course *Lie Theory* taught by Prof Christian Korff. These notes were last updated at 14:07 on August 6, 2025.

E_8



Chapters

Page

Contents

Page

One

Introduction

We fix some standard notation here:

- $\Bbbk$ will denote an algebraically closed field, except for when we explicitly mention that the field needn't be algebraically closed.
- A will denote an associative unital algebra.
- Letters like V, U, and W will denote vector spaces over $\Bbbk$.
- Letters like M and N will denote modules.

Part I

Algebra Representations

Two

Initial Definitions

2.1 Algebra

Definition 2.1.1 — Algebra An **algebra** is a $\Bbbk$-vector space, A, equipped with a bilinear map,

$$m : A \times A \to A \tag{2.1.2}$$
$$(a, b) \mapsto m(a, b) = ab. \tag{2.1.3}$$

If this map satisfies the condition that

$$m(a, m(b, c)) = m(m(a, b), c), \text{ or equivalently } a(bc) = (ab)c, \tag{2.1.4}$$

for all $a, b, c \in A$ then we call A an **associative algebra**.
If A posses a distinguished element, $1 \in A$, such that $m(1, a) = a = m(a, 1)$, or equivalently $1a = a = a1$ for all $a \in A$ then we say that A is a **unital algebra**.
If $m(a, b) = m(b, a)$, or equivalently $ab = ba$, for all $a, b \in A$ then we say that A is a **commutative algebra**.

Whenever we say, otherwise unqualified, "algebra" we will mean associative unital algebra unless we specify otherwise. We will not assume commutativity of a general algebra.

The condition of associativity can be written as a commutative diagram,

$$\begin{array}{ccc} A \times A \times A & \xrightarrow{m \times \mathrm{id}_A} & A \times A \\ {\scriptstyle \mathrm{id}_A \times m}\big\downarrow & & \big\downarrow{\scriptstyle m} \\ A \times A & \xrightarrow[m]{} & A, \end{array} \tag{2.1.5}$$

Remark 2.1.6 This diagram goes part of the way to the more abstract definition that "an associative unital (commutative) algebra is a (commutative) monoid in the category of vector spaces". This definition is nice because it is both very general and dualises to the notion of a coalgebra. See the *Hopf Algebra* notes for more details.

Example 2.1.7

- $A = \Bbbk$ is an algebra with the product given by the product in the field;
- $A = \Bbbk[x_1, \dots, x_n]$, the ring of polynomials in the variables x_i with coefficients in $\Bbbk$, is an algebra under the addition and multiplication of polynomials.
- $A = \Bbbk\langle x_1, \dots, x_n\rangle$, the free algebra on x_i, may be considered as the algebra of polynomials in non-commuting variables, x_i.
- $A = \operatorname{End} V$ for V a $\Bbbk$-vector space is an algebra with multiplication given by composition of morphisms.

Definition 2.1.8 — Group Algebra Let G be a group. The **group algebra** or **group ring** $\Bbbk G = \Bbbk[G]$ is defined to be the set of finite formal linear combinations

$$\sum_{g\in G} c_g g \tag{2.1.9}$$

where $c_g \in \Bbbk$ is nonzero for only finitely many values g. Addition is defined by

$$\sum_{g\in G} c_g g + \sum_{g\in G} d_g g = \sum_{g\in G} (c_g + d_g)g. \tag{2.1.10}$$

Multiplication is defined by requiring that it distributes over addition and that the product of two terms in the above sums is given by

$$(c_g g)(d_h h) = (c_g d_h)(gh) \tag{2.1.11}$$

where multiplication on the left is in $\Bbbk G$, the multiplication $c_g d_h$ is in $\Bbbk$, and the multiplication gh is in G.
If we do the same construction replacing $\Bbbk$ with a ring, R, then we get the group ring, RG, which is not an algebra but instead an R-module.

Definition 2.1.12 — Algebra Homomorphism Let A and B be $\Bbbk$-algebras. An **algebra homomorphism** is a linear map $f : A \to B$ such that $f(ab) = f(a)f(b)$ for all $a, b \in A$.
If A and B are unital, with units 1_A and 1_B respectively, then we further require that $f(1_A) = 1_B$.
We denote by $\operatorname{Hom}(A, B)$ or $\operatorname{Hom}_{\Bbbk}(A, B)$ the set of all algebra homomorhpisms $A \to B$.

If m_A and m_B denote the multiplication maps of A and B respectively then we may think of a homomorphism, f, as a linear map which “commutes” with the multiplication map, that is $f \circ m_A = m_B \circ f$.

Alternatively, an algebra, A is both an abelian group under addition, and a monoid under multiplication, and an algebra homomorhpism is both a group and monoid homomorphism with respect to these structures.

2.2 Representations and Modules

There are two competing terminologies in the field, with slightly different notation and emphasis depending on which we use. We'll use the more modern notion of modules most of the time, but will occasionally and interchangeably use the notion of representations as well.

Definition 2.2.1 — Representation Let V be a $\Bbbk$-vector space and A a $\Bbbk$-algebra. Any $\rho \in \operatorname{Hom}(A, \operatorname{End} V)$ is called a **representation** of A. That is, a representation of A is an algebra homomorphism $\rho : A \to \operatorname{End} V$.

Definition 2.2.2 — Module Let A be a $\Bbbk$-algebra. A **left A-module**, M, is an abelian group, with the binary operation denoted $+$, equipped with a **left action**

$$. : A \times M \to M \tag{2.2.3}$$

$$(a, m) \mapsto a \,.\, m \tag{2.2.4}$$

such that for all $a, b \in A$ and $m, n \in M$ we have[a]

M1 $(ab) \,.\, m = a \,.\, (b \,.\, m)$ (note that (ab) is the product in A);

M2 $1 \,.\, m = m$.

M3 $a \,.\, (m + n) = a \,.\, m + a \,.\, n$;

M4 $(a + b) \,.\, m = a \,.\, m + b \,.\, m$;

One can similarly define a **right A-module**, M, as an abelian group with a **right action**

$$. : M \times A \to M \tag{2.2.5}$$

$$(m, a) \mapsto m \,.\, a \tag{2.2.6}$$

such that for all $a, b \in A$ and $m, n \in M$ we have

M1 $(m + n) \,.\, a = m \,.\, a + n \,.\, a$;

M2 $m \,.\, (a + b) = m \,.\, a + m \,.\, b$;

M3 $m \,.\, (ab) = (m \,.\, a) \,.\, b$;

M4 $m \,.\, 1 = m$.

A **two-sided A-module** is then an abelian group, M, which is simultaneously a left and right A-module satisfying

$$a \,.\, (m \,.\, b) = (a \,.\, m) \,.\, b \tag{2.2.7}$$

for all $a, b \in A$ and $m \in M$.

[a] Note that M1 and M2 simply say that this is a group action on the set M, and M3 and M4 two impose that this group action is compatible with both the group operation and addition in the algebra.

When it doesn't risk confusion we will write $a \,.\, m$ as am and $m \,.\, a$ as ma.

Note that a module is a generalisation of the notion of a vector space. In fact, if $A = \Bbbk$ then a module is exactly a vector space.

More compactly, one can define a right A-module as a left A^{op}-module, where A^{op} is the **opposite algebra** of A, defined to be the same underlying vector space with multiplication $*$ defined by $a * b = ba$, where ba is the multiplication in A. Because of this we will almost never have reason to work with right modules, we can always turn them into a left module over the opposite algebra instead.

Note that if A is commutative every left A-module is a right A-module and vice versa, and also a two-sided module.

Without further clarification the term "module" will mean

- a left module if A is not necessarily commutative;
- a two sided module if A is commutative.

A representation of A and an A-module carry exactly the same information. Given a representation, $\rho : A \to \operatorname{End} V$ we may define a group action on V by $a \,.\, v = \rho(a)v$. Composition in $\operatorname{End} V$ is exactly repeated application of this action: $[\rho(a)\rho(b)]v = \rho(a)[\rho(b)v]$ (M1). The unit of $\operatorname{End} V$ is the identity morphism, id_V, and $1 \in A$ must map to id_V, so $\rho(1)v = \mathrm{id}_V v = v$ (M2). Linearity of $\rho(a)$ means that $\rho(a)(v + w) = \rho(a)v + \rho(v)w$ (M3). Linearity of ρ means that $\rho(a + b)v = \rho(a)v + \rho(b)v$ (M4).

Conversely, given an A-module, M, we can define scalar multiplication by $\lambda \in \Bbbk$ on M by $\lambda m = (\lambda 1)m$ where $\lambda 1$ is scalar multiplication in A. This makes M a vector space, and we may define a morphism $\rho : A \to \operatorname{End} M$ by defining $\rho(a)$ by $\rho(a) = a \,.\, m$, which uniquely determines $\rho(a)$, say by considering the action on some fixed basis of M.

Further, these two constructions are inverse, given a module if we construct the corresponding representation then construct the corresponding module from that we get back the original module, and vice versa. This means that the notion of a representation and a module really are the same, and we don't need to distinguish between them. We will use whichever terminology and notation is better suited to the problem, which is usually the module terminology and notation.

Proposition 2.2.8 Let V be a $\Bbbk$-vector space, G a group, and $\rho : G \to GL(V)$ a group homomorphism. We may define a $\Bbbk G$-module by extending this map linearly, defining

$$\left(\sum_{g \in G} c_g g\right) . \, v = \sum_{g \in G} c_g \rho(g) v. \tag{2.2.9}$$

Conversely, given a left $\Bbbk G$-module on V we may define a group homomorphism $\rho : G \to \mathrm{GL}(V)$ by defining $\rho(g)$ to be the linear operation $v \mapsto g \,.\, v$.

Proof. This is just a special case of the equivalence of representations and modules discussed above. □

Note that a **group representation** is defined to be a group homomorphism $\rho\colon G \to \mathrm{GL}(V)$. The above result shows that a group representation of G is exactly the same as an algebra representation of $\Bbbk G$, so we can just study algebras.

Definition 2.2.10 — Regular Representation Let $V = A$ be an algebra and define $\rho\colon A \to \operatorname{End} A$ by $\rho(a)b = ab$. This is called the **left regular representation**. Similarly, the **right regular representation** is given by defining $\rho(a)b = ba$.

2.3 Direct Sums

The goal of much of representation theory is to classify possible representations. To do this we usually decompose representations into smaller parts that can be more easily classified. This decomposition is done by the direct sum.

Definition 2.3.1 — Direct Sum Let M and N be A-modules. The **direct sum**, $M \oplus N$, is the A-module given by the direct sum of the underlying abelian groups equipped with the action

$$a(m \oplus n) = am \oplus an \tag{2.3.2}$$

for all $a \in A$, $m \in M$ and $n \in N$.

The required properties follow immediately from the definition:

M1 $(ab)(m\oplus n) = (ab)m\oplus(ab)n = a(bm)\oplus a(bn) = a(bm\oplus bn) = a(b(m\oplus n))$;

M2 $1(m \oplus n) = 1m \oplus 1n = m \oplus n$;

M3 $a((m \oplus n) + (m' \oplus n')) = a((m + m') \oplus (n + n')) = a(m + m') \oplus a(n + n') = (am + am') \oplus (an + an') = (am \oplus an) + (am' \oplus an') = a(m \oplus n) + a(m' \oplus n')$;

M4 $(a + b)(m \oplus n) = (a + b)m \oplus (a + b)n = (am + bm) \oplus (an + bn) = (am \oplus an) + (bm \oplus bn) = a(m \oplus n) + b(m \oplus n)$.

Definition 2.3.3 — Submodule Let M be a left A-module. An abelian subgroup $N \trianglelefteq M$ is a **A-submodule** if $AN \subseteq N$. In this case we say that N is **invariant** under the action of A.

Note that by AN we mean

$$AN = \{an \mid a \in A, n \in N\}. \tag{2.3.4}$$

So $AN \subseteq N$ means that $an \in N$ for all $a \in A$ and $n \in N$. Thus, invariance means that no element of N leaves N under the action of A.

Definition 2.3.5 — Trivial Submodule Every A-module, M, admits two submodules, M itself and the zero module, 0, which contains only 0. We call these **trivial submodules**.

Note that some texts call only 0 the trivial submodule, and make the distinction of a submodule vs a *proper* submodule, the distinction being that M is not a proper submodule of M. Then when we say "nontrivial submodule" these texts will say "nontrivial proper submodule".

Definition 2.3.6 — Simple Submodule Let M be an A-module. We say that M is **simple** or **irreducible** if it contains no nontrivial submodules.

Typically "simple" is used for modules and "irreducible" is used more for representations, although irreducible is used for both.

Definition 2.3.7 — Semisimple Let M be an A-module. Then M is **semisimple** or **completely reducible** if it can be written as a direct sum of finitely many simple modules.

That is, M is semisimple if

$$M = \bigoplus_{i=1}^{n} N_i = N_1 \oplus \cdots \oplus N_n \tag{2.3.8}$$

where each N_i is simple. Note that we define the empty sum to be the zero module, so the zero module is considered semisimple (and also simple, since it contains only itself as a submodule).

Again, "semisimple" is typically used only for modules, and "completely reducible" is used primarily for representations.

Definition 2.3.9 — Indecomposable Let M be an A-module. Then M is **indecomposable** if M cannot be written as a direct sum of nontrivial modules.

The nontrivial requirement here just rules out decompositions of the form[1] $M = M \oplus 0$.

Note that every simple (irreducible) module is indecomposable, since if it had a decomposition $M = N_1 \oplus N_2$ with N_i nontrivial then their is a canonical copy of each N_i as a submodule of M. The converse does not hold in general, not all indecomposable modules are irreducible. It is possible that M contains a submodule, N, but that there is no submodule N' such that $M = N \oplus N'$. Contrast this to finite dimensional vector spaces where we can take N' to be the orthogonal complement (with respect to some inner product) of N and this direct sum holds. We can still form the orthogonal complement of a submodule, but it will not, in general, be a submodule. There are, however, many special cases, such as finite dimensional complex representations of (group algebras) finite groups, where the orthogonal complement can be defined in such a way that it is a submodule, and in this case indecomposable and irreducible coincide.

[1] Note that with our definition of the direct sum this really only holds up to isomorphism, since M has elements m whereas $M \oplus 0$ has elements $(m, 0)$. However, we're yet to define morphisms between modules, and once we do we'll see that $\oplus$ is the product in the category of modules, and as such is only defined up to isomorphism, so we may as well momentarily take the isomorphism that makes this equality true.

One of the main goals of representation theory is to classify all indecomposable modules of a given algebra. This then gives us an understanding of *all* modules over that algebra, since any nonsimple or decomposable module may be realised as a direct sum of these classified indecomposable modules.

2.4 Module Homomorphisms

Definition 2.4.1 — Module Homomorphism Let M and N be A-modules. An **A-module homomorphism** or **intertwiner** is a homomorphism of the underlying abelian groups $\varphi : M \to N$ which “commutes” with the action of A, by which we mean

$$\varphi(a \,.\, m) = a \,.\, \varphi(m) \tag{2.4.2}$$

for all $a \in A$ and $m \in M$.
An invertible A-module homomorphism is called an **isomorphism** of A-modules.
Homomorphisms of right A-modules may be defined similarly.

Notation 2.4.3 We write $\mathrm{Hom}_A(M,N)$ for the set of A-module homomorphisms $M \to N$. Note that $\mathrm{Hom}_A(M,N) \subseteq \mathrm{Hom}_{\mathsf{Ab}}(M,N)$ where $\mathrm{Hom}_{\mathsf{Ab}}(M,N)$ is the set of all homomorphisms $M \to N$ of the underlying abelian groups.

Note that in $\varphi(a.m)$ a is acting on an element of M, and in $a.\varphi(m)$ a is acting on an element of N, so these are in general different actions. Writing $a\,.$ for the map $x \mapsto a\,.\,x$ we can express the condition of commuting action as the commutativity of the diagram

$$\begin{tikzcd} M \arrow[r, "\varphi"] \arrow[d, "a."'] & N \arrow[d, "a."] \\ M \arrow[r, "\varphi"'] & N \end{tikzcd} \tag{2.4.4}$$

for all $a \in A$.

Lemma 2.4.5 Isomorphisms of A-modules are exactly bijective morphisms of A-modules.

Proof. Let $\varphi : M \to N$ be a bijective morphism of A-modules. Then the (set-theoretic) inverse, $\varphi^{-1} : N \to M$, exists. We claim that this is a morphism of A-modules. This follows by taking $n \in N$ to be the image of $m \in M$ under φ, giving

$$\varphi^{-1}(a\,.\,n) = \varphi^{-1}(a\,.\,\varphi(m)) = \varphi^{-1}(\varphi(a\,.\,m)) = a\,.\,m = a\,.\,\varphi^{-1}(m). \tag{2.4.6}$$

Conversely, if $\varphi : M \to N$ is an isomorphism of A-modules it must necessarily be that φ^{-1} is the (set-theoretic) inverse of the underlying function of φ, and so φ must be bijective. □

If we instead talk of representations (V, ρ) and (W, σ) then a homomorphism of representations, $\varphi: V \to W$, must satisfy $\varphi(\rho(a)v) = \sigma(a)\varphi(v)$. Further, by linearity of ρ and σ and the fact that $\rho(1) = \mathrm{id}_V$ and $\sigma(1) = \mathrm{id}_W$ we have that for $\lambda \in \Bbbk$

$$\varphi(\lambda m) = \varphi(\rho(1)\lambda m) = \varphi(\rho(\lambda 1)m) = \sigma(\lambda 1)\varphi(m) = \lambda\sigma(1)\varphi(m) = \lambda\varphi(m). \quad (2.4.7)$$

This shows that φ must be a linear map $\varphi: V \to W$. In fact, we can *define* a homomorphism of representations to be a linear map $\varphi: M \to N$ satisfying $\varphi(\rho(a)m) = \sigma(a)\varphi(m)$. We will also write $\mathrm{Hom}_A(V, W)$ for the set of representation morphisms $V \to W$. Note then that $\mathrm{Hom}_A(V, W) \subseteq \mathrm{Hom}_{\Bbbk\text{-}\mathsf{Vect}}(V, W)$ where $\mathrm{Hom}_{\Bbbk\text{-}\mathsf{Vect}}(V, W)$ is the set of linear maps $V \to W$ of the underlying vector spaces. Using the notation Hom_A for both modules and representations is justified by the following remark.

Remark 2.4.8 There is a category, A-Mod (Mod-A), with left (right) A-modules as objects and A-module homomorphisms as morphisms. Similarly, there is a category $\mathsf{Rep}(A)$ of representations of A with objects being representations (V, ρ) and morphisms being homomorphisms of representations.

In Section 2.2 we showed that we have a mapping $F: A\text{-}\mathsf{Mod} \to \mathsf{Rep}(A)$ constructing a representation from a module, and a mapping $G: \mathsf{Rep}(A) \to A\text{-}\mathsf{Mod}$ constructing a module from a representation. In the discussion above we extend this mapping to define a representation homomorphism from a module homomorphism. We can also ignore the requirement of linearity with respect to scalar multiplication in the definition of a representation homomorphism to recover a module homomorphism. Further, applying either of these constructions to the appropriate identity map just gives the identity, and both constructions preserve composition. These operations on homomorphisms are also inverses of each other. Thus, F and G are functors and we have $FG = \mathrm{id}_{\mathsf{Rep}(A)}$ and $GF = \mathrm{id}_{A\text{-}\mathsf{Mod}}$. Thus, A-Mod and $\mathsf{Rep}(A)$ are isomorphic as categories, justifying the fact that we will soon cease to distinguish between them.

Lemma 2.4.9 The category A-Mod defined above is indeed a category.

Proof. First note that $\mathrm{id}_M: M \to M$ is an A-module homomorphism for any A-module, M, since we have

$$\mathrm{id}_M(a \,.\, m) = a \,.\, m = a \,.\, \mathrm{id}_M(m). \quad (2.4.10)$$

Now note that if $\varphi: M \to N$ and $\psi: N \to P$ are module homomorphisms then $\psi \circ \varphi: M \to P$ is a module homomorphism since

$$(\psi \circ \varphi)(a \,.\, m) = \psi(\varphi(a \,.\, m)) = \psi(a \,.\, \varphi(m)) = a \,.\, \psi(\varphi(m)) = a \,.\, (\psi \circ \varphi)(m)$$

for all $a \in A$ and $m \in M$. Finally, composition is just composition of the underlying functions, which is associative. □

2.5 Schur's Lemma

We can now give one of the first results of representation theory. It places a restriction on the types of morphisms we can have between modules when one or more of the modules is simple. We give the result as a proposition and a corollary, although for historical reasons it's called a lemma. The proposition is more general, and the corollary is a special case. Both are known as Schur's lemma, with context determining if we use the more general result or the special case.

Before we can prove this result however we need a couple of results about kernels and images of module morphisms.

Lemma 2.5.1 Let $\varphi\colon M \to N$ be a morphism of modules. Then $\ker\varphi$ is a submodule of M and $\operatorname{im}\varphi$ is a submodule of N.

Proof. Step 1: $\ker\varphi$
We know that $\ker\varphi$ is a subgroup of M, so we only need to show that it is invariant under the action of A. Take $m \in \ker\varphi$, that is $m \in M$ is such that $\varphi(m) = 0$, and $a \in A$. Then

$$\varphi(a \,.\, m) = a \,.\, \varphi(m) = a \,.\, 0. \tag{2.5.2}$$

For arbitrary $m' \in M$ we have

$$a \,.\, 0 = a \,.\, (m' - m') = (a \,.\, m') - (a \,.\, m') = 0 \tag{2.5.3}$$

so $a \,.\, 0 = 0$ for any $a \in A$, and thus $\varphi(a \,.\, m) = a \,.\, 0 = 0$, so $a \,.\, m \in \ker\varphi$.

Step 2: $\operatorname{im}\varphi$
We know that $\operatorname{im}\varphi$ is a subgroup of M, so we only need to show that it is invariant under the action of A. Take $n \in \operatorname{im}\varphi$ and $a \in A$. There exists some $m \in M$ such that $n = \varphi(m)$. Then

$$a \,.\, n = a \,.\, \varphi(m) = \varphi(a \,.\, m) \tag{2.5.4}$$

and $a \,.\, m \in M$ so $a \,.\, n \in \operatorname{im}\varphi$. □

Proposition 2.5.5 — Schur's Lemma Let $\Bbbk$ be any (not necessarily algebraically closed) field, and let A be an algebra over $\Bbbk$. Let M and N be A-modules and let $\varphi\colon M \to N$ be a morphism of A-modules. Then

1. if M is simple either $\varphi = 0$ or φ is injective;
2. if N is simple either $\varphi = 0$ or φ is surjective.

Combined if M and N are simple then either $\varphi = 0$ or φ is an isomorphism.

Proof. Step 1: M Simple
Let M be simple, so its only submodules are 0 and M. We know that $\ker\varphi$ is a submodule of M, so there are two cases to consider:

- If $\ker\varphi = M$ then every element of M maps to 0, so $\varphi = 0$.
- If $\ker\varphi = 0$ then φ is injective[a].

STEP 2: N SIMPLE

Let N be simple, so its only submodules are 0 and N. We know that $\operatorname{im}\varphi$ is a submodule of N, so there are two cases to consider:

- If $\operatorname{im}\varphi = 0$ then every element of M maps to 0, so $\varphi = 0$.
- If $\operatorname{im}\varphi = N$ then φ is surjective.

□

[a] We know that for group homomorphisms if the kernel is trivial then the map is injective, and injectivity is a set-theoretic property, so it still holds when we add the extra structure of the A-action

Corollary 2.5.6 — Schur's Lemma Let $\Bbbk$ be an algebraically closed field, and let A be an algebra over $\Bbbk$. Let V be a finite dimensional representation of A. Then any representation homomorphism $\varphi\colon V \to V$ is a multiple of the identity. That is, $\varphi = \lambda \mathrm{id}_V$ for $\lambda \in \Bbbk$. Note that $\lambda = 0$ subsumes the trivial case.

Proof. Let $\lambda \in \Bbbk$ be an eigenvalue of φ with corresponding eigenvector $v \in V$. Note that eigenvalues exist because

a) V is finite dimensional so the determinant may be defined as a polynomial in the entries of some matrix representing φ in a fixed basis; and

b) $\Bbbk$ is algebraically closed, so this polynomial has roots.

Then by definition $\varphi(v) = \lambda v$ which we can rearrange to $(\varphi - \lambda\mathrm{id}_V)v = 0$. Thus, $v \in \ker(\varphi - \lambda\mathrm{id}_V)$, and since eigenvectors are, by definition, nonzero this means that $\ker(\varphi - \lambda\mathrm{id}_V) \neq 0$, so $\varphi - \lambda\mathrm{id}_V$ is not injective, so by Schur's lemma (Proposition 2.5.5) we must have that $\varphi - \lambda\mathrm{id}_V = 0$. Thus, $\varphi = \lambda\mathrm{id}_V$. □

Corollary 2.5.7 Let A be a commutative algebra over an algebraically closed field, $\Bbbk$. Then all nontrivial finite dimensional irreducible representations of A are one dimensional.

Proof. Let V be a finite dimensional irreducible representation of A. For $a \in A$ define a map $\varphi_a\colon V \to V$ by $v \mapsto \varphi_a(v) = a \,.\, v$. This is an intertwiner: take $b \in A$ and $v \in V$, then we have

$$\varphi_a(b\,.\,v) = a\,.\,(b\,.\,v) = (ab)\,.\,v = (ba)\,.\,v = b\,.\,(a\,.\,v) = b\,.\,\varphi_a(v). \tag{2.5.8}$$

Note that this is only true because $ab = ba$.
By Schur's lemma (Corollary 2.5.6) there exists some $\lambda_a \in \Bbbk$ such that $\varphi_a = \lambda_a \mathrm{id}_V$. Then $a \,.\, v = \varphi_a(v) = \lambda_a v$, so every $a \in A$ acts as scalar multiplication. This means that any subspace is invariant, since every subspace is, by definition, invariant under scalar multiplication. Thus, the only way that a representation can have no nontrivial invariant subspaces if if it only has trivial subspaces, which is only true if it is one dimensional (zero dimensional being ruled out by the assumption that the representation is nontrivial). □

Example 2.5.9 Consider $A = \Bbbk[x]$, which is a commutative algebra. We can determine all irreducible representations of A.
A representation, $\rho : \Bbbk[x] \to \operatorname{End} V$, is fully determined by the value of $\rho(x)$, since given an arbitrary polynomial, $f(x) = \sum_{i=1}^n a_i x^i$, its action on $v \in V$ is determined through linearity by

$$f(x) \,.\, v = \rho(f(x))v = \rho\left(\sum_{i=1}^n a_i x^i\right) v = \sum_{i=1}^n a_i \rho(x)^i v. \tag{2.5.10}$$

Further, by Corollary 2.5.7 we know that any irreducible representation of $\Bbbk[x]$ is one dimensional, so it must be that $\rho(v) = \lambda v$ for some $\lambda \in \Bbbk$.
Let V_λ denote the one-dimensional representation in which x acts as scalar multiplication by λ. We claim that $V_\lambda \cong V_\mu$ if and only if $\lambda = \mu$. Suppose that $\varphi : V_\lambda \to V_\mu$ is an isomorphism. Then $\varphi(x \,.\, v) = \varphi(\lambda v) = \lambda\varphi(v)$ and $\varphi(x \,.\, v) = x \,.\, \varphi(v) = \mu\varphi(v)$. Thus, $\lambda = \mu$.
So, we have classified all irreducible representations of $\Bbbk[x]$, they are precisely the one dimensional vector spaces, V_λ for $\lambda \in \Bbbk$ in which $\rho(x) = \lambda \mathrm{id}_{V_\lambda}$.
This result generalises to polynomials in an arbitrary number of variables, $\Bbbk[x_1, \dots, x_n]$. Then a representation is fully determined by the values of $\rho(x_1)$ through $\rho(x_n)$. Thus an irreducible representation is a one dimensional vector space, $V_{\lambda_1,\dots,\lambda_n}$ in which $\rho(x_i) = \lambda_i \mathrm{id}_{V_{\lambda_1,\dots,\lambda_n}}$.
Go back to the case of $A = \Bbbk[x]$. For a nontrivial ($\lambda \neq 0$) finite dimensional irreducible representation, V_λ, instead of starting with the action of x we can perform a change of variables and work with $y = x/\lambda$. Then we get the representation V_1. This means that all finite dimensional irreducible representations of $\Bbbk[x]$ are essentially the same, up to rescaling. This also means that they're pretty boring.
Indecomposable representations of $\Bbbk[x]$ are more interesting on the other hand. Let V be a finite dimensional representation. We can fix a basis and look at matrices. Suppose $B \in \operatorname{End} V$, then since we work over an algebraically closed field we know that the Jordan normal form of B exists after a basis change, allowing us to write the matrix of B as

$$B = \begin{pmatrix} J_{\lambda_1,n_1} & & & \\ & J_{\lambda_2,n_2} & & \\ & & \ddots & \\ & & & J_{\lambda_k,n_k} \end{pmatrix} \tag{2.5.11}$$

where J_{λ_i,n_i} is the $n_i \times n_i$ Jordan block matrix

$$J_{\lambda_i,n_i} = \begin{pmatrix} \lambda_i & 1 & & & \\ & \lambda_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda_i & 1 \\ & & & & \lambda_i \end{pmatrix}. \tag{2.5.12}$$

This block diagonal decomposition of B gives us a corresponding direct sum decomposition of V. Each Jordan block cannot be diagonalised (with the exception of the 1×1 Jordan blocks which are trivially diagonal). Thus we cannot further decompose B and so we cannot further decompose V. The result is that

$$V = \bigoplus_{i=1}^{k} V_{\lambda_i,n_i} \tag{2.5.13}$$

where $V_{\lambda_i,n_i} = \Bbbk^{n_i}$ is an n_i-dimensional vector space upon which the action of B is given by J_{λ_i,n_i}. Then taking $B = \varphi(x)$ defines a representation of $\Bbbk[x]$ on V, and specifically we have the subrepresentations V_{λ_i,n_i} in which x acts as the Jordan block J_{λ_i,n_i}.

2.6 Ideals and Quotients

Definition 2.6.1 — Ideals Let A be an algebra. A subspace, $I \subseteq A$, such that $AI \subseteq I$ is called a **left ideal**. Similarly if $IA \subseteq I$ then we call I a **right ideal**. A **two-sided ideal** is simultaneously a left and right ideal.

Note that by AI we mean $AI = \{ai \mid a \in A, i \in I\}$, so the condition that I is a left ideal is that $ai \in I$ for all $a \in A$ and $i \in I$.

Example 2.6.2

- Any algebra, A, always has 0 and A as ideals. If these are the only ideals then we call A **simple**.
- Any left (right) ideal is a submodule of the left (right) regular representation. This is simply identifying that A is an A-module with the action being left (right) multiplication and as such the notion of an ideal coincides with that of a submodule. Note that the notion of a simple module coincides with the notion of a simple algebra under this identification.
- If $f : A \to B$ is an algebra morphism then $\ker f$ is a two-sided ideal. We know that $\ker f$ is a subspace of A, so just note that if $a \in \ker f$ then $f(a) = 0$ and we have

$$f(ba) = f(b)f(a) = f(b)0 = 0 \tag{2.6.3}$$

and

$$f(ab) = f(a)f(b) = 0f(a) = 0 \tag{2.6.4}$$

so ab and ba are in $\ker f$.

We will say "ideal" when we mean either a left ideal. Note that in the commutative case all left ideals are right ideals and hence two-sided ideals, so we don't need to distinguish the three cases.

Notation 2.6.5 Let A be an algebra and $S \subseteq A$ a subset of A. Denote by $\langle S \rangle$ the two-sided ideal generated by S. That is,

$$\langle S \rangle = \operatorname{span}\{asb \mid s \in S, \text{ and } a, b \in A\}. \tag{2.6.6}$$

For example, consider $\Bbbk[x]$. Then $\langle x \rangle$ consists of all polynomials that can be factorised as $xf(x)$ where $f(x)$ is an arbitrary polynomial, so $f(x) = \sum_{i=0}^{n} a_i x^i$. Thus, $xf(x) = \sum_{i=0} a_i x^{i+1}$, so $\langle x \rangle$ consists of all polynomials with zero constant term. More generally, $\rangle x - a\rangle$ for $a \in \Bbbk$ consists of all polynomials which factorise as $(x-a)f(x)$ for an arbitrary polynomial $f(x)$, and thus this is the ideal consisting of all polynomials with a as a root.

The point of defining ideals is really in order to define quotients. In this way ideals are to algebras as normal subgroups are to groups.

Definition 2.6.7 — Quotient Let A be an algebra and $I \subseteq A$ an ideal. We define the **quotient** to be the algebra A/I whose elements are equivalence classes

$$[a] = a + I \coloneqq \{a' \in A \mid a - a' \in I\}. \tag{2.6.8}$$

Addition and scalar multiplication are defined by

$$[a] + [b] = (a + I) + (b + I) = [a + b] = a + b + I \tag{2.6.9}$$

and

$$\lambda[a] = [\lambda a] \tag{2.6.10}$$

for $a, b \in A$ and $\lambda \in \Bbbk$.

Lemma 2.6.11 The quotient of an algebra by an ideal is again an algebra.

Proof. Let A be an algebra and $I \subseteq A$ an ideal. Note that the quotient of a vector space by any subspace is again a vector space, so we need only define a multiplication operation on this vector space. We do so by defining

$$[a][b] = (a + I)(b + I) \coloneqq [ab] = ab + I. \tag{2.6.12}$$

We need to show that this is well-defined and satisfies the properties of

multiplication in an algebra.
STEP 1: WELL-DEFINED
Let $a, a' \in A$ be representatives of the same equivalence class, $[a] = [a']$. Then by definition $a - a' \in I$. For $b \in A$ we then have

$$[a][b] = [ab] = [a'b + (a - a')b] = [a'b] = [a'][b]. \tag{2.6.13}$$

Here we've used the fact that $a - a' \in I$ and I is an ideal so $(a - a')b \in I$, and we can add any element of I inside an equivalence class without leaving the equivalence class. Similarly, one can show that $[a][b] = [a][b']$ whenever $[b] = [b']$. Thus, this product is well-defined.

STEP 2: ALGEBRA
Linearity in the first argument follows from a direct calculation using the properties of quotient spaces:

$$\begin{aligned}[(a + \lambda a')b] &= [ab + \lambda a'b] = [ab] + \lambda[a'b] \\ &= [a][b] + \lambda[a'][b] = ([a] + \lambda[a'])[b] = [a + \lambda a'][b]\end{aligned} \tag{2.6.14}$$

for $a, a', b \in A$ and $\lambda \in \Bbbk$. Linearity in the second argument follows similarly. Associativity follows from

$$[a]([b][c]) = [a][bc] = [a(bc)] = [(ab)c] = [ab][c] = ([a][b])[c]. \tag{2.6.15}$$

Unitality follows from

$$[1][a] = [1a] = [a], \qquad \text{and} \qquad [a][1] = [a1] = [a]. \tag{2.6.16}$$

□

2.6.1 Generators and Relations

One of the most common ways to define an algebra is as a quotient of another algebra by some ideal given in terms of generators. The most common starting place is the free algebra, $\Bbbk\langle x_1, \dots, x_m\rangle$. We can then take $f_1, \dots, f_n \in \Bbbk\langle x_1, \dots, x_m\rangle$, and form an ideal, $\langle f_1, \dots, f_n\rangle$. Then we may form the algebra

$$A = \Bbbk\langle x_1, \dots, x_m\rangle / \langle f_1, \dots, f_n\rangle. \tag{2.6.17}$$

Intuitively, elements of this are non-commutative polynomials in the x_i subject to the constraint that anywhere that we can manipulate the polynomial to be written with f_i we can set that f_i equal to zero.

For example, let $f_{i,j} = x_i x_j - x_j x_i$ for $i, j = 1, \dots, m$. Consider the algebra $A = \Bbbk\langle x_1, \dots, x_m\rangle / \langle f_{i,j}\rangle$ consists of non-commutative polynomials in x_i subject to the condition that $x_i x_j - x_j x_i = 0$, which is to say $x_i x_j = x_j x_i$, which is exactly the condition that the x_i *do* commute with each other.

Another example is $A = \Bbbk\langle x_1, \dots, x_n\rangle / \langle x_i^2 - e, x_i x_{i+1} x_i - x_{i+1} x_i x_{i+1}\rangle$. This sets $x_i^2 = e$ and $x_i x_{i+1} x_i = x_{i+1} x_i x_{i+1}$ (called the **braid relation**). These are exactly the relations defining the symmetric group, S_n, when we interpret x_i as the transposition $(i\, i+1)$. We're also taking linear combinations of these x_i, so $A = \Bbbk S_n$.

2.6.2 Quotient Modules

Definition 2.6.18 — Quotient Module Let M be an A-module and N a submodule of M. We define the **quotient module**, M/N, to be the module consisting of equivalence classes

$$[m] = m + N \coloneqq \{m' \in M \mid m - m' \in M\}. \tag{2.6.19}$$

Addition in this module is defined by

$$[m] + [m'] = [m + m'] \tag{2.6.20}$$

for $m, m' \in M$ and the action of A is given by

$$a\,.\,[m] = [a\,.\,m] \tag{2.6.21}$$

for $a \in A$ and $m \in M$.

Lemma 2.6.22 The quotient of a module by a submodule is again a module.

Proof. Let M be an A-module with $N \subseteq M$ a submodule. Then N is a subgroup of an abelian group, and so is automatically a normal subgroup. Then we know that M/N is an abelian group also.
Suppose that $[m] = [m']$, that is m and m' are representatives of the same equivalence class. Then $m' - m \in N$. We then have

$$\begin{aligned} a\,.\,[m] &= a\,.\,[m' + (m - m')] = [a\,.\,(m' + (m - m'))] \\ &= [a\,.\,m' + a\,.\,(m - m')] = [a\,.\,m'] = a\,.\,[m']. \end{aligned} \tag{2.6.23}$$

Here we've used the fact that $m' - m \in N$ and N is a submodule so $a\,.\,(m' - m) \in N$ as well. So, the action of $a \in A$ on $[m] = [m']$ is well-defined.
It remains to show that the action of A on M/N makes it an A-module:

M1 $(ab)\,.\,[m] = [(ab)\,.\,m] = [a\,.\,(b\,.\,m)] = a\,.\,[b\,.\,m] = a\,.\,(b\,.\,[m])$;

M2 $1\,.\,[m] = [1\,.\,m] = [m]$;

M3 $a\,.\,([m] + [n]) = a\,.\,[m + n] = [a\,.\,(m + n)] = [a\,.\,m + a\,.\,n] = [a\,.\,m] + [a\,.\,n] = a\,.\,[m] + a\,.\,[n]$;

M4 $(a + b)\,.\,[m] = [(a + b)\,.\,m] = [a\,.\,m + b\,.\,m] = [a\,.\,m] + [b\,.\,m] = a\,.\,[m] + b\,.\,[m]$

for all $a, b \in A$ and $m, n \in M$. □

Remark 2.6.24 Consider the left regular representation of A. As we have mentioned ideals of A are precisely submodules of the regular representation. It follows that A/I is a left A-module precisely when I is a left ideal.

Three

Tensor Products

3.1 Tensor Product of Modules

We first define the tensor product of R-modules (R a ring). This definition can also be applied to A-modules (A an algebra) without modification.

Definition 3.1.1 — Tensor Product Let R be a ring, M a right R-module, and N a left R-module. Then the **tensor product**, $M \otimes_R N$, is the abelian group

$$\frac{F(\{m \otimes n \mid m \in M, n \in N\})}{I} \tag{3.1.2}$$

where $F(X)$ denotes the free abelian group on the set X and I is the normal subgroup generated from all elements of the form

- $(m + m') \otimes n - m \otimes n - m' \otimes n$;
- $m \otimes (n + n') - m \otimes n - m \otimes n'$;
- $(m\,.\,r) \otimes n - m \otimes (r\,.\,n)$

with $m, m' \in M$, $n, n' \in N$ and $r \in R$.

! The tensor product does not, in general, have the structure of an R-module. It is just an abelian group. In a sense the R-actions of M and N are "used up" in the construction and don't "survive" to produce a sensible notion of an R-action on $M \otimes_R N$.

Notation 3.1.3 When R is clear from context we will write $M \otimes N$ instead of $M \otimes_R N$. Conversely, if needed we'll write $m \otimes_R n$ for elements of $M \otimes_R N$ if there are multiple ways to define the tensor product.

Intuitively, $M \otimes_R N$ consists of sums of elements which we write as[1] $m \otimes n$ with $m \in M$ and $n \in N$. So, one element of $M \otimes_R N$ might be

$$m_1 \otimes n_1 + m_2 \otimes n_2 + m_3 \otimes n_3 \tag{3.1.4}$$

with $m_i \in M$ and $n_i \in N$. Note that there are no factors of R here, this is purely an operation in the free group. The quotient imposes that in $M \otimes_R N$ we have the

[1] We should write $[m \otimes n]$ or something similar, since what we actually have is the equivalence class of $m \otimes n$ in $F(\{m \otimes n\})/I$.

relations

$$(m + m') \otimes n = m \otimes n + m' \otimes n; \tag{3.1.5}$$
$$m \otimes (n + n') = m \otimes n + m \otimes n'; \tag{3.1.6}$$
$$(m \,.\, r) \otimes n = m \otimes (r \,.\, n). \tag{3.1.7}$$

As we mentioned the tensor product of a right and left R-module is not, in general, an R-module in any consistent way. In order for the tensor product to be a module we need to have some extra module structure present in one of the two modules which then remains after the tensor product is formed. Of course, this extra structure must be compatible with the existing structure, and it turns out that the following is exactly the right definition for this purpose.

Definition 3.1.8 — Bimodule Left A and B be associative unital $\Bbbk$-algebras. An $(\boldsymbol{A}, \boldsymbol{B})$**-bimodule** is an abelian group, M, which is both a left A-module and a right B module in such a way that

$$(a \,.\, m) \,.\, b = a \,.\, (m \,.\, b) \tag{3.1.9}$$

for all $a \in A$, $b \in B$, and $m \in M$.

Example 3.1.10 Let V be a $\Bbbk$-vector space and a left A-module. Then V is an $(A, \Bbbk)$-bimodule where $a \,.\, v$ is just the action of A on V as an A-module and $v \,.\, \lambda = \lambda v$ is just scalar multiplication by elements of $\Bbbk$. That this is a bimodule follows because

$$a \,.\, (v \,.\, \lambda) = a \,.\, (\lambda v) = \lambda(a \,.\, v) = (a \,.\, v) \,.\, \lambda \tag{3.1.11}$$

having used the fact that the action of a on v is $\Bbbk$-linear.
In fact, we can define a bimodule first (just combining the definitions of a left and right module), then a left A-module is an $(A, \Bbbk)$-bimodule, and a right A-module is a $(\Bbbk, A)$-bimodule.

Lemma 3.1.12 Let M be an (A, B)-bimodule, and N a left B-module. Then $M \otimes_B N$ is a left A-module with $a \,.\, (m \otimes n) \coloneqq (a \,.\, m) \otimes n$.

Proof. First note that as an (A, B)-bimodule M is, in particular, a right B-module. Thus, the tensor product $M \otimes_B N$ is defined as the quotient of a free abelian group by an ideal, and so is again an abelian group. It remains only to show that this abelian group equipped with the action of A on the first factor is an A-module.
To do so take an arbitrary element of $M \otimes_B N$, which is of the form $\sum_{i \in I} m_i \otimes n_i$ where I is some finite indexing set, $m_i \in M$ and $n_i \in N$. We are free to define the action of A on this element to be

$$a \,.\, \left(\textstyle\sum_{i \in I} m_i \otimes n_i\right) \coloneqq \textstyle\sum_{i \in I} (a \,.\, m_i) \otimes n_i. \tag{3.1.13}$$

Then when I is a singleton this reduces to $a \,.\, (m \otimes n) = (a \,.\, m) \otimes n$ as

required.
We can now prove that this makes $M \otimes_B N$ a left A-module:

M1 $(ab) \,.\, \sum_i m_i \otimes n_i = \sum_i ((ab) \,.\, m_i) \otimes n_i = \sum_i (a \,.\, (b \,.\, m_i)) \otimes n_i = a \,.\, \sum_i (b \,.\, m_i) \otimes n_i = a \,.\, \big(b \,.\, \sum_i m_i \otimes n_i\big)$;

M2 $1 \,.\, \sum_i m_i \otimes n_i = \sum_i (1 \,.\, m_i) \otimes n_i = \sum_i m_i \,.\, n_i$;

M3 $a.\big(\sum_{i\in I} m_i \otimes n_i + \sum_{j\in J} m_j \otimes n_j\big) = a.\big(\sum_{i\in I\sqcup J} m_i \otimes n_i\big) = \sum_{i\in I\sqcup J}(a.\, m_i) \otimes n_i = \sum_{i\in I}(a \,.\, m_i) \otimes n_i + \sum_{j\in J}(a \,.\, m_j) \otimes n_j$;

M4 $(a+b).\sum_i m_i \otimes n_i = \sum_i ((a+b).m_i) \otimes n_i = \sum_i (a.m_i + b.m_i) \otimes n_i = \sum_i (a \,.\, m_i) \otimes n_i + (b \,.\, m_i) \otimes n_i = a \,.\, \sum_i m_i \otimes n_i + b \,.\, \sum_i m_i \otimes n_i$.

□

Similarly, if M is a right A-module and N is an (A, B)-bimodule then $M \otimes_A N$ is a right B-module with the action given by $(m \otimes n) \,.\, b = m \otimes (n \,.\, b)$.

Example 3.1.14 Any $\Bbbk$-vector space, V, is a $(\Bbbk, \Bbbk)$-bimodule, defining $\lambda \,.\, v = \lambda v = v \,.\, \lambda$ for $\lambda \in \Bbbk$ and $v \in V$. If U is some other vector space then we can form the $\Bbbk$-module $V \otimes_\Bbbk U$, which is of course just the usual tensor product of vector spaces.
In fact, this works for any commutative algebra, A, we can take any A-module as an (A, A)-bimodule, so if M and N are A-modules then $M \otimes_A N$ is an A-module.

3.1.1 Universal Property

The tensor product may also be defined via a universal property.

Lemma 3.1.15 Let M be an right A-module, and let N be a left A-module. Then for any abelian group, G, and any group homomorphism $f : M \times N \to G$ satisfying ... there is a unique group homomorhpism $\bar{f} : M \otimes_A N \to G$ such that $\bar{f}(m \otimes n) = f(m, n)$ for all $m \in M$ and $n \in N$. That is, the diagram

$$\begin{tikzcd} M \times N \arrow[r, "- \otimes -"] \arrow[dr, "f"'] & M \otimes_A N \arrow[d, "\exists! \bar{f}"] \\ & G \end{tikzcd} \tag{3.1.16}$$

commutes.

Proof. To make this diagram commutes we can define $\bar{f}(m \otimes n) = f(m, n)$. The fact that $\bar{f}$ is a group homomorhpism means that this uniquely defines the value of $\bar{f}$ on any element of $M \otimes_A N$ by

$$\bar{f}\left(\sum_i m_i \otimes n_i\right) = \sum_i f(m_i, n_i). \tag{3.1.17}$$

Note that $\mathrm{Hom}_A(M,N)$ inherits the module structure of N via pointwise operations. Let M be an (A,B)-bimodule, N a (B,C)-bimodule, and P an (A,C)-bimodule for three algebras, A, B, and C. Then we can form the tensor product $M \otimes_B N$, which is an A-module, and we can consider the hom-set $\mathrm{Hom}_A(M \otimes_B N, P)$, of left A-module homomorphisms, this is itself an A-module, and in fact is an (A,A)-bimodule. We can also form the hom-set $\mathrm{Hom}_C(N,P)$ of right C-module homomorhpisms, which is an left A-module under pointwise action using the A-module structure of P. Then we can take the hom-set $\mathrm{Hom}_B(M, \mathrm{Hom}_C(N,P))$, which is an A-module under pointwise the action. Then it turns out that we actually have an isomorphism

$$\mathrm{Hom}_A(M \otimes_B N, P) \xrightarrow{\cong} \mathrm{Hom}_B(M, \mathrm{Hom}_C(N,P)) \tag{3.1.18}$$

given by sending f to g defined by $g(m)(n) = f(m \otimes n)$. This isomorphism is natural in all objects, and thus this is an adjunction.

3.2 Tensor Algebra

Definition 3.2.1 — Tensor Algebra Let V be a vector space over $\Bbbk$. Then the **tensor algebra**, TV, is defined to be

$$\bigoplus_{n=0}^{\infty} V^{\otimes n} = \Bbbk \oplus V \oplus (V \otimes V) \oplus (V \otimes V \otimes V) \oplus \cdots . \tag{3.2.2}$$

Multiplication is defined by $ab = a \otimes b \in V^{\otimes(n=m)}$ for $a \in V^{\otimes n}$ and $b \in V^{\otimes m}$, and extended linearly.

Lemma 3.2.3 Let V be an n-dimensional vector space over $\Bbbk$. Then TV is isomorphic to $\Bbbk\langle x_1, \ldots, x_n \rangle$, the free algebra on n indeterminates.

Proof. Pick a basis for V. Identify this basis with the x_i. Elements of TV are linear combinations of tensor products of these basis elements, so we can identify them with polynomials in non-commuting variables. For example, given the basis $\{e_i\}$ for V we have that $e_1 \otimes e_2 \otimes e_1$ maps to $x_1 x_2 x_1$, and $e_1 \otimes e_2 + e_1 \otimes e_3 \otimes e_2$ maps to $x_1 x_2 + x_1 x_3 x_2$. □

The nice thing about the tensor algebra is that it gives us a basis free way to work with the free algebra, that is a way that is independent of the choice of generators. As it is there is no commutativity imposed on the product in TV, we can impose some commutativity condition by taking quotients.

Definition 3.2.4 — Quotients of the Tensor Algebra Let V be a vector space over $\Bbbk$. We define following quotients:

- $SV := TV/\langle v \otimes w - w \otimes v\rangle$, the **symmetric algebra**; and
- $\Lambda V := TV/\langle v \otimes w + w \otimes v\rangle$, the **exterior algebra**.

If $\mathfrak{g} = V$ is a Lie algebra then we may define the quotient $\mathcal{U}(\mathfrak{g}) := TV/\langle v \otimes w - w \otimes v - [v, w]\rangle$, the **universal enveloping algebra**.

The idea is that for SV we impose that[2] $v \otimes w = w \otimes v$, which makes SV isomorphic to $\Bbbk[x_1, \dots, x_n]$ for $n = \dim V$. For ΛV we impose that $v \otimes w = -w \otimes v$ (usually the product here is written as $v \wedge w$). Finally, for $\mathcal{U}(\mathfrak{g})$ we impose that the bracket, $[v, w]$ is exactly the commutator $v \otimes w - w \otimes v$. This last case is nice because it allows us to treat the abstract bracket as if it were a commutator.

[2] identifying elements with their equivalence class

Note that the tensor algebra, as well as the quotients SV and ΛV, are graded algebras, meaning that they have decompositions as direct sums:

$$SV = \bigoplus_{n=0}^{\infty} S^n V, \qquad \text{and} \qquad \Lambda V = \bigoplus_{n=0}^{\infty} \Lambda^n V. \tag{3.2.5}$$

Here $S^n V$ ($\Lambda^n V$) is the nth (anti)symmetric tensor power of V, that is, it's $V^{\otimes n}$ modulo the relation that factors (anti)commute. Note that $S^n V$ is isomorphic to the subalgebra of $\Bbbk[x_1, \dots, x_n]$ consisting of homogeneous polynomials of degree n.

Four

Jacobson's Density Theorem

4.1 Semisimple Representations

Recall that a module is semisimple if it is a direct sum of simple modules, and a simple module is one with no nontrivial submodules.

Example 4.1.1 Let V be an n-dimensional simple A-module. Then $\operatorname{End} V$ is an A-module as well, with A acting by left matrix multiplication (after fixing some basis so that elements of $\operatorname{End} V$ can be identified with matrices and then identifying elements of A acting on $\operatorname{End} V$ with the corresponding linear operator on V). With this construction $\operatorname{End} V$ is semisimple, in particular

$$\operatorname{End} V \cong \underbrace{V \oplus \cdots \oplus V}_{n \text{ terms}} =: nV. \tag{4.1.2}$$

This isomorphism is given by fixing some basis, $\{v_1, \dots, v_n\} \subseteq V$, and then defining a linear map $\operatorname{End} V \to nV$ by $\varphi \mapsto (\varphi(v_1), \dots, \varphi(v_n))$. Viewing v_i as column matrices $\varphi(v_i)$ is simply the ith column of the matrix corresponding to φ in this basis.

In this example $\operatorname{End} V$ ends up being a direct sum of a single simple module. In the general semisimple case any simple module can appear in the decomposition. If we restrict ourselves to finite dimensions then we can get a pretty good handle on which simple modules appear in such a decomposition. In particular, any finite-dimensional semisimple module, V, may be decomposed as

$$V = \bigoplus_{i \in I} m_i V_i \tag{4.1.3}$$

with $m_i \in \mathbb{Z}_{\geq 0}$ and V_i running over all finite dimensional simple modules. We call m_i the **multiplicity** of V_i in V. Note that since this decomposition is unique up to the order of the terms.

Lemma 4.1.4 Let V be a finite dimensional semisimple A-module, with decomposition

$$V = \bigoplus_{i \in I} m_i V_i \tag{4.1.5}$$

with $m_i \in \mathbb{Z}_{\geq 0}$ and V_i simple. Then the multiplicity, m_i, is given by

$$m_i = \dim(\mathrm{Hom}_A(V_i, V)). \tag{4.1.6}$$

Proof. We make use of the fact that[a]

$$\mathrm{Hom}_A(V_i, V' \oplus V'') \cong \mathrm{Hom}_A(V_i, V') \oplus \mathrm{Hom}_A(V_i, V''). \tag{4.1.7}$$

This extends to all finite direct sums.
Note that $\mathrm{Hom}_A(V_i, V)$ is an $(A, \Bbbk)$-bimodule with the left action $(a.\varphi)(v) = \varphi(a\,.\,v)$ and right action $(\varphi\,.\,\lambda)(v) = \lambda\varphi(v)$. Further, V_i is a right $\Bbbk$-module with the action $v\,.\,\lambda = \lambda v = (\lambda 1_A)\,.\,v$. Thus, $\mathrm{Hom}_A(V_i, V) \otimes_{\Bbbk} V_i$ is a left A-module.
We can define a map

$$\begin{aligned} \psi \colon \bigoplus_{i\in I} \mathrm{Hom}_A(V_i, V) \otimes_{\Bbbk} V_i &\to V \\ \bigoplus_{i\in I} \varphi_i \otimes v_i &\mapsto \sum_i \varphi_i(v_i). \end{aligned} \tag{4.1.8}$$

This is an A-module isomorphism:

$$\begin{aligned} \psi\left(a\,.\,\textstyle\bigoplus_{i\in I}\varphi_i \otimes v_i\right) &= \psi\left(\textstyle\bigoplus_{i\in I}\varphi_i \otimes (a\,.\,v_i)\right) && (4.1.9)\\ &= \textstyle\sum_{i\in I}\varphi_i(a\,.\,v_i) && (4.1.10)\\ &= \textstyle\sum_{i\in I} a\,.\,\varphi_i(v_i) && (4.1.11)\\ &= a\,.\,\textstyle\sum_{i\in I}\varphi_i(v_i) && (4.1.12)\\ &= a\,.\,\psi\left(\textstyle\bigoplus_{i\in I}\varphi_i \otimes v_i\right). && (4.1.13) \end{aligned}$$

Linearity is clear from the definition. It remains only to show that this map is invertible. By linearity it is sufficient to show that the map

$$\begin{aligned} \mathrm{Hom}(V_i, V) \otimes V_i &\to V && (4.1.14)\\ \varphi_i \otimes v_i &\mapsto \varphi_i(v_i) && (4.1.15) \end{aligned}$$

is an isomorphism. Since V_i is simple Schur's lemma tells us that this map is either zero or surjective. It is clearly not zero, since we can simply choose some vector v_i and some nonzero map φ_i on which $\varphi_i(v_i) \neq 0$. Thus, this map is surjective. A surjective linear map between finite dimensional modules is an isomorphism. Hence, the map in Equation (4.1.8) is an isomorphism.
We then have

$$\begin{aligned} \dim V &= \dim\left(\textstyle\bigoplus_{i\in I}\mathrm{Hom}_A(V_i, V)\right) && (4.1.16)\\ &= \textstyle\sum_{i\in I}\dim(\mathrm{Hom}_A(V_i, V))\dim(V_i) && (4.1.17) \end{aligned}$$

and

$$\begin{aligned} \dim V &= \dim\left(\textstyle\bigoplus_{i\in I} m_i V_i\right) && (4.1.18)\\ &= \textstyle\sum_{i\in I} m_i \dim(V_i). && (4.1.19) \end{aligned}$$

Since these are finite sums and this must hold for arbitrary semisimple modules V, including the case where $V = V_i$ is actually simple, we must have that

$$m_i = \dim(\mathrm{Hom}_A(V_i)).$$ □

[a] $\mathrm{Hom}(V_i, -)$ is right adjoint (to $- \otimes_A V_i$) and as such preserves colimits

The decomposition into simple submodules also puts restrictions on the non-simple submodules that we can have. First, every submodules of a semisimple module must itself be semisimple, meaning it has its own decomposition into simple modules. Further, the simple modules that can appear in the decomposition of the submodule are only the ones that appear in the decomposition of the module. Finally, the multiplicity with which these simple modules appear in the submodule must be at most the multiplicity with which they appear in the original module. That is, the only way to form a submodule of a semisimple module is to take some subset of the simple modules that appear in the decomposition and take their direct sum.

Proposition 4.1.20 Let V be a semisimple finite-dimensional A-module with decomposition

$$V = \bigoplus_{i=1}^{m} n_i V_i \tag{4.1.21}$$

with the V_i pairwise-nonisomorhpic simple A-modules. Let $W \subseteq V$ be a submodule. Then

$$W = \sum_{i=1}^{m} r_i V_i \tag{4.1.22}$$

with $0 \leq r_i \leq n_i$ for all i, and the inclusion $\varphi : W \hookrightarrow V$ decomposes as

$$\varphi = \bigoplus_{i=1}^{m} \varphi_i \tag{4.1.23}$$

where $\varphi_i : r_i V_i \to n_i V_i$ are maps given by $\varphi_i(v_1, \dots, v_{r_i}) = (v_1, \dots, v_{r_i}) . X_i$ where $X_i \in \mathrm{Mat}_{r_i \times n_i}(\Bbbk)$ acts on the row vector by right matrix multiplication and has rank r_i.

Proof. The proof is by induction on $n = \sum_{i=1}^{m} n_i$. For the base case we just have that V is simple, and so its only submodules are the zero module (the empty direct sum) or V itself, in which case the statement clearly holds. Now suppose that this is the case when $\sum_i n_i = n-1$. Fix some submodule, $W \subseteq V$. If $W = 0$ then we're done, so suppose $W \neq 0$. Fix some simple submodule, $P \subseteq W$. Such a P exists as a consequence of Lemma 4.1.29. By Schur's lemma P must be isomorphic to V_i for some i, and the inclusion

$\varphi|_P : P \to V$ factors through $n_i V_i$ by

$$P \xrightarrow{\cong} V_i \hookrightarrow n_i V_i \hookrightarrow V. \tag{4.1.24}$$

Identifying P with V_i this map is given by

$$v \mapsto (vq_1, \dots, vq_{n_i}) \tag{4.1.25}$$

with $q_i \in \Bbbk$ not all zero.
The group $G_i = \mathrm{GL}_{n_i}(\Bbbk)$ acts on $n_i V_i$ by right matrix multiplication. We can also act trivially on $n_j V_j$ for $j \neq i$. Then G_i acts on V. This gives an action of G_i on the set of submodules of V, and this action preserves the property that we're trying to establish, that under the action of $g_i \in G_i$ the matrix X_i goes to $X_i g_i$ while the matrices X_j ($j \neq i$) are left unchanged. Taking $g_i \in G_i$ such that $(1_1, \dots, q_{n_i})g_i = (1, 0, \dots, 0)$, which is always possible as g_i is invertible, we have that Wg_i contains the first summand, V_i, of $n_i V_i$. Thus, $Wg_i \cong V_i \oplus W'$ where

$$W' \subseteq n_1 V_1 \oplus \cdots \oplus (n_i - 1)V_i \oplus \cdots \oplus n_m V_m \tag{4.1.26}$$

is the kernel of the projection of Wg_i onto the first summand V_i. The inductive hypothesis then holds for this subspace, and so it has a decomposition

$$W' \cong \bigoplus_{j=1}^{m} r_j' V_i \tag{4.1.27}$$

with $0 \le r_i' \le n_i - 1$ and $0 \le r_j \le n_j$ for $j \neq i$, and so taking

$$W \cong V_i \oplus W \cong \bigoplus_{j=1}^{m} r_j V_i \tag{4.1.28}$$

with $r_i = r_i' + 1$ and $r_j = r_j'$ we get the desired result. □

Lemma 4.1.29 Any nonzero finite dimensional A-module contains a simple submodule.

Proof. The proof is by induction on dimension. Let V be a finite dimensional nonzero A-module. We start with $\dim V = 1$. Then V is itself simple, and we are done. Suppose then that all A-modules of dimension at most k contain a simple submodule. Consider the case when $\dim V = k+1$. If V is simple we are done. If V is not simple then it contains a proper submodule, W. Since W is a *proper* submodule it has dimension less than $k + 1$, and thus the induction hypothesis holds. Thus, W has a simple submodule, which is then also a simple submodule of V. Then, by induction, the statement holds for all finite dimensional A-modules. □

Remark 4.1.30 We assumed that $\Bbbk$ was algebraically closed in the use of Schur's lemma above. However, this is not required for a modified result to hold. If we replace $\mathrm{Mat}_{r_i \times n_i}(\Bbbk)$ with $\mathrm{Mat}_{r_i \times n_i}(D_i)$ where $D_i = \mathrm{End}_A(V_i)$ then the result holds for any field $\Bbbk$. The D_i are division algebras (algebras in which division by any nonzero element is defined). When $\Bbbk$ *is* algebraically closed Schur's lemma applies and tells us that the maps $V_i \to V_i$ are just scalar multiplication, allowing us to identify D_i with $\Bbbk$ to get the result as stated above.

Corollary 4.1.31 Let V be a finite dimensional simple A-module. Given two subsets $\{x_1, \dots, x_n\}, \{y_1, \dots, y_n\} \subseteq V$ with the first being linearly independent there exists some $a \in A$ such that $a \,.\, x_i = y_i$.

Proof. The proof is by contradiction, so suppose that this is not the case. Then $W = \{(a \,.\, x_1, \dots, a \,.\, x_n) \mid a \in A\}$ must be a proper submodule of nV, that is there is some element of V we can pick for one of the y_i such that we cannot reach $(y_1, \dots, y_n)$ by the action of a. Then since V is simple we know that $W = rV$ for some $r < n$, a strict inequality since we have a *proper* submodule. By Proposition 4.1.20 we know that there is some $X \in \mathrm{Mat}_{r\times n}(()\Bbbk)$ and some $u_1, \dots, u_r \in V$ such that

$$(u_1, \dots, u_r) \,.\, X = (x_1, \dots, x_n). \tag{4.1.32}$$

To achieve this result we've just considered the $a = 1$ case to get $(x_1, \dots, x_n) \in W = rV$. Since $r < n$ we know that there is some $(z_1, \dots, z_n) \in \Bbbk^n \setminus \{0\}$ such that $X \,.\, (z_1, \dots, z_n)^\top = 0$, because X only has rank r. Thus, we can consider

$$\begin{aligned} 0 &= (u_1, \dots, u_r) \,.\, X \,.\, (z_1, \dots, z_n)^\top && (4.1.33)\\ &= (x_1, \dots, x_n) \cdot (z_1, \dots, z_n)^\top && (4.1.34)\\ &= \sum_{i=1}^{n} z_i x_i. && (4.1.35) \end{aligned}$$

Since the x_i are linearly independent this means that $z_i = 0$, a contradiction. □

4.2 Density Theorem

We're now ready to start working towards a result known as the density theorem. This result says that a certain class of algebras are basically just direct sums of matrix algebras. We have to prove some technical results first though.

Theorem 4.2.1. Let V be a finite dimensional A-module.

1. If V is simple then the associated algebra morphism $r\colon A \to \mathrm{End}\, V$ is surjective.

2. If $V = \oplus_{i=1}^m V_i$ with the V_i pairwise nonisomorphic finite dimensional simple A-modules then

$$r = \bigoplus_{i=1}^{m} r_i : A \to \bigoplus_{i=1}^{m} \operatorname{End} V_i \tag{4.2.2}$$

is surjective.

Proof. 1. Fix some basis, $\{v_1, \dots, v_n\} \subseteq V$, and let $w_i = \varphi(v_i)$ for some $\varphi \in \operatorname{End} V$. Then by Corollary 4.1.31 there exists some $a \in A$ such that $a \,.\, v_i = w_i$, and thus $r(a) = \varphi$, so r is surjective.

2. Let B_i be the image of A in $\operatorname{End} V_i$. Notice that $\operatorname{End} V_i \cong d_i V_i$ where $d_i = \dim V_i$. Let B be the image of A in $\bigoplus_i \operatorname{End} V_i$. Then $B \cong \bigoplus_i B_i \cong \bigoplus_i d_i V_i$, and the first part tells us that $B_i = \operatorname{End} V_i$ by surjectivity of each representation map, and thus $B \cong \bigoplus \operatorname{End} V_i$, so r is surjective.

□

The next result considers what happens when we have an algebra that is a direct sum of matrix algebras. Before the proof however we need the following definition.

Definition 4.2.3 — Dual Module Let V be a left A-module. Then the **dual module** is $V^* = \operatorname{Hom}_{\Bbbk}(V, \Bbbk)$ with the action defined by $(f.a)(v) = f(a.v)$ for all $f \in V^*$, $a \in A$, and $v \in V$.

Theorem 4.2.4. Let $\Bbbk$ be a field which is not necessarily algebraically closed. Let A be the $\Bbbk$-algebra given by

$$A = \bigoplus_{i=1}^{r} \operatorname{Mat}_{d_i}(\Bbbk) \tag{4.2.5}$$

for some $d_i \in \mathbb{N}$. Then

1. the simple A-modules are $\Bbbk^{d_i}$ with $(X_1, \dots, X_r)$ acting by matrix multiplication by X_i; and

2. any finite dimensional A-module is semisimple.

Proof. 1. Let $v, w \in \Bbbk^{d_i}$ be such that $v \neq 0$. Then there exists some linear map sending v to w, and hence some matrix $X \in \operatorname{Mat}_{d_i}(\Bbbk)$ such that $Xv = w$. Thus, $V_i = \Bbbk^{d_i}$ must be simple since any nonzero subspace containing v and not w cannot be a submodule.

2. Let W be a finite dimensional left A-module. Consider its dual, W^*, which we can think of as a left A^{op}-module. The algebra A^{op} is given

by

$$A^{\mathrm{op}} = \bigoplus_i \mathrm{Mat}_{d_i}(\Bbbk)^\top \cong \bigoplus_i \mathrm{Mat}_{d_i}(\Bbbk) \tag{4.2.6}$$

and we identify $a \in A$ with $a^\top \in A^{\mathrm{op}}$ where $(X_1, \dots, X_r)^\top = (X_1^\top, \dots, X_r^\top)$. Really nothing is going on here since we're considering square matrices so taking the transpose changes individual elements but doesn't change the set of all matrices under consideration.

What this lets us do is interpret W^* as an A-module with $a.f = f.a^\top$. We can fix a basis $\{f_1, \dots, f_n\} \subseteq W^*$, and then define a surjection

$$\varphi : nA \twoheadrightarrow W^* \tag{4.2.7}$$
$$a_1 \oplus \cdots \oplus a_n \mapsto a_1 . f_1 + \cdots + a_n . f_n. \tag{4.2.8}$$

This is a surjection by Theorem 4.2.1. We can consider the dual map, $\varphi^* : W \hookrightarrow (nA)^* \cong nA$, which will be an injection. Further, $W \cong \operatorname{im}\varphi^* \subseteq nA$ is a submodule of the semisimple module nA (where $a.(b_1 \oplus \cdots \oplus b_n) = ab_1 \oplus \cdots \oplus ab_n$) and we can apply Proposition 4.1.20 to conclude that W is semisimple.

□

What we have just shown is that matrix algebras, and their direct sums, have particularly nice properties. We understand their simple modules well, they're just $\Bbbk^d$ with d appearing as the number of rows of some matrix, and all finite dimensional modules are semisimple, so all are just some direct sum $\bigoplus_i \Bbbk^{d_i}$. The logical next question is when is a given algebra, A, isomorphic to some direct sum of matrix algebras? It turns out that there's a simple subspace we can consider that vanishes only when A is a direct sum of matrix algebras.

Definition 4.2.9 — Radical Let A be an algebra. We call

$$\operatorname{Rad} A = \{a \in A \mid a \text{ acts as zero on any simple } A\text{-module}\} \subseteq A \tag{4.2.10}$$

the **radical** of A.

Definition 4.2.11 — Nilpotent Ideal Let A be an algebra. We call $a \in A$ a **nilpotent element** if there exists some $k \in \mathbb{N}$ such that $a^k = 0$. A **nilpotent ideal** is an ideal in which all elements are nilpotent.

Proposition 4.2.12

1. $\operatorname{Rad} A$ is a two-sided ideal.
2. If A is finite dimensional then any nilpotent two-sided ideal is contained in $\operatorname{Rad} A$.

3. $\operatorname{Rad} A$ is the largest two-sided nilpotent ideal.

Proof. 1. We first show that $\operatorname{Rad} A$ is a subspace. Let V be a simple A-module. Then if $a, b \in \operatorname{Rad} A$ we have

$$(a+b) \,.\, v = a \,.\, v + b \,.\, v = 0 + 0 = 0 \tag{4.2.13}$$

for all $v \in V$, and thus $\operatorname{Rad} A$ is closed under addition. If $\lambda \in \Bbbk$ we also have

$$(\lambda a) \,.\, v = \lambda(a \,.\, v) = \lambda 0 = 0, \tag{4.2.14}$$

and so $\operatorname{Rad} A$ is closed under scalar multiplication. Thus, $\operatorname{Rad} A$ is a subspace of A.

Let $a \in \operatorname{Rad} A$ and $b \in A$. Then we know that if V is a simple A-module $a \,.\, v = 0$ for all $v \in V$. We therefore have

$$(ab).v = a.(b.v) = 0, \quad \text{and} \quad (ba).v = b.(a.v) = b.0 = 0 \tag{4.2.15}$$

since $b \,.\, v \in V$ so a acts on it by zero, and b acts linearly so it sends 0 to 0. Thus, $ab, ba \in \operatorname{Rad} A$, so $\operatorname{Rad} A$ is a two-sided ideal.

2. Let V be a simple A-module and I a nilpotent ideal. Fix some nonzero $v \in V$. Then $I \,.\, v \subseteq V$ is a submodule. By simplicity of V there are two possibilities
   - $I \,.\, v = V$, and since $v \in V$ there must be some $x \in I$ such that $x \,.\, v = v$, but then we cannot have that $x^k = 0$ for any $k \in \mathbb{N}$ as we must have $x^k \,.\, v = v$, so we can't have $I \,.\, v = V$ if I is nilpotent;
   - $I \,.\, v = 0$, in which case every element of I acts as zero on any element of V, and so $I \subseteq \operatorname{Rad} A$.

3. Let

$$0 = A_0 \subseteq A_1 \subseteq A_1 \subseteq \cdots \subseteq A_n = A \tag{4.2.16}$$

be a filtration of the regular representation of A such that A_{i+1}/A_i is simple. Such a filtration exists by Lemma 4.2.19.

Let $x \in \operatorname{Rad} A$, then x acts on the simple A-module A_{i+1}/A_i by zero, and so x must map any element of A_{i+1} to some element of A_i, since that will then be sent to zero in the quotient. Thus x^n acts as zero on all of $A_n = A$, and so $\operatorname{Rad} A$ is nilpotent. By the previous part we also know that $\operatorname{Rad} A$ contains any nilpotent two-sided ideal, and so $\operatorname{Rad} A$ is the largest two-sided nilpotent ideal (ordered by inclusion). □

Definition 4.2.17 — Filtration Let V be an A-module. A finite **filtration** of

V is a sequence of submodules

$$0 = V_0 \subseteq V_1 \subseteq \cdots \subseteq V_n = V. \tag{4.2.18}$$

Lemma 4.2.19 Let V be a finite dimensional A-module. Then there is a filtration

$$0 = V_0 \subseteq V_1 \subseteq \cdots \subseteq V_n = V \tag{4.2.20}$$

for which V_{i+1}/V_i is a simple A-module for all i.

Proof. We induct on $\dim V$. If $\dim V = 0$ then we have the filtration $0 = V_0 = V$ and we are done. Suppose the result holds for all dimensions less than $\dim V$. If V is simple then we have the filtration $0 = V_0 \subseteq V_1 = V$ and $V/0 \cong V$ is simple, so we're done. Suppose then that V is not simple, and pick some nontrivial submodule $V_1 \subsetneq V$. Take the module $U = V/V_1$. Since $V_1 \neq 0$ we know that $\dim(V/V_1) < \dim V$, and so by the induction hypothesis there is a filtration

$$0 = U_0 \subseteq U_1 \subseteq \cdots \subseteq U_{n-1} = U \tag{4.2.21}$$

such that U_{i+1}/U_i is simple. Let $\pi : V \twoheadrightarrow V/V_1$ be the canonical projection. For $i \geq 2$ define $V_i = \pi^{-1}(U_i)$ to be the preimage of U_i under this projection. Then we have the filtration

$$0 = V_0 \subseteq V_1 \subseteq \cdots \subseteq V_n = V. \tag{4.2.22}$$

Note that here we've used the fact that the preimage under a module morphism of a submodule of the codomain is a submodule of the domain, which can be seen as follows: take $v \in V_i$ and we have some $u \in U_i$ such that $\pi(v) = u$, then

$$a . u = a . \pi(v) = \pi(a . v) \in U_i \tag{4.2.23}$$

which shows that $a . v \in V_i$ also, so V_i is closed under the action of A, and the preimage of a subspace is again a subspace.
All we have to do now is show that the given filtration has the desired property. To see that this is indeed the case consider $V_{i+1}/V_i = \pi^{-1}(U_{i+1})/\pi^{-1}(U_i) \cong \pi^{-1}(U_{i+1}/U_i)$ which shows that V_{i+1}/V_i is the preimage of a simple module, and must therefore be simple itself, if it wasn't then the image of any nontrivial submodule of V_{i+1}/V_i would provide a nontrivial submodule of U_{i+1}/U_i. $\square$

The following result gives us a handle on the number of simple A-modules in the finite dimensional case. It also shows that given any algebra we can always quotient by the radical to get something isomorphic to a direct sum of endomorphism spaces, which is isomorphic to a direct sum of matrix algebras. In this way the radical consists of the elements which obstruct our attempt to understand A as being formed from matrix algebras.

Notation 4.2.24 We write $\mathrm{Irr}(A)$ for the set of isomorphism classes of simple A-modules. We further assume that each isomorphism class has some canonical choice of representative, which we'll call V_i, so we can take $\mathrm{Irr}(A) = \{V_i\}$. We assume that sums over the index i in V_i run over all simple A-modules.

Theorem 4.2.25. Any finite dimensional algebra, A, has only finitely many simple A-modules, V_i, (up to isomorphism) and

$$\sum_i (\dim V_i)^2 \leq \dim A. \tag{4.2.26}$$

Further,

$$A/\operatorname{Rad} A \cong \bigoplus_i \operatorname{End} V_i. \tag{4.2.27}$$

Proof. Let V be a simple A-module and take some $v \in V$ with $v \neq 0$. Then $A \,.\, v \neq 0$ since $1 \in A$ so $v \in A \,.\, v$. Thus, by simplicity we must have that $A \,.\, v = V$. Further, V is finite dimensional since A is finite dimensional, and if we could construct infinitely many linearly independent elements by acting on v with elements of A those infinitely many elements of A would be linearly independent in A, a contradiction.
Now let $\{V_i\} = \mathrm{Irr}(A)$ be the set of simple A-modules. Then by Theorem 4.2.1 we have a surjection

$$\bigoplus_i \rho_i : A \twoheadrightarrow \operatorname{End} V_i. \tag{4.2.28}$$

Thus, we have

$$\dim\left(\bigoplus\nolimits_i \operatorname{End} V_i\right) = \sum\nolimits_i \dim(\operatorname{End} V_i) \tag{4.2.29}$$

$$= \sum\nolimits_i (\dim V_i)^2 \tag{4.2.30}$$

where we've used the fact that the dimension of a direct sum is the sum of the dimensions, and $\operatorname{End} V$ has dimension $(\dim V)^2$, which can be seen by fixing a basis for V and considering elements of $\operatorname{End} V$ as $(\dim V) \times (\dim V)$ matrices. Finally, since the above map is a surjection the dimension is bounded by $\dim A$, and thus we have

$$\sum_i (\dim V_i)^2 \leq \dim A \tag{4.2.31}$$

as claimed.
We have that

$$\ker\left(\bigoplus\nolimits_i \rho_i\right) = \operatorname{Rad} A \tag{4.2.32}$$

since by definition elements of this kernel are sent to the zero map when when they act on each simple module, V_i, and this is exactly the definition of said elements being in $\operatorname{Rad} A$. Thus, by the first isomorphism theorem we have that

$$A/\ker\left(\bigoplus_i \rho_i\right) = A/\operatorname{Rad} A \cong \bigoplus_i \operatorname{End} V_i. \qquad \square$$

We now give a definition of a semisimple algebra. Note that several equivalent definitions are in use, and some of these are covered in Proposition 4.2.34.

Definition 4.2.33 — Semisimple Algebra A finite dimensional algebra, A, is **semisimple** if $\operatorname{Rad} A = 0$.

Proposition 4.2.34 Let A be a finite dimensional algebra, then the following are equivalent:

(I) A is semisimple, that is $\operatorname{Rad} A = 0$;

(II) $\dim A = \sum_i (\dim V_i)^2$ where V_i runs over all simple A-modules;

(III) $A \cong \bigoplus_i \operatorname{Mat}_{d_i}(\Bbbk)$ for some $d_i \in \mathbb{N}$;

(IV) Any finite dimensional A-module is semisimple. In particular, the regular representation is semisimple.

Proof. Step 1: (I) $\implies$ (II)
We have that

$$A/\operatorname{Rad} A \cong \bigoplus_i \operatorname{End} V_i \tag{4.2.35}$$

and taking dimensions we have

$$\dim(A/\operatorname{Rad} A) = \sum_i (\dim V_i)^2. \tag{4.2.36}$$

If A is semisimple then $\operatorname{Rad} A = 0$ and this reduces to the equality

$$\dim A = \sum_i (\dim V_i)^2. \tag{4.2.37}$$

Step 2: (I) $\implies$ (III)
By Theorem 4.2.25 we know that

$$A/\operatorname{Rad} A \cong \bigoplus_i \operatorname{End} V_i \tag{4.2.38}$$

and if A is semisimple then $\operatorname{Rad} A = 0$ so this reduces to

$$A \cong \bigoplus_i \operatorname{End} V_i. \tag{4.2.39}$$

Fixing some basis for V_i we may identify elements of End V_i with matrices in $\mathrm{Mat}_{d_i}(\Bbbk)$ where $d_i = \dim V_i$. Thus, we have

$$A \cong \bigoplus_i \mathrm{Mat}_{d_i}(\Bbbk). \tag{4.2.40}$$

STEP 3: (III) $\implies$ (IV)
By the second part of Theorem 4.2.4 we have that any finite dimensional A-module is semisimple.

STEP 4: (IV) $\implies$ (I)
Consider the regular representation of A which decomposes as

$$A \cong \bigoplus_i n_i V_i \tag{4.2.41}$$

with V_i simple and $n_i \in \mathbb{Z}_{\geq 0}$. Take some $x \in \operatorname{Rad} A$, then by definition x acts as zero on each V_i submodule, and so acts as zero on all of A, in particular $x \,.\, 1 = 0$. In the regular representation the action of x is just multiplication, so $x \,.\, 1 = x1 = x$, thus we must have $x = 0$, and hence $\operatorname{Rad} A = 0$. □

One question that we may ask is how many simple A-modules are there (up to isomorphism)? Of course, if we can find the decomposition $A \cong \bigoplus_i \operatorname{End} V_i$ then we have answered the question, but we can often answer the question much faster with the following result definition and result.

Definition 4.2.42 — Centre Let A be an algebra. The **centre** of A, denoted $Z(A)$, is the subalgebra

$$Z(A) := \{a \in A \mid ab = ba \forall b \in A\}. \tag{4.2.43}$$

That is, the centre is the subspace consisting of all elements of A that commute with all other elements of A. This is clearly a subspace since if $a, a' \in Z(A)$ then $(a + \lambda a')b = ab + \lambda a'b = ba + \lambda ba' = b(a + \lambda a')$ for all $b \in A$ and $\lambda \in \Bbbk$. This is in fact a subalgebra since if $a, a' \in Z(A)$ then $aa'b = aba' = aa'b$ so $aa' \in Z(A)$.

Lemma 4.2.44 Let A be a finite dimensional semisimple algebra. Then

$$|\mathrm{Irr}(A)| = \dim Z(A). \tag{4.2.45}$$

Proof. First note that if A_1 and A_2 are algebras then

$$Z(A_1 \oplus A_2) = Z(A_1) \oplus Z(A_2), \tag{4.2.46}$$

since if $(a_1, a_2) \in Z(A_1 \oplus A_2)$ then we have

$$(a_1, a_2)(b_1, b_2) = (b_1, b_2)(a_1, a_2) \tag{4.2.47}$$

for all $b_1, b_2 \in A_1 \oplus A_2$, and evaluating the left hand side gives $(a_1 b_1, a_2 b_2)$

and the right hand side gives (b_1a_1, b_2a_2), so this equality holds if and only if $a_ib_i = b_ia_i$ for all $b_i \in A_i$, in other words, if $a_i \in Z(A_i)$ and thus if and only if $(a_1, a_2) \in Z(A_1) \oplus Z(A_2)$.
Since A is semisimple we know that $\operatorname{Rad} A = 0$, and thus

$$A/\operatorname{Rad} A = A/0 \cong A \cong \bigoplus_i \operatorname{End} V_i \tag{4.2.48}$$

by Theorem 4.2.25. Thus, we have

$$Z(A) = \bigoplus_i Z(\operatorname{End}(V_i)). \tag{4.2.49}$$

Further, since V_i is a simple module we know by Schur's lemma (Proposition 2.5.5) that if an element commutes with all other elements then said element is just scalar multiplication, and further any multiplication by a scalar gives such a map, so

$$Z(\operatorname{End} V_i) \cong \Bbbk. \tag{4.2.50}$$

Combining these two results we have

$$Z(A) \cong \bigoplus_i \Bbbk = |\operatorname{Irr} A|\Bbbk \tag{4.2.51}$$

and so

$$\dim Z(A) = |\operatorname{Irr} A| \tag{4.2.52}$$

where we've used the fact that the sum is indexed by simple A-modules, so has exactly as many terms as there are simple A-modules, and of course, $\dim \Bbbk = 1$. □

Note that if A is not semisimple then this result no longer holds, since $A/\operatorname{Rad} A \ncong A$. However, given a simple A-module, V, we know that all elements of $\operatorname{Rad} A$ act on V by zero, and thus there is a corresponding $(A/\operatorname{Rad} A)$-module V', which has the same underlying space, but now elements of $A/\operatorname{Rad} A$ act by $[a] . v = a . v$ for any representative a of this equivalence class. This gives a well-defined action precisely because elements of $\operatorname{Rad} A$ act by zero, so if a' is some other representative then $a - a' \in \operatorname{Rad} A$ and thus $0 = (a - a') . v = a . v - a' . v$ and thus $a . v = a' . v$ as required.

In fact, more generally if I is an ideal of A and V is an A-module on which all elements of I act as zero then A/I acts on V by $[a].v = a.v$. This can be quite useful when we define algebras via a quotient, first construct an A-module, V, then show that the ideal $I \subseteq A$ acts as zero on V, then we automatically get an (A/I)-module structure for V.

Five

Character Theory

In this chapter we study character theory. The general idea being that for finite dimensional representations we can identify elements of A with linear maps $V \to V$ which we can identify with matrices. We can then take the trace of these matrices, which is a nice thing to do because the trace is basis independent, despite the identification of elements and matrices requiring us to pick a basis. We can then learn a surprising amount just looking at these traces, which we call characters.

5.1 Definitions

Definition 5.1.1 — Character Let A be an algebra and V a finite dimensional A-module with the corresponding algebra homomorphism $\rho : A \to \operatorname{End} V$. Then the **character** of V is the map

$$\chi_V : A \to \Bbbk \tag{5.1.2}$$
$$a \mapsto \chi_V(a) = \operatorname{tr}_V \rho(a) \tag{5.1.3}$$

Note that we write tr_V to denote the trace of matrices corresponding to elements of $\operatorname{End} V$ after fixing some basis. We do this because later we'll want to take characters over different modules, and it's helpful to be able to distinguish which space the matrices we're taking the trace of act on. When there's no chance of confusion we'll drop the subscript V.

Definition 5.1.4 Let A be an algebra with subalgebras $B, C \subseteq A$. Then we denote by $[B, C]$ the subspace

$$[B, C] = \operatorname{span}\{[b, c] \mid b \in B \text{ and } c \in C\} \tag{5.1.5}$$

where $[b, c] = bc - cb$.

Note that for any A-module, V, with corresponding character χ_V, we have

$[A,A] \subseteq \ker \chi_V$, since

$$\begin{aligned}\chi_V([a,b]) &= \mathrm{tr}(\rho([a,b])) &(5.1.6)\\ &= \mathrm{tr}(\rho(a)\rho(b) - \rho(b)\rho(a)) &(5.1.7)\\ &= \mathrm{tr}(\rho(a)\rho(b)) - \mathrm{tr}(\rho(b)\rho(a)) &(5.1.8)\\ &= \mathrm{tr}(\rho(a)\rho(b)) - \mathrm{tr}(\rho(a)\rho(b)) &(5.1.9)\\ &= 0, &(5.1.10)\end{aligned}$$

having used the cyclic property of the trace. Thus $[a,b] \in \ker \chi_V$ for all $a, b \in A$, and since the kernel is a subspace any linear combination of commutators will also vanish under χ_V, showing that $[A,A] \subseteq \ker \chi_V$.

This tells us that the character also gives a well-defined map

$$\tilde{\chi}_V : A/[A,A] \to \Bbbk \tag{5.1.11}$$

defined by

$$\tilde{\chi}_V([a]) = \chi_V(a) = \mathrm{tr}_V(\rho(a)). \tag{5.1.12}$$

In fact, it will prove more useful to define the character to be such a map. This allows us to view the character as an element of the dual space

$$\tilde{\chi}_V \in (A/[A,A])^* = \hom_{\Bbbk}(A/[A,A], \Bbbk). \tag{5.1.13}$$

We will do this, and do not distinguish between χ_V and $\tilde{\chi}_V$ in the notation.

This is a useful thing to do because now the characters live in a vector space, and that lets us do linear-algebra-things to them, like look for a basis of this space.

Theorem 5.1.14. Let A be a finite dimensional algebra. The characters of distinct finite-dimensional simple A-modules are linearly independent in $(A/[A,A])^*$. Further, if A is finite dimensional and semisimple then the characters of simple A-modules provide a basis for $(A/[A,A])^*$.

Proof. STEP 1: LINEAR INDEPENDENCE

Let that A be a finite dimensional (not necessarily semisimple) algebra. Then there is a finite number, n, of simple A-modules, V_i for $i = 1, \dots, n$, with corresponding algebra homomorphisms $\rho_i : A \to \mathrm{End}\, V_i$. Then by the density theorem we have a surjection

$$\rho_1 \oplus \cdots \oplus \rho_n : A \twoheadrightarrow \mathrm{End}\, V_1 \oplus \cdots \oplus \mathrm{End}\, V_n. \tag{5.1.15}$$

Suppose that

$$\sum_i \lambda_i \chi_{V_i} = 0 \tag{5.1.16}$$

with $\lambda_i \in \Bbbk$. If $a \in A$ we must therefore have

$$\sum_i \lambda_i \chi_{V_i}(a) = 0. \tag{5.1.17}$$

Now take some arbitrary $M \in \operatorname{End} V_1 \oplus \cdots \oplus \operatorname{End} V_n$, which we view as a matrix by fixing some basis, which fixes a basis for each V_i. We can then identify that $M = M_1 \oplus \cdots \oplus M_n$, where each $M_i \in \operatorname{End} V_i$ is viewed as a matrix through the corresponding fixed basis. We can then consider the sum

$$\sum_i \lambda_i \operatorname{tr}_{V_i} M_i \tag{5.1.18}$$

where the λ_i are the same coefficients as before. By surjectivity of $\rho_1 \oplus \cdots \oplus \rho_n$ we know that there is some $a \in A$ such that $M = (\rho_1 \oplus \cdots \oplus \rho_n)(a)$, and thus $M_i = \rho_i(a)$. This then gives that the sum above is

$$\sum_i \lambda_i \operatorname{tr}_{V_i}(\rho_i(a)) = \sum_i \lambda_i \chi_{V_i}(a) = 0. \tag{5.1.19}$$

Now, we are free to choose M, and hence M_i, such that $\operatorname{tr}_{V_i} M_i$ takes on any value in $\Bbbk$, which means that the only way this equation can hold for an arbitrary choice of M is if $\lambda_i = 0$ for all $i = 1, \dots, n$. Thus, the χ_{V_i} are linearly independent.

Step 2: Basis

Now suppose that A is a finite dimensional semisimple algebra. We have shown that the characters, χ_{V_i}, corresponding to simple A-modules, are linearly independent elements of $(A/[A,A])^*$. We now show that they are also a spanning set of $(A/[A,A])^*$.

Since A is semisimple we have that

$$A \cong \bigoplus_{i=1}^{n} \operatorname{Mat}_{d_i}(\Bbbk) \tag{5.1.20}$$

where $d_i = \dim V_i$. We have the following well known fact about derived subalgebras of Lie algebras (Lemma 5.1.29):

$$[\operatorname{Mat}_d(\Bbbk), \operatorname{Mat}_d(\Bbbk)] = [\mathfrak{gl}_{d_i}(\Bbbk), \mathfrak{gl}_{d_i}(\Bbbk)] = (\mathfrak{gl}_{d_i}(\Bbbk))' = \mathfrak{sl}_d(\Bbbk). \tag{5.1.21}$$

The Lie algebra $\mathfrak{sl}_d(\Bbbk)$ consists precisely of the $d \times d$ matrices over $\Bbbk$ with zero trace. Further, for algebra B and C, we have

$$[B \oplus C, B \oplus C] = [B,B] \oplus [C,C], \tag{5.1.22}$$

which follows immediately by linearity. Thus, we have

$$[A,A] \cong \bigoplus_{i=1}^{n} \mathfrak{sl}_{d_i}(\Bbbk). \tag{5.1.23}$$

It then follows that

$$A/[A,A] \cong \left(\bigoplus_{i=1}^{n} \mathrm{Mat}_{d_i}(\Bbbk)\right) / \left(\bigoplus_{i=1}^{n} \mathfrak{sl}_{d_i}(\Bbbk)\right) \tag{5.1.24}$$
$$= \left(\bigoplus_{i=1}^{n} \mathfrak{gl}_{d_i}(\Bbbk)\right) / \left(\bigoplus_{i=1}^{n} \mathfrak{sl}_{d_i}(\Bbbk)\right) \tag{5.1.25}$$
$$\cong \bigoplus_{i=1}^{n} \mathfrak{gl}_{d_i}(\Bbbk)/\mathfrak{sl}_{d_i}(\Bbbk) \tag{5.1.26}$$
$$\cong \bigoplus_{i=1}^{n} \Bbbk \tag{5.1.27}$$
$$= \Bbbk^n. \tag{5.1.28}$$

This shows that we have n-linearly independent elements, χ_{V_i}, and $\dim(A/[A,A]) = n$, so these linearly independent elements are actually a basis. □

Lemma 5.1.29 The derived subalgebra of $\mathfrak{gl}_n(\Bbbk)$ is $\mathfrak{sl}_n(\Bbbk)$.

Proof. First note that $\mathfrak{gl}_n(\Bbbk) = \mathrm{Mat}_n(\Bbbk)$ is the (Lie algebra) of $n \times n$ matrices with entries in $\Bbbk$. The elementary matrices, E_{ij}, form a basis of $\mathfrak{gl}_n(\Bbbk)$. Note that E_{ij}, for $i, j = 1, \dots, n$, are matrices which are zero everywhere except in row i and column j, where they have a 1. So, it is sufficient to show that the commutator of any two elementary matrices is in $\mathfrak{sl}_n(\Bbbk)$, and then any linear span of such commutators will be in $\mathfrak{sl}_n(\Bbbk)$. To do this first note that

$$E_{ij}E_{kl} = \delta_{jk}E_{il}. \tag{5.1.30}$$

Then we have

$$[E_{ij}, E_{kl}] = E_{ij}E_{kl} - E_{kl}E_{ij} \tag{5.1.31}$$
$$= \delta_{jk}E_{il} - \delta_{li}E_{kj}. \tag{5.1.32}$$

Now we consider cases:

1. if $i \neq l$ and $j \neq k$ we get 0;
2. if $l \neq i$ and $j = k$ we get E_{il};
3. if $l = i$ and $j \neq k$ we get $-E_{kj}$;
4. if $i = l$ and $j = k$ we get $E_{ii} - E_{jj}$.

We see that in each case the matrix we get is traceless, specifically in the last case if $i \neq j$ then the diagonal contains a 1 and a -1, and if $i = j$ then we have zero, and the second and third case have zero on the diagonal since $i \neq l$ and $k \neq j$ in these two cases. Thus, each matrix we get from $[E_{ij}, E_{kl}]$ is an element of $\mathfrak{sl}_n(\Bbbk)$. □

Lemma 5.1.33 Characters are invariant under isomorphism.

Proof. Let V and W be isomorphic finite dimensional A-modules. Then V and W are related by an isomorphism, $V \to W$, but fixing bases for both we can view this isomorphism as a basis change, and the character is independent of basis choice. □

Lemma 5.1.34 Let V be a finite dimensional A-module, and let $W \subseteq V$ be a submodule. Then

$$\chi_V = \chi_W + \chi_{V/W}. \tag{5.1.35}$$

Proof. Fix a basis for W and extend this to a basis of V. This can be done since $V = W \oplus V/W$ as vector spaces. Then any linear map $\varphi \colon V \to V$ such that $\varphi(W) \subseteq W$ decomposes into a linear map $W \to W$ and a linear map $V/W \to V/W$. Since W is a submodule $\rho(a)$ is exactly such a linear map for all $a \in A$, and thus

$$\operatorname{tr}_V \rho(a) = \operatorname{tr}_W \rho(a) + \operatorname{tr}_{V/W} \rho(a), \tag{5.1.36}$$

and so

$$\chi_V = \chi_W + \chi_{V/W}.$$ □

5.2 Jordan–Hölder and Krull–Schmidt Theorems

We can now prove two standard results about filtrations using character theory.

Theorem 5.2.1 — Jordan–Hölder. Let V be a finite dimensional A-module with filtrations

$$0 = V_0 \subseteq V_1 \subseteq \cdots \subseteq V_n = V \tag{5.2.2}$$

and

$$0 = V_0' \subseteq V_1' \subseteq \cdots \subseteq V_m' = V \tag{5.2.3}$$

such that $W_i = V_i/V_{i-1}$ and $W_i{}' = V_i'/V_{i-1}'$ are simple. Then

1. $n = m$; and
2. There exists some $\sigma \in S_n$ such that $W_i \cong W'_{\sigma(i)}$, that is, the two series give rise to the same simple A-modules (up to isomorphism), but possibly in different orders.

Proof.

This proof holds only in characteristic 0. The result does hold in general though, and can be proven in positive characteristic by induction on the dimension of V. The problem in characteristic p is that the coefficients only end up being determined mod p.

Consider the character χ_V. Using the first series and Lemma 5.1.34 we know that

$$\chi_V = \bigoplus_{i=1}^{n} \chi_{W_i}, \tag{5.2.4}$$

and using the second series we know that

$$\chi_V = \bigoplus_{i=1}^{m} \chi_{W_i'}. \tag{5.2.5}$$

Since the characters of the simple A-modules form a basis of $(A/[A,A])^*$ any decomposition such as the above must be unique, and thus we have $n = m$ and there is some permutation, $\sigma \in S_n$ such that $\chi_{W_i} = \chi_{W'_{\sigma(i)}}$, and thus $W_i \cong W'_{\sigma(i)}$. □

Definition 5.2.6 — Jordan–Hölder Series Given a finite dimensional A-module, V, admitting a filtration

$$0 = V_0 \subseteq V_1 \subseteq \cdots \subseteq V_n = V \tag{5.2.7}$$

such that V_i/V_{i-1} are simple we call n the **length** of V, and the set of simple modules $\{V_i/V_{i-1}\}$ is called the **Jordan–Hölder series** of V.

Note that by the Jordan–Hölder theorem the length and Jordan–Hölder series are well-defined, being independent of the choice of filtration, so long as the quotient of successive modules is simple.

The following result holds for finite length modules. Note that finite length is a strictly weaker condition than finite dimension, since finite dimension guarantees the existence of

Theorem 5.2.8 — Krull–Schmidt. Every finite length A-module, V, is a direct sum of indecomposable modules. Further, this decomposition is unique up to isomorphism and permutation of the summands.

Proof. Step 1: Existence
Let V be a finite length A-module. We may suppose that $V = V_1 \oplus V_2$ with V_i A-modules, and without loss of generality we assume that V_1 cannot be written as a sum of indecomposables. Then we must be able to decompose

V_1 again. Continuing on we see that this gives rise to an infinite length filtration, contradicting the assumption that V has finite length.

STEP 2: UNIQUENESS
We make use of Lemma 5.2.14. Using this result take two decompositions into indecomposables

$$V = V_1 \oplus \cdots \oplus V_m = V_1' \oplus \cdots \oplus V_m'. \tag{5.2.9}$$

We will prove that $V_k \cong V_k'$ for some k. Let

$$i_k : V_k \hookrightarrow V, \qquad \text{and} \qquad i_k' : V_k' \hookrightarrow V \tag{5.2.10}$$

be the natural inclusions, and

$$p_k : V \twoheadrightarrow V_k, \quad \text{and} \quad p_k' : V \twoheadrightarrow V_k' \tag{5.2.11}$$

be the natural projections. Then we have the map

$$\theta_k : p_1 \circ i_k' \circ p_k' \circ i_1 : V_1 \to V_1, \tag{5.2.12}$$

which is a composite of module morphisms, so is itself a module morphism. We also have that $\sum_k \theta_k = \mathrm{id}_V$, since summing over all k the image of $i_k' \circ p_k'$ in the middle runs over all of V, We know that id_V is not nilpotent, so by the contrapositive of Lemma 5.2.14 we know that at least one of the θ_ks must be an isomorphism. Without loss of generality we assume that θ_1 is an isomorphism. Then we have that

$$V_1 = \mathrm{im}(p_1' \circ i_1) \oplus \ker(p_1 \circ i_1'), \tag{5.2.13}$$

but V_1 is indecomposable, so $p_1' \circ i_1 : V_1 \to V_1'$ must be an isomorphism. We may then consider $V_2 \oplus \cdots V_m \cong V_2' \oplus \cdots V_m$, and by the same logic we may take $V_2 \cong V_2'$. Repeating this eventually terminates after m applications. $\square$

Lemma 5.2.14 Let W be a finite dimensional indecomposable A-module. Then

1. any module morphism $\theta : W \to W$ is either an isomorphism or nilpotent;
2. if $\theta_i : W \to W$ for $i = 1, \dots, n$ is a set of nilpotent module morphisms then $\theta = \sum_i \theta_i$ is also a nilpotent module morphism.

Proof. We work over an algebraically closed field, thus W splits into a sum of generalised eigenspaces. These are submodules of W. Thus, θ can have only one eigenvalue, call it λ. If $\lambda = 0$ then θ is nilpotent, and if $\lambda \neq 0$ then θ is an isomorphism.
We prove that the sum of nilpotents is nilpotent by induction on n. For the

base case, $n = 1$, we clearly have that $\theta = \theta_1$ is nilpotent. Suppose then that the hypothesis holds up to n summands, and that at n summands θ is not nilpotent. Then θ must be an isomorphism, and thus its inverse exists, and we have $\mathrm{id}_W = \theta\theta^{-1} = \theta^{-1}\sum_{i=1}^{n}\theta_i = \sum_{i=1}^{n}\theta^{-1}\theta_i$. Since the morphisms $\theta^{-1}\theta_i$ are not isomorphisms they are nilpotent, and thus $\mathrm{id}_W - \theta^{-1}\theta_n = \theta^{-1}\theta_1 + \cdots + \theta^{-1}\theta_{n-1}$ is an isomorphism, but it's also a sum of $n-1$ nilpotents, so it should be nilpotent, a contradiction. Thus by induction any such sum of nilpotents is itself nilpotent. ☐

5.3 Tensor Products

Let A and B be $\Bbbk$-algebras. Then $A \otimes_{\Bbbk} B$ is also a $\Bbbk$-algebra when equipped with the product

$$(a \otimes b)(a' \otimes b') = aa' \otimes bb' \tag{5.3.1}$$

for $a, a' \in A$ and $b, b' \in B$.

Theorem 5.3.2. Let A and B be $\Bbbk$-algebras. Let V be a simple finite dimensional A-module, and W a simple finite dimensional B-module. Then $V \otimes_{\Bbbk} W$ is a simple $(A \otimes_{\Bbbk} B)$-module. Further, any finite dimensional simple $(A \otimes_{\Bbbk} B)$-module is of this form with V and W unique.

Proof. By the density theorem we have surjections $A \twoheadrightarrow \operatorname{End} V$ and $B \twoheadrightarrow \operatorname{End} W$. Thus, we have a surjection

$$A \otimes B \twoheadrightarrow \operatorname{End} V \otimes \operatorname{End} W \cong \operatorname{End}(V \otimes W). \tag{5.3.3}$$

Thus, $V \otimes W$ must be simple, as any submodules would only arise as submodules of V and W.
Now suppose that U is a simple $(A \otimes B)$-module, and let A' and B' denote the images of A and B in $\operatorname{End} U$. Then A' and B' are finite dimensional, and we can assume without loss of generality that A and B are also finite dimensional. By Claim 5.3.6 we have that

$$\operatorname{Rad}(A \otimes B) = \operatorname{Rad}(A) \otimes B + A \otimes \operatorname{Rad}(B) \tag{5.3.4}$$

and thus, we have

$$(A \otimes B)/\operatorname{Rad}(A \otimes B) = A/\operatorname{Rad}(A) \otimes B/\operatorname{Rad}(B). \tag{5.3.5}$$

Since all of the algebras in question are matrix algebras the assertion follows. ☐

Claim 5.3.6 For $\Bbbk$-algebras A and B we have

$$\operatorname{Rad}(A \otimes B) = \operatorname{Rad}(A) \otimes B + A \otimes \operatorname{Rad}(B). \tag{5.3.7}$$

Proof. Consider the simple module $V \otimes W$, where V is a simple A-module and W is a simple B-module. We know that if $a \otimes b \in \operatorname{Rad}(A \otimes B)$ then $a \otimes b$ acts as zero on $V \otimes W$. We also know that if $v \otimes w \in V \otimes W$ then $a \otimes b$ acts as

$$(a \otimes b) \,.\, (v \otimes w) = (a \,.\, v) \otimes (b \,.\, w). \tag{5.3.8}$$

If this is to vanish then it must be that either $a \,.\, v = 0$ or $b \,.\, w = 0$. Thus, $a \in \operatorname{Rad} A$ or $b \in \operatorname{Rad} B$, and so $a \otimes b \in \operatorname{Rad} A \otimes B + A \otimes \operatorname{Rad} B$. Conversely, clearly any element of this set acts trivially on $V \otimes W$, and thus we have containment the other way. $\square$

Part II

Group Representations

Six

Representation Theory of Finite Groups

Throughout this chapter G will be a finite group.

In this chapter we will look at representations of finite groups. We have already developed much of the required theory because group representations, $\rho\colon G \to \mathrm{GL}(V)$, are in one-to-one correspondence with $\Bbbk G$-modules. Note that we write G-module and $\mathrm{Hom}_G(V, W)$ for $\Bbbk G$-module and $\mathrm{Hom}_{\Bbbk G}(V, W)$.

6.1 Maschke's Theorem

Theorem 6.1.1 — Maschke. Let char $\Bbbk$ be coprime to $|G|$. Then

1. $\Bbbk G$ is semisimple;
2. $\Bbbk G \cong \bigoplus_i \operatorname{End} V_i$ with the isomorphism given on the basis by $g \mapsto \bigoplus_i \rho_i(g)$ where $\rho_i\colon G \to \mathrm{GL}(V_i)$ are the irreducible representations of G.

Proof. We know that semisimplicity of $\Bbbk G$ implies that $\Bbbk G$ decomposes as in the second point (Proposition 4.2.34), so we need only show that $\Bbbk G$ is semisimple.

To prove that $\Bbbk G$ is semisimple it is sufficient to prove that given a G-module, V, and a G-submodule $W \subseteq V$ there is some G-submodule, W' such that $V = W \oplus W'$. This will show that any finite-dimensional $\Bbbk G$-module is semisimple, and hence that $\Bbbk G$ is semisimple by Proposition 4.2.34.

Given a G-module, V, and a G-submodule, W, we always have *as vector spaces* some $\bar{W} \subseteq V$ such that $V = W \oplus \bar{W}$. We will construct from $\bar{W}$ a G-submodule W' such that $V = W \oplus W'$.

Let $p\colon V \twoheadrightarrow W$ be projection onto the subspace W. That is, $p|_W = \mathrm{id}_W$ and $p|_{\bar{W}} = 0$. We may define

$$P = \frac{1}{|G|} \sum_{g\in G} \rho(g) p \rho(g)^{-1} \tag{6.1.2}$$

where $\rho\colon G \to \mathrm{GL}(V)$ is our representation map. Now consider $W' = \ker P$. We claim that W' is a submodule and $V = W \oplus W'$.

To verify these we need to show that $G \,.\, W' \subseteq W'$ and that P is projection onto W. Suppose that $w \in W'$, that is $Pw = 0$. Then for $h \in G$ we have

$$\begin{aligned} P(h \,.\, w) &= P\rho(h)w && (6.1.3)\\ &= \frac{1}{|G|}\sum_{g\in G}\rho(g)p\rho(g)^{-1}\rho(h)w && (6.1.4)\\ &= \frac{1}{|G|}\sum_{g\in G}\rho(g)p\rho(g^{-1}h)w && (6.1.5)\\ &= \frac{1}{|G|}\sum_{g'\in G}\rho(hg')p\rho(g'^{-1})w && (6.1.6)\\ &= \rho(h)\frac{1}{|G|}\sum_{g'\in G}\rho(g')p\rho(g'^{-1})w && (6.1.7)\\ &= \rho(h)Pw && (6.1.8)\\ &= 0 && (6.1.9) \end{aligned}$$

where we've reparametrised the sum using $g'^{-1} = g^{-1}h$, so $g' = h^{-1}g$ and $g = hg'$. This is a common trick when dealing with sums over group elements like this one. We have successfully shown that $h \,.\, w \in \ker P$ if $w \in \ker P$, and thus $h \,.\, w \in W'$.
We can now verify that P is a projection onto W. For this we have to show that $P|_W = \mathrm{id}_W$, and $P(V) \subseteq W$, which combined imply that $P^2 = P$. For the first if $w \in W$ consider

$$Pw = \frac{1}{|G|}\sum_{g\in G}\rho(g)p\rho(g)^{-1}w. \tag{6.1.10}$$

Since W is a submodule we know that $\rho(g)^{-1}w \in W$, then since p is a projection onto W we know that $p\rho(g)^{-1}w = \rho(g)^{-1}w$, and thus $\rho(g)p\rho(g)^{-1}w = \rho(g)\rho(g)^{-1}w = w$. So, the sum reduces to

$$Pw = \frac{1}{|G|}\sum_{g\in G} w = \frac{|G|}{|G|}w = w. \tag{6.1.11}$$

Thus, $P|_W = \mathrm{id}_W$ as claimed. Now we can show that $P(V) \subseteq W$. For $v \in V$ consider

$$Pv = \frac{1}{|G|}\sum_{g\in G}\rho(g)p\rho(g)^{-1}v. \tag{6.1.12}$$

By definition V is closed under the action of g, so $\rho(g)^{-1}v \in V$, then by definition $p\rho(g)^{-1}v \in W$, and since W is a submodule $\rho(g)p\rho(g)^{-1}v \in W$ for all $g \in G$. Submodules are closed under taking linear combinations, so $Pv \in W$. Thus, P is a projection onto W, and so we have the decomposition of vector spaces $V = W \oplus W'$, and we've already shown that W' is actually a submodule, so this is a decomposition of G-modules. □

Corollary 6.1.13 We have

$$\Bbbk G \cong \bigoplus_i (\dim V_i) V_i \tag{6.1.14}$$

and

$$|G| = \sum_i (\dim V_i)^2. \tag{6.1.15}$$

Proof. This is simply Maschke's theorem applied to the regular representation, which is just G acting on itself by multiplication, where we've used $|G| = \dim \Bbbk G$. □

The converse of Maschke's theorem holds also.

Proposition 6.1.16 If $\Bbbk G$ is semisimple then $\operatorname{char} \Bbbk$ and $|G|$ are coprime.

Proof. By Maschke's theorem we can write

$$\Bbbk G \cong \bigoplus_{i=1}^{r} \operatorname{End} V_i \tag{6.1.17}$$

where the V_i are simple G-modules and $V_1 = \Bbbk$ is the trivial representation. Then we have

$$\Bbbk G \cong \Bbbk \oplus \bigoplus_{i=2}^{r} \operatorname{End} V_i \cong \Bbbk \oplus \bigoplus_{i=2}^{r} d_i V_i \tag{6.1.18}$$

with $d_i = \dim V_i$. Schur's lemma then tells us that every homomorphism of G-modules $\Bbbk \to \Bbbk G$ is a scalar multiple of some fixed homomorphism $\Lambda : \Bbbk \to \Bbbk G$, and every G-module homomorphism $\Bbbk G \to \Bbbk$ is a scalar multiple of some fixed homomorphism $\varepsilon : \Bbbk G \to \Bbbk$. More symbolically, the hom-spaces $\operatorname{Hom}_{\Bbbk G}(\Bbbk, \Bbbk G)$ and $\operatorname{Hom}_{\Bbbk G}(\Bbbk G, \Bbbk)$ are one-dimensional with bases Λ and ε respectively, so are simply $\Bbbk\Lambda$ and $\Bbbk\varepsilon$. We are free to choose these maps to be such that $\varepsilon(g) = 1$ for all $g \in G$, and $\Lambda(1) = \sum_{g \in G} g$. Then we have

$$\varepsilon(\Lambda(1)) = \varepsilon\left(\textstyle\sum_{g \in G} g\right) = \sum_{g \in G} \varepsilon(g) = \sum_{g \in G} 1 = |G|. \tag{6.1.19}$$

Now, if $|G| = kp$ where $p = \operatorname{char} \Bbbk$ then $|G| = 0$ in $\Bbbk G$ and so this sum says that $\varepsilon \circ \Lambda(1) = 0$, which means that Λ has no left-inverse since $a\varepsilon \circ \Lambda(1) = 0$ for all $a \in \Bbbk$, which rules out all maps $\Bbbk G \to \Bbbk$ (since all are of the form $a\varepsilon$ for some $a \in \Bbbk$) as inverses for Λ, since these would have to give $a\varepsilon \circ \Lambda(1) = 1$. □

Example 6.1.20 Consider $G = \mathbb{Z}/p\mathbb{Z}$, and $\Bbbk$ a field of characteristic p. Clearly, char $\Bbbk = p$ and $|G| = p$ are not coprime.
A consequence of this is that every simple $\mathbb{Z}/p\mathbb{Z}$-module over $\Bbbk$ is trivial. This follows because in a finite group of order p we have that $x^p = 1$, so $x^p - 1$ acts as zero, but over a field of characteristic p we have that $x^p - 1 = (x-1)^p$, and thus $(x-1)^p$ acts as zero, so $x - 1$ acts as 0 (as 0 is the only element of the group which doesn't act as 1 when raised to the power of p), so x must act as 1.

6.2 Group Characters

Definition 6.2.1 — Group Character Let G be a group and $\rho\colon G \to \mathrm{GL}(V)$ a representation on a finite dimensional space, V. Then the **character** of V is the map

$$\chi_V\colon G \to \Bbbk \tag{6.2.2}$$
$$g \mapsto \chi_V(g) = \mathrm{tr}_V(\rho(g)). \tag{6.2.3}$$

Of course, if $\tilde{\chi}_V\colon \Bbbk G \to \Bbbk$ is the character of the corresponding representation of the group algebra $\Bbbk G$ then $\chi_V = \tilde{\chi}_V|_G$, viewing G as a subset of $\Bbbk G$ in the canonical way (i.e., restricting to the canonical basis).

Definition 6.2.4 — Class Function Let G be a group. A **class function** of G is a map $f\colon G \to \Bbbk$ such that $f(g) = f(hgh^{-1})$ for all $g, h \in G$. We write

$$\mathcal{X}(G) = \{f\colon G \to \Bbbk \mid f(g) = f(hgh^{-1}) \forall g, h \in G\} \tag{6.2.5}$$

for the set of all class functions.

That is, class functions are functions which are invariant under conjugation of their argument. Another way of putting this, which explains the name, is that class functions are exactly those functions which are constant on each conjugacy class. Because of this we can identify

$$\mathcal{X}(G) \cong_{\mathsf{Set}} \mathrm{Func}(\mathcal{C}(G), \Bbbk) \tag{6.2.6}$$

where $\mathcal{C}(G)$ is the set of all conjugacy classes and

$$\mathrm{Func}(A, B) = \{f\colon A \to B\} = \mathsf{Set}(A, B). \tag{6.2.7}$$

Actually, under pointwise addition and scalar multiplication $\mathcal{X}(G)$ is a vector space. Further, under mild conditions the irreducible characters provide a basis for this space.

Theorem 6.2.8. If char $\Bbbk$ and $|G|$ are coprime then the irreducible characters, χ_{V_i}, of G form a basis for $\mathcal{X}(G)$.

Proof. From Maschke's theorem we know that $A = \Bbbk G$ is semisimple. We have proven that the irreducible algebra characters χ_{V_i} form a basis for $(A/[A,A])^*$. We then have

$$\begin{aligned}(A/[A,A])^* &= \{f \in \mathrm{Hom}_{\Bbbk}(\Bbbk G, \Bbbk) \mid gh - hg \in \ker f \forall g, h \in G\}\\ &= \{f \in \mathrm{Hom}_{\Bbbk}(\Bbbk G, \Bbbk) \mid f(gh) - f(hg) = 0 \forall g, h \in G\}\\ &= \{f \in \mathrm{Hom}_{\Bbbk}(\Bbbk G, \Bbbk) \mid f(gh) = f(hg) \forall g, h \in G\}\\ &\cong_{\Bbbk\text{-Vect}} \{f \in \mathrm{Func}(G, \Bbbk) \mid f(gh) = f(hg) \forall g, h \in G\}\\ &= \mathcal{X}(G).\end{aligned}$$

□

Corollary 6.2.9 The number of irreducible representations of G is equal to the number of conjugacy classes:

$$|\mathrm{Irr}(G)| = |\mathcal{C}(G)|. \tag{6.2.10}$$

Example 6.2.11 This example makes use of ideas from the representation theory of the symmetric group, something we'll cover in more detail in Chapter 9 Consider the symmetric group, $G = S_n$. Using cycle notation if we write every element as a product of disjoint cycles then two elements are in the same conjugacy class if and only if they have the same cycle type. More concretely, take S_4, then the cycle type of $(1\,2\,3\,4)$ is (4), the cycle type of $(1\,2)(3\,4)$ is $(2,2)$, the cycle type of $(1\,2\,3)$ is $(3,1)$ (note that $(1\,2\,3) = (1\,2\,3)(4)$, and we have to include all elements of $\{1,2,3,4\}$). So, for example, $(1\,2\,3)$ and $(2\,3\,4)$ are conjugate, and so are $(1\,2)(3\,4)$ and $(1\,3)(2\,4)$.
We can identify conjugacy classes with cycle types, and we can identify cycle types with partitions of n. A **partition** of n being a tuple $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_k)$ such that $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_k \ge 0$ $\lambda_1 + \lambda_2 + \cdots + \lambda_k = n$. We write $\lambda \vdash n$ to denote that λ is a partition of n.
A common, and useful notation, for partitions is that of **Young diagrams**. Here we take a partition, λ, and write a row of λ_i boxes in the ith row (rows counted from the top down). For example, $(1\,2)(3\,4)$ has cycle type $\lambda = (2,2)$, and the corresponding Young diagram is

$$\lambda = \begin{array}{|c|c|}\hline & \\ \hline & \\ \hline \end{array}. \tag{6.2.12}$$

Similarly, $(1\,2\,3)$ has cycle type $\mu = (3,1)$, and the corresponding Young diagram is

$$\mu = \begin{array}{|c|c|c|}\hline & & \\ \hline & \multicolumn{2}{c}{} \\ \cline{1-1} \end{array}. \tag{6.2.13}$$

So, we have a bijection between

- conjugacy classes of S_n;
- partitions of n;
- Young diagrams with n boxes.

It will turn out that Young diagrams, and the related Young tableaux, come up a lot when we start counting things related to the symmetric group. Later, we will explicitly define the irreducible representation, V_λ, of S_n corresponding to a partition $\lambda \vdash n$.

Note that if char $\Bbbk$ divides $|G|$ then $\Bbbk G$ is not generally semisimple and we typically have $|\mathcal{C}(G)| \geq |\mathrm{Irr}(G)|$.

Corollary 6.2.14 For a field of characteristic 0 two G-modules, V and W, are isomorphic if and only if $\chi_V = \chi_W$.

Proof. Under these conditions $\Bbbk G$ is semisimple, and thus we can decompose both representations as

$$V = \bigoplus_i n_i V_i, \qquad \text{and} \qquad W = \bigoplus_i m_i V_i \tag{6.2.15}$$

where V_i are irreducible representations and $n_i, m_i \in \mathbb{Z}_{\geq 0}$. Then we have

$$\begin{aligned} \chi_V(g) &= \mathrm{tr}_V(\rho_V(g)) && (6.2.16)\\ &= \mathrm{tr}_{\bigoplus_i n_i V_i}(n_i \rho_{V_i}(g)) && (6.2.17)\\ &= \sum_i n_i \,\mathrm{tr}_{V_i}(\rho_{V_i}(g)) && (6.2.18)\\ &= \sum_i n_i \chi_{V_i}(g) && (6.2.19) \end{aligned}$$

and similarly

$$\chi_W(g) = \sum_i m_i \chi_{V_i}(g). \tag{6.2.20}$$

Since the characters are a basis we have equality between these only if $n_i = m_i$, and thus both representations have the same decomposition, so are isomorphic. □

There is an isomorphism of vector spaces $\Bbbk G \cong_{\Bbbk\text{-Vect}} \mathrm{Func}(G, \Bbbk)$ on the basis by identifying g with δ_g for $g \in G$ where

$$\delta_g(h) = \delta_{g,h} = \begin{cases} 1 & g = h \\ 0 & g \neq h \end{cases} \tag{6.2.21}$$

is the **Kronecker delta**.

We can define the **convolution** product, $*$, on $\mathrm{Func}(G, \Bbbk)$ by

$$(\psi * \varphi)(g) = \sum_{h \in G} \psi(h)\varphi(h^{-1}g). \tag{6.2.22}$$

This product makes $\mathrm{Func}(G, \Bbbk)$ an algebra, and extends the above isomorphism to an isomorphism of algebras, $\Bbbk G \cong_{\Bbbk\text{-}\mathsf{Alg}} \mathrm{Func}(G, \Bbbk)$, since

$$(\delta_g * \delta_h)(k) = \sum_{\ell \in G} \delta_g(\ell)\delta_h(\ell^{-1}k) \tag{6.2.23}$$

$$= \sum_{\ell \in G} \delta_{g,\ell}\delta_{h,\ell^{-1}k} \tag{6.2.24}$$

and terms in this sum vanish except for when $g = \ell$ and $h = \ell^{-1}k$, which means that $h = g^{-1}k$, or $k = gh$. So we only get a nonzero output if $k = gh$, which means that this convolution is exactly δ_{gh}, which is of course the same as taking the product of g and h in $\Bbbk G$ then mapping to $\mathrm{Func}(G, \Bbbk)$.

Proposition 6.2.25 Let

$$\boldsymbol{c} = \sum_{g \in c} g \tag{6.2.26}$$

where $c \in \mathcal{C}(G)$ is some conjugacy class. Then $Z(\Bbbk G) = \langle \boldsymbol{c} \mid C \in \mathcal{C}(G) \rangle$ and $Z(\Bbbk G) \cong \mathcal{X}(G)$.

Proof. We first show that for each conjugacy class, c, $\boldsymbol{c}$ is in $Z(\Bbbk G)$. To do so we show that $\boldsymbol{c}$ commutes with all elements of G, so taking $g \in G$ we have

$$\boldsymbol{c}g = \sum_{h \in C} hg \tag{6.2.27}$$

$$= \sum_{h \in C} ghg^{-1}g \tag{6.2.28}$$

$$= \sum_{h \in C} gh \tag{6.2.29}$$

$$= g\sum_{h \in C} h \tag{6.2.30}$$

$$= g\boldsymbol{c}. \tag{6.2.31}$$

Here we've used the fact that conjugation by g is a permutation on c, and thus changing h to ghg^{-1} in the sum doesn't change the sum, it just permutes the terms.
The result follows from Lemma 6.2.32 applied to the special case where $X = G$ with the action given by conjugation, in which case the invariant subspace is exactly the centre of $\Bbbk G$. □

Lemma 6.2.32 Let G be a finite group acting on a finite set, X. The invariant subspace of the free vector space $\Bbbk X$ is spanned by elements of the form

$\boldsymbol{o} = \sum_{x \in o} x$ where o ranges over all orbits of the group action.

Proof. Consider $\boldsymbol{o}$ for some orbit, o, we have

$$g \,.\, \boldsymbol{o} = \sum_{x \in o} g \,.\, o = \boldsymbol{o}. \tag{6.2.33}$$

This follows since acting with g is just a permutation of the orbit, o, and thus the sum is unchanged, it's just a permutation of the terms in the sum. Conversely, suppose that $v = \sum_{x \in X} v_x x$ is invariant under the action of G. Then we have

$$g \,.\, v = \sum_{x \in X} v_x (g \,.\, x) \tag{6.2.34}$$

and by invariance we demand that this is equal to

$$v = \sum_{x \in X} v_x x = \sum_{g^{-1}.x \in X} v_{g^{-1}.x} x, \tag{6.2.35}$$

so we can conclude that $v_x = v_{g^{-1}.x}$ for all $g \in G$, and thus $v_x = v_y$ whenever x and y lie in the same orbit. Hence, v is a linear combination of the elements $\boldsymbol{o}$, and so the $\boldsymbol{o}$ are a basis of the invariant subspace of $\Bbbk X$. □

Example 6.2.36 — Finite Abelian Group Let G be a finite abelian group. Since G is abelian every element of G is in its own conjugacy class, so

$$|\mathrm{Irr}(G)| = |\mathcal{C}(G)| = |G|. \tag{6.2.37}$$

By the structure theorem we know that

$$G \cong \mathbb{Z}_{n_1} \times \dots \times \mathbb{Z}_{n_k} \tag{6.2.38}$$

for some $n_i \in \mathbb{Z}_{\geq 0}$. Since G is abelian Schur's lemma tells us that all representations are one dimensional. Further, these irreducible representations form a group under pointwise multiplication:

$$(\rho_1 \cdot \rho_2)(g) = \rho_1(g)\rho_2(g). \tag{6.2.39}$$

The identity, ε, is the trivial representation, $\varepsilon(g) = 1$. The inverse of ρ is the representation $g \mapsto 1/\rho(g)$.
Each irreducible representation is a map $\rho : G \to \Bbbk^\times \cong \mathrm{GL}(\Bbbk)$. Thus, in this case the representations coincide with the characters.
We call the group $G^\vee \coloneqq (\mathrm{Irr}(G), \cdot)$ the **character group** or **dual group** of G.
Consider now $G = \mathbb{Z}_n$ and $\Bbbk = \mathbb{C}$. Then we have the irreducible representation

$$\rho : \mathbb{Z}_n \to \mathbb{C} \tag{6.2.40}$$
$$m \mapsto \mathrm{e}^{2\pi i m/n} \tag{6.2.41}$$

and $\mathbb{Z}_n^\vee = \{\rho^k \mid k = 1, \dots, n\}$, which clearly gives an isomorphism $\mathbb{Z}_n^\vee \cong \mathbb{Z}_n$. In fact, for any finite abelian group we have $G^\vee \cong G$, but not uniquely. However, we do have a canonical isomorphism $G \cong (G^\vee)^\vee$ given by $g \mapsto (\chi \mapsto \chi(g))$.

6.3 Dual Representations

Definition 6.3.1 — Dual Representation Let $\rho : G \to \mathrm{GL}(V)$ be a representation of a finite group on a finite dimensional vector space. Then the dual space, V^*, gives rise to a representation, $\rho^* : G \to \mathrm{GL}(V^*)$, with the on $f \in V^*$ given by

$$(g \,.\, f)(v) = (\rho^*(g)f)(v) = f(\rho(g^{-1})v) \tag{6.3.2}$$

for all $v \in V$.

For $\Bbbk = \mathbb{C}$ we can further simply this by identifying that $\rho^*(g) = \overline{\rho(g^{-1})}^{\mathsf{T}}$. That is, g acts on V^* by the Hermitian conjugate of the action of g^{-1} on V.

Lemma 6.3.3 We have $\chi_{V^*}(g) = \chi_V(g^{-1})$.

Proof. This follows from a direct calculation:

$$\begin{aligned}
\chi_{V^*}(g) &= \mathrm{tr}_{V^*}(\rho^*(g)) && (6.3.4)\\
&= \mathrm{tr}_V(\rho(g^{-1})) && (6.3.5)\\
&= \chi_V(g^{-1}). && \square
\end{aligned}$$

Note that $\chi_V(g) = \sum_i \lambda_i$ where λ_i are the eigenvalues of $\rho(g)$. We also know that for a finite group we have $\rho(g)^{|G|} = \rho(g^{|G|}) = \rho(e) = I$, and thus the eigenvalues of $\rho(g)$ must be roots of unity. For $\Bbbk = \mathbb{C}$ we have $\chi_{V^*}(g) = \sum_i \lambda_i^{-1} = \overline{\chi_V(g)}$, and thus $V \cong V^*$ as G-modules if and only if $\chi_V(g) \in \mathbb{R}$ for all $g \in G$.

6.4 Tensor Products of Representations

Definition 6.4.1 Let $\rho_V : G \to \mathrm{GL}(V)$ and $\rho_W : G \to \mathrm{GL}(W)$ be representations of G. Then there is a representation

$$\rho_V \otimes \rho_W : G \to \mathrm{GL}(V) \otimes \mathrm{GL}(W) \cong \mathrm{GL}(V \otimes W) \tag{6.4.2}$$

given by

$$(\rho_V \otimes \rho_W)(g) = \rho_V(g) \otimes \rho_W(g). \tag{6.4.3}$$

Note that the character of a tensor product of representations is given by

$$\chi_{V\otimes W}(g) = \chi_V(g)\chi_W(g). \tag{6.4.4}$$

Example 6.4.5 — Schur–Weyl Duality Consider the group $G = \mathrm{GL}(V)$. Then $V^{\otimes n}$ carries a left G-module structure given on simple tensors by

$$g \,.\, (v_{i_1} \otimes \cdots \otimes v_{i_n}) = (g \,.\, v_{i_1} \otimes \cdots \otimes g \,.\, v_{i_n}) \tag{6.4.6}$$

where $g \,.\, v_{i_k}$ is the obvious action of $g \in \mathrm{GL}(V)$ on $v_{i_k} \in V$.
The space $V^{\otimes n}$ also naturally carries a right S_n-module action, given on simply tensors by

$$(v_{i_1} \otimes \cdots \otimes v_{i_n}) \,.\, w = v_{i_{w(1)}} \otimes \cdots \otimes v_{i_{w(n)}}. \tag{6.4.7}$$

That is, $w \in S_n$ just permutes the terms in the tensor product.
These two actions are compatible, in a sense they "commute", since it doesn't matter if we act with $g \in \mathrm{GL}(V)$ on v_{i_k} then rearrange the order of the factors, or if we rearrange the order of the factors then act with g.
The result is that $V^{\otimes n}$ is a $(\mathrm{GL}(V), S_n)$-bimodule.

6.5 Orthogonality of Characters

For this section we will work over $\Bbbk = \mathbb{C}$.

Lemma 6.5.1 Let G be a finite group. Then we may define a bilinear form

$$\langle -, - \rangle \colon \mathcal{X}(G) \times \mathcal{X}(G) \to \mathbb{C} \tag{6.5.2}$$

by

$$\langle \psi, \varphi \rangle := \frac{1}{|G|} \sum_{g \in G} \psi(g)\overline{\varphi(g)}. \tag{6.5.3}$$

This gives a well-defined Hermitian inner product on $\mathcal{X}(G)$.

Proof. Linearity in the first argument and conjugate linearity in the second follow because we defined the inner product as a sum over ψ and $\overline{\varphi}$. Conjugate symmetry is clear from the definition. This is positive definite, for $\psi \neq 0$ we have

$$\langle \psi, \psi \rangle = \frac{1}{|G|} \sum_{g \in G} \psi(g)\overline{\psi(g)} = \frac{1}{|G|} \sum_{g \in G} |\psi(g)|^2 \tag{6.5.4}$$

which is clearly a sum of non-negative terms and so is positive, since at least one term must be nonzero as $\psi \neq 0$. □

Theorem 6.5.5. Let V and W be G-modules, then

$$\langle \chi_V, \chi_W \rangle = \dim(\mathrm{Hom}_G(V, W)). \tag{6.5.6}$$

In particular, if V and W are irreducible then

$$\langle \chi_V, \chi_W \rangle = \begin{cases} 1 & V \cong W, \\ 0 & \text{otherwise.} \end{cases} \tag{6.5.7}$$

Proof. By definition we have

$$\langle \chi_V, \chi_W \rangle = \frac{1}{|G|} \sum_{g \in G} \chi_V(g) \overline{\chi_W(g)} \tag{6.5.8}$$

$$= \frac{1}{|G|} \sum_{g \in G} \chi_V(g) \chi_{W^*}(g) \tag{6.5.9}$$

$$= \frac{1}{|G|} \sum_{g \in G} \chi_{V \otimes W^*}(g) \tag{6.5.10}$$

$$= \frac{1}{|G|} \sum_{g \in G} \mathrm{tr}_{V \otimes W^*}(\rho(g)) \tag{6.5.11}$$

$$= \mathrm{tr}_{V \otimes W^*} \left(\frac{1}{|G|} \sum_{g \in G} \rho(g) \right). \tag{6.5.12}$$

Now, we can identify that

$$P = \frac{1}{|G|} \sum_{g \in G} g \in Z(\mathbb{C}G). \tag{6.5.13}$$

Thus, what we have above is $\mathrm{tr}_{V \otimes W^*}(\rho(P))$.
If $X \in \mathrm{Irr}(G)$ then

$$P|_X = \begin{cases} \mathrm{id}_X & X \cong \mathbb{C}, \\ 0 & \text{otherwise.} \end{cases} \tag{6.5.14}$$

Thus, for any representation, X, $P|_X$ is projection onto X^G, the subspace fixed by the action of G. Hence,

$$\mathrm{tr}_{V \otimes W^*}(\rho(P)) = \dim(\mathrm{Hom}_G(\mathbb{C}, V \otimes W^*)) \tag{6.5.15}$$

$$= \dim(V \otimes W^*)^G \tag{6.5.16}$$

$$= \dim \mathrm{Hom}_G(V, W) \tag{6.5.17}$$

having used the fact that $V \otimes W^* \cong \mathrm{Hom}_{\mathbb{C}}(V, W)$ and $\mathrm{Hom}_{\mathbb{C}}(V, W)^G \cong \mathrm{Hom}_G(V, W)$. □

Corollary 6.5.18 A G-module, V, is simple if and only if $\langle \chi_V, \chi_V \rangle = 1$.

Theorem 6.5.19. Let $g, h \in G$, then

$$\sum_{X \in \operatorname{Irr}(G)} \chi_X(g)\overline{\chi_X(h)} = \begin{cases} |Z_g| & g \text{ conjugate to } h, \\ 0 & \text{otherwise,} \end{cases} \tag{6.5.20}$$

where $Z_g = \{h \in G \mid gh = hg\}$ is the centraliser of g in G.

Proof. We start with the following calculation:

$$\begin{aligned} \sum_{X \in \operatorname{Irr}(G)} \chi_X(g)\overline{\chi_X(h)} &= \sum_{X \in \operatorname{Irr}(G)} \chi_X(g)\chi_{X^*}(h) && (6.5.21) \\ &= \sum_{X \in \operatorname{Irr}(G)} \operatorname{tr}_X(\rho_X(g)) \operatorname{tr}_{X^*}(\rho_{X^*}(h)) && (6.5.22) \\ &= \operatorname{tr}_{\bigoplus_{X \in \operatorname{Irr}(G)} X \otimes X^*}(\rho_X(g) \otimes \rho_{X^*}(h)) && (6.5.23) \\ &= \operatorname{tr}_{\bigoplus_{X \in \operatorname{Irr}(G)} X \otimes X^*}(\rho_X(g) \otimes \rho_X(h^{-1})) && (6.5.24) \\ &= \operatorname{tr}_{\bigoplus_{X \in \operatorname{Irr}(G)} \operatorname{End} X}(x \mapsto \rho(g) x \rho(h^{-1})) && (6.5.25) \\ &= \operatorname{tr}_{\mathbb{C}G}(y \mapsto g y h^{-1}). && (6.5.26) \end{aligned}$$

Here we've used the fact that $X \otimes X^* \cong \operatorname{End} X$, with the isomorphism given by $A \otimes B \mapsto (x \mapsto AxB^*)$. We've then used the fact that

$$\mathbb{C}G \cong \bigoplus_{X \in \operatorname{Irr}(G)} \operatorname{End} X, \tag{6.5.27}$$

since $\mathbb{C}G$ is semisimple.

We now consider cases, the first being when g and h are not conjugate. Suppose that g_i generate G. Then $g g_i h^{-1} \neq g_i$. Thus, the map $y \mapsto g y h^{-1}$, viewed as a matrix, has no on-diagonal elements, and so has vanishing trace.

If instead g and h are conjugate then using the fact that characters are class functions and applying the same logic as above we have

$$\begin{aligned} \sum_{X \in \operatorname{Irr}(G)} \chi_X(g)\overline{\chi_X(h)} &= \sum_{X \in \operatorname{Irr}(G)} \chi_X(g)\overline{\chi_X(g)} && (6.5.28) \\ &= \operatorname{tr}_{\mathbb{C}G}(y \mapsto g y g^{-1}). && (6.5.29) \end{aligned}$$

Further, viewing $y \mapsto g y g^{-1}$ as a matrix we can see that the (y, y) component on the diagonal is 1 precisely if $yg = gy$, and 0 otherwise. That is, there are precisely as many 1s on the diagonal as elements of Z_g, and so $\operatorname{tr}_{\mathbb{C}G}(y \mapsto g y g^{-1}) = |Z_g|$. □

6.5.1 Unitary Representations

Definition 6.5.30 — Unitary Representation Let G be a group and consider a complex vector space, V, equipped with an inner product, $\langle -, - \rangle$. We say that the representation $\rho : G \to \mathrm{GL}(V)$ is **unitary** if $\rho(g)$ is a unitary operator, that is, if

$$\langle \rho(g)v, \rho(g)w \rangle = \langle v, w \rangle \tag{6.5.31}$$

for all $g \in G$ and $v, w \in V$.
Alternatively, a **unitary representation** of G is a homomorphism $\rho : G \to \mathrm{U}(V) \subseteq \mathrm{GL}(V)$ where

$$\mathrm{U}(V) = \{\varphi \in \mathrm{GL}(V) \mid \langle \varphi(v), \varphi(u) \rangle = \langle v, u \rangle\} \tag{6.5.32}$$

is the **unitary group**.

Unitary representations are particularly important in quantum mechanics. The idea is that V is a state space, that is V is the space of possible wave functions, ψ (or $|\psi\rangle$). As is standard we restrict to normalised wavefunctions To each quantity we may want to measure we associate some element of V^*, which we write as $\langle \varphi|$ if the corresponding element of V is $|\varphi\rangle$ (note that there is a canonical isomorphism $V \cong V^*$ because we have the inner product (Riesz representation theorem)). Then the probability of being measured to be in the state $|\varphi\rangle$ when in the state $|\psi\rangle$ is $\langle \varphi|\psi\rangle = \langle \varphi, \psi \rangle$.

A unitary representation, $\rho : G \to \mathrm{U}(V)$, is then interpreted as a symmetry of our system, since the probabilities that we measure are unaffected by this action.

Consider a complex vector space, V. Note that $V \otimes V$ inherits the inner product $\langle u_1 \otimes v_1, u_2 \otimes v_2 \rangle_{V \otimes V} = \langle u_1, v_1 \rangle_V \langle u_2, v_2 \rangle_V$. Without further knowledge of V there are two unitary representations of S_2 on $V \otimes V$, they are $u \otimes v \mapsto v \otimes u$ and $u \otimes v \mapsto -v \otimes u$.

The physical interpretation of this is that if V is the state space of a single particle then $V \otimes V$ is the state space of two identical particles. The two options for S_2 actions then correspond to the two fundamental types of particles. If $u \otimes v \mapsto v \otimes u$ we call the particles **bosons**, and if $u \otimes v \mapsto -v \otimes u$ we call the particles **fermions**.

It turns out that if we're given a finite dimensional complex representation, $\rho : G \to \mathrm{GL}(V)$, of a *finite* group we can always construct a new inner product on V such that this is a unitary representation.

Theorem 6.5.33. Let G be a finite group and V a complex finite-dimensional inner product space with inner product $\langle -, - \rangle$. Let $\rho : G \to \mathrm{GL}(V)$ be a representation of G. Then there exists an inner product, $(-, -)$, on V with respect to which ρ gives a unitary representation.

Proof. We define an inner product on V by

$$(u, v) = \sum_{g \in G} \langle \rho(g)u, \rho(g)v \rangle. \tag{6.5.34}$$

That this is linear follows from the fact that the action of G is linear and

$\langle -, - \rangle$ is linear. The fact that this is positive definite follows because each term in the sum is nonnegative, and for $u \neq v$ we must have $\rho(g)u \neq \rho(g)v$ since $\rho(g)$ is invertible, and thus $\langle \rho(g)u, \rho(g)v \rangle \neq 0$ for $u \neq v$.
That this new inner product is invariant under the action of G follows from a simple calculation:

$$(\rho(g)u, \rho(g)v) = \sum_{h \in G} \langle \rho(h)\rho(g)u, \rho(h)\rho(g)v \rangle \tag{6.5.35}$$

$$= \sum_{h \in G} \langle \rho(hg)u, \rho(hg)v \rangle \tag{6.5.36}$$

$$= \sum_{k \in G} \langle \rho(k)u, \rho(k)v \rangle \tag{6.5.37}$$

$$= (u, v), \tag{6.5.38}$$

where we've reindexed the sum with $k = hg$. $\square$

Another nice property of unitary representations is that since they respect the inner product we get all of the structure of vector spaces that comes with it, including the splitting of short exact sequences, which is just a fancy way of saying that given a vector space, V, with subspace $W \subseteq V$ we always have the orthogonal complement, $W' = \{w' \in V \mid \langle w, w' \rangle = 0 \forall w \in W\}$, which is such that $V \cong W \oplus W'$.

Theorem 6.5.39. Any finite dimensional unitary representation of any group is completely reducible.

Proof. Let V be a finite dimensional unitary representation of a group, G. If V is irreducible we are done. Else, let $W \subseteq V$ be a subrepresentation. Then $W' = \{w' \in V \mid \langle w, w' \rangle = 0\}$ is a subrepresentation also since if $w' \in W'$ then $\rho(g)w' \in W'$ because for any $w \in W'$ we have $\langle w, \rho(g)w' \rangle = \langle \rho(g)\tilde{w}, \rho(g)w' \rangle = \langle \tilde{w}, w' \rangle = 0$ where $\tilde{w} = \rho(g)^{-1}w$ is an element of W because W is closed under the action of g^{-1}. Thus, W and W' are subrepresentations, and as vector spaces we know that $V \cong W \oplus W'$. If either of W or W' is not irreducible we may iterate this process. Eventually this process will terminate as at each iteration the dimensions of the new spaces are lower than the dimension of the original space, and we started with a finite dimensional space. $\square$

Seven

Applications of Characters

7.1 Computing Tensor Products

Suppose we have simple G-modules, V and W. Then the tensor product $V \otimes W$ is again a G-module with the action $g\,.\,(v \otimes w) = (g\,.\,v) \otimes (g\,.\,w)$. Assuming that $\Bbbk G$ is semisimple (so char $\Bbbk$ and $|G|$ are coprime) we can decompose $V \otimes W$ as a direct sum of simple G-modules:

$$V \otimes W = \bigoplus_{U \in \mathrm{Irr}(G)} N_{VW}^{U} U. \tag{7.1.1}$$

Here the coefficients, N_{VW}^{U}, are just the multiplicities of U in this decomposition. These are nonnegative integer values.

We can compute the coefficients, N_{VW}^{U}, using characters. First, note that the character of $V \otimes W$ is $\chi_{V\otimes W} = \chi_V \chi_W$ and using the above decomposition we have

$$\chi_{V\otimes W} = \sum_{U \in \mathrm{Irr}(G)} N_{VW}^{U} \chi_U. \tag{7.1.2}$$

Taking inner products on both sides and using the orthogonality of irreducible characters we have

$$\begin{aligned}
\langle \chi_{V\otimes W}, \chi_U \rangle &= \left\langle \sum_{U' \in \mathrm{Irr}(G)} N_{VW}^{U'} \chi_{U'}, \chi_U \right\rangle && (7.1.3)\\
&= \sum_{U' \in \mathrm{Irr}(G)} N_{VW}^{U'} \langle \chi_{U'}, \chi_U \rangle && (7.1.4)\\
&= \sum_{U' \in \mathrm{Irr}(G)} N_{VW}^{U'} \delta_{U'U} && (7.1.5)\\
&= N_{VW}^{U}. && (7.1.6)
\end{aligned}$$

Here $\delta_{U'U} = 0$ if $U' \ncong U$ and $\delta_{U'U} = 1$ if $U' \cong U$ as G-modules. So, by computing characters we can completely determine the decomposition of $V \otimes W$ into irreducibles, and since this decomposition is unique (up to order and isomorphism) we have completely determined $V \otimes W$.

7.2 Frobenius–Schur Indicator

7.2.1 Bilinear Forms and Dual Spaces

Suppose V is a finite dimensional vector space over $\Bbbk$. Then we know that $V \cong V^*$, but there is no canonical choice of isomorphism. If we fix some isomorphism

$\delta\colon V \to V^*$ then we can define a nondegenerate bilinear form $\langle -, -\rangle_\delta\colon V \times V \to \Bbbk$ by

$$\langle u, v\rangle_\delta = \delta(u)(v). \tag{7.2.1}$$

Conversely, if we have a nondegenerate bilinear form $\langle -, -\rangle\colon V \times V \to \Bbbk$ then we may define an isomorphism $\varphi\colon V \to V^*$ by $u \mapsto \varphi_u$ where $\varphi_u(v) = \langle u, v\rangle$.

However, this doesn't *quite* determine a *unique* isomorphism, because we made the arbitrary choice to define $\varphi_u(v)$ to be $\langle u, v\rangle$, rather than $\langle v, u\rangle$. To fix this we can just assume that $\langle -, -\rangle$ is not just a bilinear form, but either a symmetric or antisymmetric bilinear form. Then φ is uniquely determined for symmetry, or determined up to a sign for antisymmetry. We can always construct a symmetric bilinear form by symmetrising, if $(-, -)$ has no specific symmetry then $\langle u, v\rangle = [(u, v) \pm (v, u)]/2$ is symmetric for $+$ and antisymmetric for $-$.

This analysis also carries over from the theory of vector spaces to a G-module, M. The dual, M^*, is a G-module with the action defined by $g \,.\, f(v) = f(g^{-1} \,.\, v)$. The only subtlety being that to get a left action we use g^{-1} in the action. The only change we need to make is that the nondegenerate (anti)symmetric bilinear form needs to be invariant under the action of G. That is, we should have $\langle g \,.\, u, g \,.\, v\rangle = \langle u, v\rangle$ for all $u, v \in M$. For example, if M is equipped with an inner product then G should act unitarily on M. Thus, if $\langle -, -\rangle$ is a symmetric G-invariant bilinear form on M then we may define an isomorphism $\varphi\colon M \to M^*$ by $u \mapsto \varphi_u$ where $\varphi_u(v) = \langle u, v\rangle$. This is an isomorphism of vector spaces, and it's an isomorphism of G-modules because

$$\varphi(g \,.\, u)(v) = \varphi_{g.u}(v) = \langle g \,.\, u, v\rangle \tag{7.2.2}$$

and

$$(g \,.\, \varphi(u))(v) = (g \,.\, \varphi_u)(v) = \varphi_u(g^{-1} \,.\, v) = \langle u, g^{-1} \,.\, v\rangle. \tag{7.2.3}$$

These are equal, to see this simply act on the arguments of the first with g^{-1}, which doesn't change anything as $\langle -, -\rangle$ is G-invariant, and we get

$$\langle g \,.\, u, v\rangle = \langle g^{-1} \,.\, (g \,.\, u), g^{-1} \,.\, v\rangle = \langle g^{-1}g \,.\, u, g^{-1} \,.\, v\rangle = \langle u, g^{-1} \,.\, v\rangle. \tag{7.2.4}$$

The question then becomes when does a given G-module, M, admit such a nondegenerate (anti)symmetric invariant bilinear form? There are three possibilities, which we classify as follows.

Definition 7.2.5 Let G be a finite group and M a G-module. We say that M is of

(−1) **complex type** if $M^* \ncong M$ as G-modules;

(0) **real type** if M admits a nondegenerate symmetric invariant bilinear form;

(1) **quaternionic type** if M admits a nondegenerate antisymmetric invariant bilinear form.

This naming convention comes from considering $\operatorname{End}_{\mathbb{R}G} M$, for a simple G-module, M, over $\mathbb{R}$. This is the space of linear maps $M \to M$ which commute with the action of *real* linear combinations of group elements. It turns out that $\mathrm{End}_{\mathbb{R}G} M$ is isomorphic to one of $\mathbb{R}$, $\mathbb{C}$, or $\mathbb{H}$, precisely when M is of real, complex, or quaternionic type.

If instead we consider M to be a simple G-module over $\mathrm{Mat}_{2\times 2}(\mathbb{C})$ then $\mathrm{End}_{\mathbb{C}G} M$ is isomorphic to one of $\mathbb{C}$, $\mathbb{C}\times\mathbb{C}$, or $\mathbb{C}$ when M is of real, complex, or quaternionic type. Note that these endomorphism rings over $\mathbb{C}$ are the result of applying the extension of scalars functor, $-\otimes_{\mathbb{R}}\mathbb{C}$, to the endomorphism rings over $\mathbb{R}$.

7.2.2 The Frobenius–Schur Indicator

Definition 7.2.6 — Frobenius–Schur Indicator Let G be a finite group and M a simple G-module. The **Frobenius–Schur indicator** is defined to be

$$\mathrm{FS}(M) := \frac{1}{|G|}\sum_{g\in G}\chi_M(g^2) \tag{7.2.7}$$

where χ_M is the character of M.

Theorem 7.2.8 — Frobenius–Schur. Let G be a finite group. Then the number of involutions in G, that is, the number of elements of order at most 2, is precisely

$$\sum_{M\in\mathrm{Irr}(G)}\dim(M)\,\mathrm{FS}(M). \tag{7.2.9}$$

Proof. Consider some representation, M, and some $A\in\mathrm{End}_{\mathbb{C}G} M$. Let $\lambda_1,\dots,\lambda_n$ be the eigenvaluees of A. We consider S^2M and Λ^2M. These spaces are both formed as quotients of $M\otimes M$ by the ideal generated by $v\otimes w\pm w\otimes v$. Since A acts on $M\otimes M$ as $A\otimes A$ and this action factors through the quotient $A\otimes A$ acts on both of these spaces. We have that

$$\mathrm{tr}_{S^2M}(A\otimes A) = \sum_{1\le i\le j\le n}\lambda_i\lambda_j. \tag{7.2.10}$$

This holds for diagonal matrices when $\otimes$ is the Kronecker product, which is defined by $A\otimes B = (a_{ij}B)$ and so for a diagonal matrix the diagonal is just all products $\lambda_i\lambda_j$. Since the trace is invariant under a basis change this result must also hold for diagonalisable matrices. Finally, it holds for all matrices by continuity because the diagonalisable matrices are dense in all matrices. Similarly, we have

$$\mathrm{tr}_{\Lambda^2M}(A\otimes A) = \sum_{1\le i<j\le n}\lambda_i\lambda_j, \tag{7.2.11}$$

which again, clearly holds for diagonal matrices with the antisymmetrised Kronecker product, since there $\lambda_i^2 = 0$. Thus, we have

$$\mathrm{tr}_{S^2M}(A\otimes A) - \mathrm{tr}_{\Lambda^2M}(A\otimes A) = \sum_{1\le i\le n}\lambda_i^2 = \mathrm{tr}_M A^2. \tag{7.2.12}$$

Thus, for $g \in G$, we can take A to be the corresponding action of g and we get

$$\chi_M(g^2) = \chi_{S^2M}(g) - \chi_{\Lambda^2M}(g). \tag{7.2.13}$$

Note that g is *not* squared on the right because by definition of S^2M and Λ^2M g acts as $g \otimes g$ does on $M \otimes M$, so the squaring is automatic in the definition of the action.
Then summing this result over G and dividing by $|G|$ we get

$$\frac{1}{|G|}\sum_{g\in G}\chi_M(g^2) = \frac{1}{|G|}\sum_{g\in G}\chi_{S^2M}(g) - \frac{1}{|G|}\sum_{g\in G}\chi_{\Lambda^2M}(g). \tag{7.2.14}$$

The left hand side is exactly $\mathrm{FS}(M)$. We have the following vector space decomposition into symmetric and antisymmetric parts:

$$M \otimes M \cong S^2M \oplus \Lambda^2M. \tag{7.2.15}$$

In the finite-dimensional case we also have

$$M \otimes M \cong M \otimes M^* \cong \mathrm{End}_{\mathbb{C}} M. \tag{7.2.16}$$

Thus, we have

$$S^2M \oplus \Lambda^2M \cong \mathrm{End}_{\mathbb{C}} M \tag{7.2.17}$$

as vector spaces. Denote by X^G the fixed points of the action of G on X, that is, $X^G = \{x \in X \mid g\,.\,x = x\}$. This clearly distributes over direct sums, and we have

$$(S^2M)^G \oplus (\Lambda^2M)^G \cong (\mathrm{End}_{\mathbb{C}} M)^G = \mathrm{End}_{\mathbb{C}G} M \tag{7.2.18}$$

where we have identified in the last equality that an endomorphism is fixed under the action of G precisely if it commutes with the action of G. Taking dimensions we have

$$\dim(S^2M)^G + \dim(\Lambda^2M)^G = \dim(\mathrm{End}_{\mathbb{C}G} M). \tag{7.2.19}$$

Since M is simple we know that any G-module endomorphism of M is just scalar multiplication, and thus $\dim(\mathrm{End}_{\mathbb{C}G} M) \le 1$. Since dimensions are integers this leaves us with just two options on the right, either both dimensions are 0, or one is 0 and the other is 1. Thus,

$$\dim(S^2M)^G - \dim(\Lambda^2M)^G \in \{-1, 0, 1\}. \tag{7.2.20}$$

Note that the above quantity is the correct way to generalise the Frobenius–Schur indicator to fields other than $\mathbb{C}$.
Let I be the number of involutions of G. Then

$$I = \sum_{g\in G}[g^2 = 1] \tag{7.2.21}$$

where $[\varphi]$ is the Iverson bracket, $[\varphi] = 1$ if φ is true, and $[\varphi] = 0$ if φ is false. The second orthogonality relation (Theorem 6.5.19) tells us that

$$[g^2 = 1] = \frac{1}{|G|} \sum_{M \in \mathrm{Irr}(G)} \chi_M(g^2)\overline{\chi_M(1)}, \tag{7.2.22}$$

since this result should vanish if g^2 is not conjugate to 1 and should be $|Z_g|$ otherwise. Then we note that g^2 is conjugate to the identity if and only if g^2 *is* the identity. Further, $|Z_g| = |G|$ if $g^2 = 1$. Thus, we have that

$$I = \frac{1}{|G|} \sum_{g \in G} \sum_{M \in \mathrm{Irr}(G)} \chi_M(g^2)\overline{\chi_M(1)}. \tag{7.2.23}$$

Since $\chi_M(1) = \dim M$ this simplifies to

$$I = \frac{1}{|G|} \sum_{g \in G} \sum_{M \in \mathrm{Irr}(G)} \dim(M)\chi_V(g^2) = \sum_{M \in \mathrm{Irr}(G)} \dim(M)\,\mathrm{FS}(M). \tag{7.2.24}$$

□

The proof of the following result is some fairly involved linear algebra, but essentially comes down to the universal property of the tensor/symmetric/exterior product giving a correspondence between bilinear forms and linear maps, and the bilinear forms inherit the (anti)symmetry of the symmetric/exterior product.

Proposition 7.2.25 Let G be a finite group, and M a simple G-module. Then $\mathrm{FS}(M)$ is -1 if M is of complex type, 0 if M is of real type, and 1 if M is of quaternionic type.

Example 7.2.26 This example assumes some knowledge about the basics of representations of S_n, a topic we will cover in Chapter 9, so maybe come back later if you're not familiar with these ideas.
It is a fact that $\mathrm{FS}(M) = 1$ for any simple S_n-module, that is, all S_n-modules are of real type. Simple S_n-modules are indexed by standard tableaux of shape λ with λ a partition of n. Thus, the number of involutions in S_n is precisely

$$\sum_{\lambda \vdash n} |\mathrm{SYT}(\lambda)| \tag{7.2.27}$$

where $\mathrm{SYT}(\lambda)$ is the set of standard Young tableau of shape λ.

7.3 Burnside's Theorem

7.3.1 Statement of Theorem

The next example of an application of character theory is Burnside's theorem, a result in number theory. While Burnside's theorem is relatively easy to state its proof requires some number theory. The result is famous for being one of the first results in group theory which was first proven through representation theory.

Before we state the theorem recall the following definition from group theory.

Definition 7.3.1 — Solvable Group A group, G, is **solvable** if there exists a series of nested normal subgroups

$$\{1\} = G_1 \lhd G_2 \lhd \cdots \lhd G_n = G \tag{7.3.2}$$

such that G_{i+1}/G_i is abelian.

Theorem — Burnside's Theorem. Any group, G, of order $p^a q^b$ with p and q primes and $a, b \in \mathbb{Z}_{\geq 0}$ is solvable.

7.3.2 Algebraic Integers

Definition 7.3.3 — Algebraic Integers A complex number, $z \in \mathbb{C}$, is an

- **algebraic number** if it is a root of some polynomial in $\mathbb{Q}[x]$;
- **algebraic integer** if it is a root of some *monic*[a] polynomial in $\mathbb{Z}[x]$.

We write $\overline{\mathbb{Q}}$ and $\overline{\mathbb{Z}}$ for the sets of algebraic numbers and integers respectively.

[a]Recall that a polynomial is **monic** if the coefficient of the highest degree term is 1.

Example 7.3.4

- $\mathbb{Z} \subseteq \overline{\mathbb{Z}}$: $n \in \mathbb{Z}$ is a root of $x - n$.
- $\mathbb{Q} \cap \overline{\mathbb{Z}} = \mathbb{Z}$: Suppose a/b is rational and reduced, then any rational polynomial with a/b as a root has a factor of $x - a/b$, to get an integer polynomial we have to scale this to $bx - a$. Thus, any integer polynomial with a/b as a root has a factor of $bx - a$, which means it cannot be monic, since any monic polynomial factors as $(x - \alpha_1) \cdots (x - \alpha_m)$ for some roots $\alpha_i \in \mathbb{C}$. Thus, a/b is an algebraic integer only if $b = 1$, in which case $a/b = a$ is an integer.

Lemma 7.3.5 $z \in \mathbb{C}$ is an algebraic number (integer) if and only if it is an eigenvalue of some $n \times n$ matrix over $\mathbb{Q}$ ($\mathbb{Z}$).

Proof. If z is an algebraic number (integer) then it is a root of the monic polynomial

$$p(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 \tag{7.3.6}$$

where $a_i \in \mathbb{Q}$ ($a_i \in \mathbb{Z}$). Note that we are always free to rescale a rational polynomial to be monic. Let

$$A = \begin{pmatrix} 0 & 0 & \dots & 0 & -a_0 \\ 1 & 0 & \dots & 0 & -a_1 \\ 0 & 1 & \dots & 0 & -a_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & -a_{n-1} \end{pmatrix}. \tag{7.3.7}$$

Then the characteristic polynomial of A is

$$-\det(A - xI) = p(x), \tag{7.3.8}$$

and thus z is an eigenvalue of A.
Conversely, suppose that z is an eigenvalue of some $n \times n$ rational (integer) matrix, A. Then z is a root of the characteristic polynomial of A. The characteristic polynomial of a matrix over $\mathbb{Q}$ ($\mathbb{Z}$) is always monic over $\mathbb{Q}$ ($\mathbb{Z}$), and thus z is an algebraic number (integer). □

Proposition 7.3.9

- $\overline{\mathbb{Z}}$ is a ring[a]; and
- $\overline{\mathbb{Q}}$ is a field.

[a] Fun Fact™: The first use of the word "ring" is attributed to Hilbert, who used it describe $\overline{\mathbb{Z}}$, and in particular the way higher powers "loop back around" to be described in terms of lower powers, which can always be done for elements of $\overline{\mathbb{Z}}$ using the polynomial they satisfy to replace higher powers with lower ones.

Proof. STEP 1: $\overline{\mathbb{Q}}$ AND $\overline{\mathbb{Z}}$ ARE RINGS
We will prove that $\overline{\mathbb{Q}}$ is a ring, the proof for $\overline{\mathbb{Z}}$ is analogous. Take $\alpha, \beta \in \overline{\mathbb{Q}}$, then there are matrices $A \in \mathrm{Mat}_m(\mathbb{Q})$ and $B \in \mathrm{Mat}_n(\mathbb{Q})$ such that α and β are eigenvalues of A and B respectively. Let $v \in \mathbb{C}^m$ and $w \in \mathbb{C}^n$ be the corresponding eigenvectors. Consider $A \otimes \mathrm{id}_{\mathbb{C}^n} \pm \mathrm{id}_{\mathbb{C}^m} \otimes B$. A calculation shows that $v \otimes w$ is an eigenvector of this matrix with eigenvalue $\alpha \pm \beta$:

$$\begin{aligned} &(A \otimes \mathrm{id}_{\mathbb{C}^n} \pm \mathrm{id}_{\mathbb{C}^m} \otimes B)(v \otimes w) && (7.3.10)\\ &\qquad = (A \otimes \mathrm{id}_{\mathbb{C}^n})(v \otimes w) \pm (\mathrm{id}_{\mathbb{C}^m} \otimes B)(v \otimes w) && (7.3.11)\\ &\qquad = Av \otimes w \pm v \otimes Bw && (7.3.12)\\ &\qquad = \alpha v \otimes w \pm v \otimes \beta w && (7.3.13)\\ &\qquad = (\alpha \pm \beta)(v \otimes w). && (7.3.14) \end{aligned}$$

Thus, $\alpha \pm \beta$ is an eigenvalue of some $(m+n) \times (m+n)$ matrix over $\mathbb{Q}$, and hence $\alpha \pm \beta \in \overline{\mathbb{Q}}$.
Similarly, $\alpha\beta$ is an eigenvalue of $A \otimes B$ with eigenvector $v \otimes w$:

$$(A \otimes B)(v \otimes w) = Av \otimes Bw = \alpha v \otimes \beta w = (\alpha\beta)(v \otimes w). \tag{7.3.15}$$

Thus, $\alpha\beta$ is an eigenvalue of some $mn \times mn$ matrix over $\mathbb{Q}$, and so $\alpha\beta \in \overline{\mathbb{Q}}$. These results, along with the inherited distributivity law from $\mathbb{C}$, prove that $\overline{\mathbb{Q}}$ is a ring.

STEP 2: $\overline{\mathbb{Q}}$ IS A FIELD
Suppose that $\alpha \in \overline{\mathbb{Q}} \setminus \{0\}$. Then there exits some matrix, $A \in \mathrm{Mat}_{m\times m}(\mathbb{Q})$ such that α is a root of $p(x) = \det(A - xI)$. We can multiply this whole equation by α^m and it follows from properties of determinants that $\alpha^m p(x) = \det(\alpha A - \alpha x I)$. Then, $\alpha^m p(1/\alpha) = \det(\alpha A - I)$, which vanishes when αA has eigenvalue 1, and since αA has the same eigenvalues as A but multiplied by α this shows that some eigenvalue, β, is such that $\alpha\beta = 1$, in other words, $\beta = 1/\alpha$, so $1/\alpha \in \overline{\mathbb{Q}}$. Thus, $\overline{\mathbb{Q}}$ contains multiplicative inverses of nonzero elements, and so is a field (it is clearly commutative and has no zero divisors as it is a subring of $\mathbb{C}$). □

7.3.3 Towards a Proof of Burnside's Theorem

Many quantities that arise in representation theory are naturally algebraic integers. We will use this to restrict the possible values that certain quantities can take, which will be important in our proof of Burnside's theorem.

Lemma 7.3.16 Let G be a finite group and M a finite-dimensional G-module. Then $\chi_M(g)$ is an algebraic integer for every $g \in G$.

Proof. Since G is finite each $g \in G$ has finite order, n, and thus the eigenvalues of $\rho_M(g)$ are nth roots of unity, and so in $\overline{\mathbb{Z}}$ as they satisfy the monic polynomial $x^n - \alpha - 1 = 0$. The trace is the sum of the eigenvalues, and $\overline{\mathbb{Z}}$ is a ring, so is closed under addition, and thus $\chi_M(g) \in \overline{\mathbb{Z}}$. □

Proposition 7.3.17 Let G be a finite group and consider the set of conjugacy classes, $\mathcal{C}(G) = \{[g_1], \dots, [g_n]\}$, with chosen representatives. Define

$$c_i = \sum_{g\in[g_i]} \in \mathbb{C}G, \tag{7.3.18}$$

then for any simple G-module, M, we have $c_i|_M = \lambda_i \mathrm{id}_M$ where

$$\lambda_i = |[g_i]| \frac{\chi_M(g_i)}{\chi_M(1)} \tag{7.3.19}$$

are algebraic integers.

Proof. First note that the c_i are central in $\mathbb{C}G$ since

$$c_i g = \sum_{g'\in[g_i]} g'g = \sum_{g''\in[g_i]} gg'' = g \sum_{g''\in[g_i]} g'' = gc_i \tag{7.3.20}$$

where we've reindexed the sum with $g'' = g^{-}1g'g$, which doesn't change

the value as we're still summing over the whole conjugacy class, just in a different order.
Thus, by Schur's lemma we know that the c_i act as a scalar on any simple G-module. Call this scalar λ_i. Consider the group ring, $\mathbb{Z}G$. This is finitely generated (since G is a finite generating set). Thus, each c_i must satisfy some monic integer polynomial equation, and this carries through to the scalars, λ_i, which shows they are algebraic integers. Viewing c_i as an operator on M we know that $c_i = \lambda_i \mathrm{id}_M$, and we can take the trace of this to get

$$\operatorname{tr}_M c_i = \operatorname{tr}_M(\lambda_i \mathrm{id}_M) = \lambda_i \dim M = \lambda_i \chi_M(1). \tag{7.3.21}$$

We also have

$$\operatorname{tr}_M c_i = \sum_{g\in[g_i]} \operatorname{tr}_M \rho_M(g) = \sum_{g\in[g_i]} \chi_M(g) = |[g_i]|\chi_M(g_i) \tag{7.3.22}$$

since the character is constant on conjugacy classes. Equating these we get the desired result. □

Theorem 7.3.23 — Frobenius Divisibility. Let G be a finite group and M a simple G-module over $\mathbb{C}$. Then $\dim M$ divides $|G|$.

Proof. With notation as in the statement of Proposition 7.3.17 we claim that

$$\sum_i \lambda_i \overline{\chi_M(g_i)} \in \overline{\mathbb{Z}} \tag{7.3.24}$$

where the sum is over all conjugacy classes. Since $\overline{\mathbb{Z}}$ is a ring and Proposition 7.3.17 shows that the λ_i are algebraic integers it is sufficient to show that $\overline{\chi_M(g_i)}$ are algebraic integers. Since G is finite we know that $\rho_M(g_i)^{|G|} = \mathrm{id}_M$, and hence $\chi_M(g_i)$ must be sums of roots of unity, which are algebraic integers, so $\chi_M(g_i)$ are algebraic integers, and hence $\overline{\chi_M(g_i)}$ are algebraic integers, as they are roots of the conjugate polynomial.
From Proposition 7.3.17 we also have

$$\sum_i \lambda_i \overline{\chi_M(g_i)} = \sum_i |[g_i]| \frac{\chi_M(g_i)\overline{\chi_M(g_i)}}{\chi_M(1)} \tag{7.3.25}$$

$$= \sum_{g\in G} \frac{\chi_M(g)\overline{\chi_M(g)}}{\dim M} \tag{7.3.26}$$

$$= \frac{|G|}{\dim M}\langle \chi_M, \chi_M\rangle \tag{7.3.27}$$

$$= \frac{|G|}{\dim M}. \tag{7.3.28}$$

This shows that this quantity is rational, as clearly $|G|$ and $\dim M$ are integers. Since the left-hand-side is in $\overline{\mathbb{Z}}$ and the right-hand-side is in $\mathbb{Q}$ they must actually be in $\overline{\mathbb{Z}} \cap \mathbb{Q} = \mathbb{Z}$, and thus $\dim M$ divides $|G|$. □

Lemma 7.3.29 If $\xi_1, \dots, \xi_n$ are roots of unity such that $a := (\xi_1 + \cdots + \xi_n)/n$ is an algebraic integer then either $\xi_1 = \cdots = \xi_n$ or $\xi_1 + \cdots + \xi_n = 0$.

Proof. If the ξ_i are not all equal then it follows from the geometry of roots of unity that $|a| < 1$. Suppose that $p(x)$ is the minimal polynomial with a as a root, then any other root, a', of this polynomial must also be a root of unity, and as such $|a'| \leq 1$ also. The product of all roots of p is an integer, and since they all have absolute value at most 1, and $|a| < 1$ it follows that this integer has absolute value less than 1, and so must be 0. Thus, $a = 0$, and since $1/n \neq 0$ we achieve the desired result. □

Theorem 7.3.30. Let G be a finite group and M a simple G-module. Let $C \in \mathcal{C}(G)$ be a conjugacy class such that $\gcd(|C|, \dim M) = 1$. Then either $\chi_M(g) = 0$ or $\rho_M(g) = \varepsilon \mathrm{id}_M$ for some $\varepsilon \in \mathbb{C}$ for all $g \in C$.

Proof. Since $\gcd(|C|, \dim M) = 1$ there exist integers a and b such that

$$a|C| + b \dim M = 1. \tag{7.3.31}$$

Multiplying by $\chi_M(g)/\dim M$ we get

$$\frac{|C|\chi_M(g)}{\dim M} + b\chi_M(g) = \frac{\chi_M(g)}{\dim M} = \frac{\varepsilon_1 + \cdots + \varepsilon_n}{n} \tag{7.3.32}$$

where ε_i are the eigenvalues of $\rho_M(g)$ and n is the dimension of M. Then the left-hand-side is an algebraic integer, since a is an integer, $|C|\chi_M(g)/\dim M$ is an algebraic integer by Proposition 7.3.17, b is an integer, and $\chi_M(g)$ is an algebraic integer as it is a sum of the eigenvalues of $\rho_M(g)$ which are roots of unity as g has finite order as G is finite. Thus, $(\varepsilon_1 + \cdots + \varepsilon_n)/n$ is an algebraic integer by Lemma 7.3.29, so it is either 0 or $\varepsilon_1 = \cdots = \varepsilon_n = \varepsilon$, in which case $\rho_M(g) = \varepsilon \mathrm{id}_M$. □

Theorem 7.3.33. Let G be a finite group and $C \in \mathcal{C}(G)$ a conjugacy class such that $|C| = p^k$ for p some prime and $k \in \mathbb{Z}_{>0}$. Then G has a proper nontrivial normal subgroup.

Proof. We may always split the set of simple G-modules as

$$\mathrm{Irr}\, G = \{\mathbb{C}\} \sqcup D \sqcup N \tag{7.3.34}$$

where $\mathbb{C}$ is the trivial representation, and D and N are the sets of "divisible" and "not divisible" dimension irreducible representations. That is,

$$D = \{M \in \mathrm{Irr}\, G \mid p \mid \dim M\}, \qquad \text{and} \qquad N = \{M \in \mathrm{Irr}\, G \mid p \nmid \dim M\}. \tag{7.3.35}$$

We claim that there exists some $M \in N$ such that $\chi_M(g) \neq 0$. To see this first note that if $M \in D$ then p divides $\dim M$, so $(\dim M)/p$ is an integer, and hence an algebraic integer. Thus,

$$a = \sum_{M \in D} \frac{1}{p} (\dim M) \chi_M(g) \tag{7.3.36}$$

is an algebraic integer. Taking some $g \in C$ we know that $g \neq 1$ since $|C| = p^k \neq 1$ and the identity is always in a conjugacy class on its own. Thus, by the second orthogonality relation we know that Theorem 6.5.19

$$\sum_{M \in \mathrm{Irr}(G)} \overline{\chi_M(e)} \chi_M(g) = 0 \tag{7.3.37}$$

and of course the character of the identity is just the dimension, so this is nothing but

$$\sum_{M \in \mathrm{Irr}(G)} (\dim M) \chi_M(g) = 0. \tag{7.3.38}$$

We can rewrite this sum in terms of the decomposition of $\mathrm{Irr}(G)$ as

$$0 = \chi_{\mathbb{C}}(g) + \sum_{M \in D} (\dim M) \chi_M(g) + \sum_{M \in N} (\dim M) \chi_M(g) \tag{7.3.39}$$

$$= 1 + pa + \sum_{M \in N} (\dim M) \chi_M(g). \tag{7.3.40}$$

Here we've used the fact that the character of the trivial representation is identically 1, as well as Equation (7.3.36) to identify a. Since $pa \neq -1$, as p is a prime and a an integer, we know that

$$\sum_{M \in N} (\dim M) \chi_M(g) \neq 0 \tag{7.3.41}$$

and thus there must be some $M \in N$ such that $\chi_M(g) \neq 0$.
Now fix $M \in N$ to be such that $\chi_M(g) \neq 0$ for $g \in C$. Since $p \nmid \dim M$ we know that $|C| = p^k$ doesn't divide M, and since p is prime $\dim M$ doesn't divide $|C|$ either. Thus, $\gcd(|C|, \dim M) = 1$, and since $\chi_M(g) \neq 0$ we know that $\rho_M(g) = \varepsilon \mathrm{id}_M$ for some ε for all $g \in C$ by Theorem 7.3.30.
Now define the subgroup

$$H = \langle gh^{-1} \mid g, h \in C \rangle. \tag{7.3.42}$$

This is not equal to $\{1\}$ as $|C| > 1$ so there exist distinct g and h in C and $gh^{-1} \neq 1$ as inverses are unique. By construction, H is normal since conjugation simply permutes the, since for all $k \in G$ we have

$$kgh^{-1}k^{-1} = kgk^{-1}kh^{-1}k^{-1} = \hat{g}\hat{h}^{-1} \tag{7.3.43}$$

for some $\hat{g}, \hat{h} \in C$ by definition of a conjugacy class.
Further, H acts trivially on M. To see this take $g, h \in C$, and then we know that $\rho_M(g) = \varepsilon_g \mathrm{id}_M$ and $\rho_M(h) = \varepsilon_h \mathrm{id}_M$ for some scalars $\varepsilon_g, \varepsilon_h \in \mathbb{C}$. Thus,

$\chi_M(g) = \varepsilon_g \dim M$ and $\chi_M(h) = \varepsilon_h \dim M$, but characters are constant on conjugacy classes, so it must be that $\varepsilon_g = \varepsilon_h$. Thus, H simply acts by some scalar multiple, $\varepsilon = \varepsilon_1 = \varepsilon_h$, and we're free to choose $\varepsilon = 1$, as we know that H does not act as zero.
Finally, it must be that $H \subsetneq G$, since if $G = H$ then G acts trivially on M, but by definition M is not the trivial representation. □

7.3.4 Proof of Burnside's Theorem

Finally, we're ready to put all of these technical results together to prove Burnside's theorem. We'll do this in two cases. The first is to prove that if the order of G has a unique prime factor then G is solvable, then the main result can be prove assuming two distinct prime factors.

Proposition 7.3.44 Let G be a group of order p^a for some prime, p, and $a \in \mathbb{Z}_{\geq 0}$. Then G is solvable.

Proof. First note that if $a = 0$ then G is trivial and is trivially solvable. We then induct on a. Suppose that the statement is true for all $a < n$ for some integer n. Now take $|G| = p^n$.
The class equation is a result from group theory which tells us that

$$|G| = |Z(G)| + \sum_i [G : Z_{g_i}] \tag{7.3.45}$$

where the sum is over conjugacy classes. The order of any conjugacy class of G must divide $|G|$, and so it follows that all conjugacy classes have size p^{k_i} for some $k_i \in \mathbb{Z}_{\geq 0}$. Then we have that $|G| = p^n = |Z(G)| + \sum_i p^{k_i}$. Thus, p must divide $|Z(G)|$, and so $Z(G)$ is nontrivial.
If G is abelian then G is solvable. If G is not abelian then $Z(G)$ is an abelian subgroup, which is solvable, meaning there exist normal subgroups

$$\{1\} \lhd Z_1 \lhd Z_2 \lhd \cdots \lhd Z_n = Z(G). \tag{7.3.46}$$

Quotients of successive terms are abelian as every group in this chain is abelian. Then $Z(G)$ is normal in G since everything in G commutes with everything in $Z(G)$. Further, $G/Z(G)$ is abelian. Thus, we have a chain of normal subgroups,

$$\{1\} \lhd Z_1 \lhd Z_2 \lhd \cdots \lhd Z_n = Z(G) \lhd G \tag{7.3.47}$$

such that quotients of successive subgroups are abelian. This proves G is solvable. □

Theorem 7.3.48 — Burnside's Theorem. Any group, G, of order $p^a q^b$ with p and q primes and $a, b \in \mathbb{Z}_{\geq 0}$ is solvable.

Proof. First, since the trivial group is solvable and Proposition 7.3.44 shows that all p-groups (that is, groups of order p^a) are solvable we may assume that p and q are distinct with $a, b \neq 0$. Finally, if G is abelian it is solvable, so we may assume that G is nonabelian, and in particular that $Z(G) \subsetneq G$. The proof is by contradiction, so assume G has order $p^a q^b$ and isn't solvable. Further, suppose that G is the smallest such G. Then G must be simple, else one of its normal subgroups would have this property.
We then know from Theorem 7.3.33 that G cannot have a conjugacy class, $C \in \mathcal{C}(G)$, of order p^k or q^k for $k \geq 1$. Thus, all conjugacy classes are either singletons or have order divisible by pq. However, we also know that

$$p^a q^b = |G| = \sum_{C \in \mathcal{C}(G)} |C| = 1 + \sum_{C \in \mathcal{C}(g) \setminus \{1\}} |C| \tag{7.3.49}$$

and the only way this can hold is if there is some $C \in \mathcal{C}(G)$ with $|C| = 1$, as if all conjugacy classes other than $\{1\}$ have order divisible by pq then 1 plus this sum cannot be divisible by pq. Thus, whatever element is in this C with $|C| = 1$ must be central. Hence, G has nontrivial centre, and thus has a normal subgroup, the centre of G. This is a contradiction of the simplicity, and hence a contradiction of our assumption of non-solvability. □

Eight

Induced Representations and Frobenius Reciprocity

8.1 Induced Representations

Let G be a finite group, and H a subgroup of G. Any G-module, M, may be viewed as an H-module in the obvious way. We just "forget" the fact that elements in $G \setminus H$ can act on M and consider only the action of elements in H. We call the resulting module the **restriction** of M to H, since if $\rho : G \to \mathrm{GL}(M)$ is the representation map for M as a G-module then the corresponding representation map for M as an H-module is $\rho|_H : H \to \mathrm{GL}(V)$.

For example, S_3 acts on $\mathbb{C}^3$ by permuting basis vectors, and $\mathbb{Z}_2 = \{(), (1\,2)\} \subset S_3$ acts on $\mathbb{C}^3$ by just swapping the first two basis vectors back and forth and leaving the third alone.

More formally, given a G-module, M, we have a canonical method of producing an H-module, and we can encode this as a functor

$$\mathrm{Res}^G_H : G\text{-}\mathsf{Mod} \to H\text{-}\mathsf{Mod} \tag{8.1.1}$$

which sends a G-module, M, to the H-module, $\mathrm{Res}^G_H M$, given by forgetting how elements of $G \setminus H$ act. This functor is the identity on module homomorphisms since the underlying sets of M and $\mathrm{Res}^G_H M$ are the same. We call this the **restriction functor**.

A natural question now is can we go the other direction? That is, if we have an H-module, M, is there a sensible way to construct a G-module? With the more formal statement above we might guess that the reverse process should be adjoint to Res^G_H. The following definition gives us exactly this reverse process.

Definition 8.1.2 — Induced Module Let G be a finite group and H a subgroup. An H-module, M, gives a G-module defined by

$$\mathrm{Ind}^G_H M := \Bbbk G \otimes_{\Bbbk H} M. \tag{8.1.3}$$

The action of G on $\mathrm{Ind}^G_H M$ is implicit in the definition of the tensor product, explicitly, it's given on simple tensors by

$$g \,.\, (g' \otimes m) = gg' \otimes m. \tag{8.1.4}$$

This all works out because $\Bbbk G$ is a $(\Bbbk G, \Bbbk H)$-bimodule (with the right $\Bbbk H$-module simply being restriction of the right regular representation). Thus, the tensor product of $\Bbbk G$ and $\Bbbk H$ is naturally a $\Bbbk G$-module.

As with restriction we have a functor

$$\mathrm{Ind}_H^G : H\text{-}\mathsf{Mod} \to G\text{-}\mathsf{Mod} \tag{8.1.5}$$

which sends M to $\Bbbk G \otimes_{\Bbbk H} M$ and an H-module homomorphism, $\varphi : M \to N$ is sent to a G-module homomorhpism

$$\mathrm{Ind}_H^G \varphi : \Bbbk G \otimes_{\Bbbk H} M \to \Bbbk G \otimes_{\Bbbk H} N \tag{8.1.6}$$

$$g \otimes m \mapsto g \otimes \varphi(n). \tag{8.1.7}$$

Note that there are several equivalent definitions of $\mathrm{Ind}_H^G M$ yielding isomorphic, but formally distinct, G-modules. One of these is

$$\mathrm{Ind}_H^G M \cong \{f : G \to M \mid f(hx) = \rho(h)f(x) \forall x \in G, h \in H\}. \tag{8.1.8}$$

That is, we consider all maps $G \to M$ which intertwine the regular representation of G and the action of G on M. Another definition is

$$\mathrm{Ind}_H^G M \cong \mathrm{Hom}_{\Bbbk H}(\Bbbk G, M) \tag{8.1.9}$$

which is really just restating the above. With these definitions the action of $g \in G$ on f is given by

$$(g \,.\, f)(x) = f(xg) \tag{8.1.10}$$

for all $x \in G$. Everything we might want then pretty much follows because $g \in G$ acts on the right of the function argument and $h \in H$ acts on the left. For example, this is a valid representation since we have

$$(g \,.\, f)(hx) = f(hxg) = \rho(h)f(xg) = \rho(h)(g \,.\, f)(x) \tag{8.1.11}$$

so $g \,.\, f$ is again in $\mathrm{Hom}_{\Bbbk H}(\Bbbk G, M)$, and

$$(g \,.\, (g' \,.\, f))(x) = (g' \,.\, f)(xg) = f(xgg') = (gg' \,.\, f)(x) \tag{8.1.12}$$

and

$$(1 \,.\, f)(x) = f(x1) = f(x) \tag{8.1.13}$$

for all $g, g', x \in G$.

Example 8.1.14 Let $\Bbbk$ be the trivial representation in which H acts as the identity, so the representation map is $\mathbb{1} : H \to \mathrm{GL}(\Bbbk) \cong \Bbbk$ with $\mathbb{1}(h) = 1$. Then, we have

$$\mathrm{Ind}_H^G \Bbbk = \Bbbk G \otimes_{\Bbbk H} \Bbbk. \tag{8.1.15}$$

Note that the tensor product is $\otimes_{\Bbbk H}$, not $\otimes_{\Bbbk}$, so $\Bbbk G \otimes_{\Bbbk H} \Bbbk$ is not isomorphic to $\Bbbk G$.

The module structure is completely determined by elements of the form $g \otimes 1$. In fact, since $gh \otimes 1 = g \otimes (h \,.\, 1) = g \otimes 1$ the action is invariant under multiplication by elements of H. Both gh and gh' have the same action for $g \in G$ and $h, h' \in H$. Thus, the action of $g \in G$ is determined only by the coset, gH, into which it falls.
The induced module, $\mathrm{Ind}_H^G \Bbbk$ is isomorphic to the coset representation, $\Bbbk G/H$, which is a G-module constructed as the free vector space on the set of cosets, G/H, with the G-action given by $g \,.\, g'H = (gg')H$.

Example 8.1.16 Let G be a finite group with subgroup H. Let $\chi : H \to \Bbbk^\times$ be a homomorphism, and $\Bbbk_\chi$ the corresponding 1-dimensional representation of H. That is, $h \,.\, \lambda = \chi(h)\lambda$ for all $h \in H$ and $\lambda \in \Bbbk$.
Consider the induced module $\mathrm{Ind}_H^G \Bbbk_\chi = \Bbbk G \otimes_{\Bbbk H} \Bbbk_\chi$. For $h \in H$ we have $h \otimes 1 = 1_G \otimes (h \,.\, 1) = 1_G \otimes \chi(h)$. The action of $g \in G$ on $\mathrm{Ind}_H^G \Bbbk_\chi$ is given by $g \,.\, (g' \otimes 1) = gg' \otimes 1$.
Let

$$e_\chi = \frac{1}{|K|} \sum_{h \in H} \chi(h)^{-1} h \in \Bbbk H. \tag{8.1.17}$$

We claim that $\mathrm{Ind}_H^G \Bbbk_\chi \cong \Bbbk G e_\chi$, where elements of $\Bbbk G e_\chi$ are $\Bbbk$-linear combinations of elements of the form $g e_\chi$ and the action on $\Bbbk G e_\chi$ is by $g \,.\, g' e_\chi = (gg')e_\chi$, that is, it's just left multiplication.
The isomorphism, $\varphi : \mathrm{Ind}_H^G \Bbbk_\chi \to \Bbbk G e_\chi$, is given by $\varphi(g \otimes 1) = g e_\chi$. This is a G-module homomorphism since

$$\begin{aligned} \varphi(g \,.\, (g' \otimes 1)) &= \varphi(gg' \otimes 1) && (8.1.18)\\ &= gg' e_\chi && (8.1.19)\\ &= \varphi(gg' \otimes 1). && (8.1.20) \end{aligned}$$

Note that if $g = kh$ with $h \in H$ then we have $g \otimes 1 = k \otimes \chi(h)$, and so we need to check that φ is well defined with respect to this ambiguity. In particular, if $g = kh = k'h'$ for $h, h' \in H$ then we need to check that $\varphi(k \otimes \chi(h)) = \varphi(k' \otimes \chi(h'))$. This is true since $\chi(h)$ and $\chi(h')$ are scalars, so we can pull the out and we have

$$\varphi(k \otimes \chi(h)) = \chi(h)\varphi(k \otimes 1) = \chi(h) k e_\chi \tag{8.1.21}$$

and similarly, $\varphi(k' \otimes \chi(h')) = \chi(h') k' e_\chi$. To show that these are equal we start with the definition of e_χ:

$$\begin{aligned} \chi(h) k e_\chi &= \chi(h) k \frac{1}{|H|} \sum_{g \in H} \chi(g)^{-1} g && (8.1.22)\\ &= \frac{1}{|H|} \sum_{g \in H} \chi(h)\chi(g)^{-1} k g. && (8.1.23) \end{aligned}$$

We can then reindex the sum by defining $g' \in G$ such that $kg = k'g'$, so $g = k^{-1}k'g'$. Since $kh = k'h'$ this gives $h = k^{-1}k'h'$. Thus,

$$\chi(h)ke_\chi = \frac{1}{|H|}\sum_{g'\in H}\chi(k^{-1}k'h')\chi(k^{-1}k'g')^{-1}k'g' \tag{8.1.24}$$

$$= \frac{1}{|H|}\sum_{g'\in H}\chi(k^{-1}k'h')\chi(g'^{-1}k'^{-1})k'g' \tag{8.1.25}$$

$$= \frac{1}{|H|}\sum_{g'\in H}\chi(k^{-1}k'h'g'^{-1}k'^{-1}k)k'g' \tag{8.1.26}$$

using $k^{-1}k' = hh'^{-1}$ and $k'^{-1}k = h'h^{-1}$ this becomes

$$\chi(h)ke_\chi = \frac{1}{|H|}\sum_{g'\in H}\chi(hh'^{-1}h'g'^{-1}h'h^{-1})k'g' \tag{8.1.27}$$

$$= \frac{1}{|H|}\sum_{g'\in H}\chi(hg'^{-1}h'h^{-1})k'g'. \tag{8.1.28}$$

Now, since χ maps into $\mathbb{C}^\times$ and is a group homomorphism we have that $\chi(ab)\chi(a)\chi(b) = \chi(b)\chi(a) = \chi(ba)$, and it follows that $\chi(hg'^{-1}h'h^{-1}) = \chi(h^{-1}hg'^{-1}h') = \chi(g'^{-1}h') = \chi(h')\chi(g')^{-1}$, and thus

$$\chi(h)ke_\chi = \frac{1}{|H|}\sum_{g'\in H}\chi(h')\chi(g')^{-1}k'g'$$

$$= \chi(h')k'\frac{1}{|H|}\sum_{g'\in H}\chi(g')^{-1}g' = \chi(h')k'e_\chi. \tag{8.1.29}$$

This shows that φ is well defined. Clearly φ is invertible, and so we have the claimed isomorphism, $\operatorname{Ind}_H^G \Bbbk_\chi \cong \Bbbk G e_\chi$.

The first example above actually gives yet another way of characterising the induced representation. If G is a finite group and H a (not necessarily normal) subgroup then we can form the coset space, G/H. Taking $\{g_1, \dots, g_n\}$ to be a complete set of representatives, that is each coset can be written as g_iH in exactly one way, we can take

$$\operatorname{Ind}_H^G M = \bigoplus_{i=1}^n g_iM \tag{8.1.30}$$

where each g_iM is an isomorphic copy of M, and we write elements of g_iM as g_im. For each $g \in G$ there is some $h_i \in H$ and $j(i) \in \{1, \dots, n\}$ such that $gg_i = g_{j(i)}h_i$, which simply restates that $\{g_1, \dots, g_n\}$ is a complete set of representatives. Then $g \in G$ acts on this space by

$$g \,.\, g_im = g_{j(i)}\rho(h_i)m_i. \tag{8.1.31}$$

So g acts by permuting the copies of M, sending g_iM to $g_{j(i)}M$, with an extra "twist" provided by the action of h_i on M. Another way of constructing this is to take

$$g_iM = \{f \in \operatorname{Hom}_{\Bbbk H}(\Bbbk G, M) \mid f(g) = 0 \text{ unless } g \in g_iH\}. \tag{8.1.32}$$

Then the action is by

$$(g \,.\, f)(x) = f(xg) \tag{8.1.33}$$

again, and we just take evaluating to zero to be equivalent to not being in $g_i M$ as defined before.

8.2 Frobenius Formula for Induced Characters

Calculating the character of a restricted module is simple. If we have a G-module, M, with character $\chi : G \to \Bbbk$ then the character of $\mathrm{Res}_H^G M$ is just the restriction of the character to H, $\chi \downarrow_H^G := \chi|_H : H \to \Bbbk$. In this section we give a method for calculating characters of induced modules, a more involved process.

Theorem 8.2.1 — Frobenius Formula. Let G be a finite group with subgroup H. Let $\{g_1, \dots, g_n\}$ be a complete set of representatives for G/H. Let M be an H-module with character χ_M. Write $\chi_M \uparrow_H^G$ for the character of $\mathrm{Ind}_H^G M$. Then

$$\chi_M \uparrow_H^G (g) = \sum_{i=1}^{n} \chi_M(g_i^{-1} g g_i) \tag{8.2.2}$$

where χ_M has been extended from H to all of G such that $\chi_M(x) = 0$ if $x \notin H$.

Proof. We shall work with

$$\mathrm{Ind}_H^G M = \bigoplus_{i=1}^{n} g_i M. \tag{8.2.3}$$

Thus, we have

$$\chi_M \uparrow_H^G (g) = \sum_i \chi_i(g) \tag{8.2.4}$$

where $\chi_i(g) = \mathrm{tr}_{g_i M}\, \rho_i(g)$ with ρ_i defined to be the corresponding blocks in the matrix

$$\rho(g) = \begin{pmatrix} \rho_1(g) & & & \\ & \rho_2(g) & & \\ & & \ddots & \\ & & & \rho_n(g) \end{pmatrix} \tag{8.2.5}$$

extended so that $\rho_i(g) = 0$ unless $gg_i \in g_i H$.
For the nonzero terms we know that $gg_i \in g_i H$ means there is some $h^{-1} \in H$ such that $gg_i h^{-1} = g_i$, and thus $g_i^{-1} g g_i = h \in H$. Now define a map $\alpha : g_i M \to M$ by $\alpha(f) = f(g_i)$. This is an isomorphism, and we have

$$\alpha(g \,.\, f) = (g \,.\, f)(g_i) = f(g_i g) = f(hg_i) = \rho(h) f(g_i) = h \,.\, \alpha(f) \tag{8.2.6}$$

and so $g \,.\, f = \alpha^{-1}(h \,.\, \alpha(f))$. This means that $\mathrm{tr}_{g_i M}\, \rho_i(g) = \chi_M(h)$. Thus, we have

$$\chi_M \uparrow_H^G (g) = \mathrm{tr}_{\mathrm{Ind}_H^G M}\, \rho(g) = \sum_{i=1}^{n} \mathrm{tr}_{g_i M}\, \rho_i(g_i) = \sum_{i=1}^{n} \chi_M(g_i g g_i^{-1}) \tag{8.2.7}$$

as claimed. □

Corollary 8.2.8 With notation as in Theorem 8.2.1 if char $\Bbbk$ and $|H|$ are coprime then we have

$$\chi_M \uparrow_H^G (g) = \frac{1}{|H|} \sum_{\substack{x \in G \\ x^{-1}gx \in H}} \chi_M(x^{-1}gx). \tag{8.2.9}$$

Proof. We have that

$$\chi_M \uparrow_H^G (g) = \sum_{i=1}^{n} \chi_M(g_i^{-1} g g_i). \tag{8.2.10}$$

Since χ_M is a class function it is invariant under conjugation of its argument, so we can write this as

$$\chi_M \uparrow_H^G (g) = \sum_{i=1}^{n} \chi_M(h^{-1} g_i^{-1} g g_i h). \tag{8.2.11}$$

for any $h \in H$. In fact, we can actually sum over all $h \in H$, and all this does is give us $|H|$ identical terms[a], so

$$\chi_M \uparrow_H^G (g) = \frac{1}{|H|} \sum_{h \in H} \sum_{i=1}^{n} \chi_M(h^{-1} g_i^{-1} g g_i h). \tag{8.2.12}$$

We can then recognise that the argument of χ_M is g conjugated by $x = g_i h$, which is chosen such that $x^{-1}gx \in H$ since $g_i^{-1} g g_i \in H$ and conjugation by $h \in H$ doesn't take us out of H. Thus, we have

$$\chi_M \uparrow_H^G (g) = \frac{1}{|H|} \sum_{\substack{x \in G \\ x^{-1}gx \in H}} \chi_M(x^{-1}gx) \tag{8.2.13}$$

where all we've done is combine the two sums, over $h \in H$ and $i \in \{1, \dots, n\}$ into a single sum. □

[a] This is where we need char $\Bbbk \nmid |H|$, if this wasn't the case we may accidentally have everything vanish in this sum.

8.3 Frobenius Reciprocity

Frobenius reciprocity is the relationship between induced and restricted modules. The strongest form of this result is that Res_H^G and Ind_H^G are adjoint functors. Before we get to that we'll give a result that holds for characters.

8.3.1 Froebnius Reciprocity of Characters

Theorem 8.3.1 — Frobenius Reciprocity of Characters. Let G be a finite group and H a subgroup. Let M be a G-module and N an H-module, both over $\mathbb{C}$. Write $\langle -, - \rangle_G$ and $\langle -, - \rangle_H$ for the inner product on the space of class functions of G and H respectively. Write χ_M and χ_N for the characters of M and N respectively. Write $\chi_N \uparrow_H^G$ for the character of $\mathrm{Ind}_H^G N$ and $\chi_M \downarrow_H^G$ for the character of $\mathrm{Res}_H^G M$. Then

$$\langle \chi_N \uparrow_H^G, \chi_M \rangle_G = \langle \chi_N, \chi_M \downarrow_H^G \rangle_H. \tag{8.3.2}$$

Proof. Write $\chi = \chi_M$. Then, by definition of the inner product of class functions we have

$$\langle \chi_N \uparrow_H^G, \chi_M \rangle_G = \frac{1}{|G|} \sum_{g \in G} \chi_W \uparrow_H^G (g) \overline{\chi(g)}. \tag{8.3.3}$$

Define a function

$$\psi(g) = \begin{cases} \chi_N(g) & g \in H, \\ 0 & \text{else.} \end{cases} \tag{8.3.4}$$

Then, using the Frobenius formula to calculate $\chi_N \uparrow_H^G (g)$ we have

$$\begin{aligned} \langle \chi_N \uparrow_H^G, \chi_M \rangle_G &= \frac{1}{|G|} \sum_{g \in G} \frac{1}{|H|} \sum_{\substack{x \in G \\ xgx^{-1} \in H}} \psi(x^{-1}gx) \overline{\chi(g)} & (8.3.5) \\ &= \frac{1}{|G|} \frac{1}{|H|} \sum_{g \in G} \sum_{\substack{x \in G \\ x^{-1}gx \in H}} \psi(x^{-1}gx) \chi(g^{-1}) & (8.3.6) \end{aligned}$$

where in the last step we've just rearranged some terms and used $\overline{\chi(g)} = \chi(g^{-1})$. Now we can reindex the sum by taking $y = x^{-1}gx$, which means $g^{-1} = xy^{-1}x^{-1}$, and the condition that $x^{-1}gx \in H$ becomes that $y \in H$, so we have

$$\langle \chi_N \uparrow_H^G, \chi_M \rangle_G = \frac{1}{|G|} \frac{1}{|H|} \sum_{x \in G} \sum_{y \in H} \psi(y) \chi(xy^{-1}x^{-1}). \tag{8.3.7}$$

We also have that $\psi(y) = \chi_N(y)$, since $y \in H$, and thus this becomes

$$\langle \chi_N \uparrow_H^G, \chi_M \rangle_G = \frac{1}{|G|} \frac{1}{|H|} \sum_{x \in G} \sum_{y \in H} \chi_N(y) \chi(xy^{-1}x^{-1}). \tag{8.3.8}$$

Since χ is a class function $\chi(xy^{-1}x^{-1}) = \chi(y^{-1})$, and so we get $|G|$ terms which are all equal to $\chi(y^{-1}) = \overline{\chi(y)}$. This perfectly cancels with the sum over $x \in G$, leaving us with

$$\langle \chi_N \uparrow_H^G, \chi_M \rangle_G = \frac{1}{|H|} \sum_{y \in G} \chi_N(y) \overline{\chi(y)} = \langle \chi_N, \chi \rangle_H. \tag{8.3.9}$$

This proves the result once we realise that since the sum is over $y \in H$ we can replace $\chi = \chi_M : G \to \mathbb{C}$ with $\chi_M \downarrow_H^G = \chi_M|_H : H \to \mathbb{C}$. □

One thing that this result tells us is that the multiplicities of induced modules and restricted modules are related. In particular, we have

$$\dim(\mathrm{Hom}_G(M, \mathrm{Ind}_H^G N)) = \dim(\mathrm{Hom}_H(\mathrm{Res}_H^G M, N)). \tag{8.3.10}$$

Thus, there exists, at the level of vector spaces, an isomorphism between these hom-spaces.

8.3.2 Frobenius Reciprocity

Theorem 8.3.11. Let G be a finite group with subgroup H. Then the functors

$$\mathrm{Res}_H^G : G\text{-}\mathsf{Mod} \to H\text{-}\mathsf{Mod}, \quad \text{and} \quad \mathrm{Ind}_H^G : H\text{-}\mathsf{Mod} \to G\text{-}\mathsf{Mod} \tag{8.3.12}$$

are left and right adjoints. That is, there is a (natural) isomorphism

$$\mathrm{Hom}_G(M, \mathrm{Ind}_H^G N) \cong \mathrm{Hom}_H(\mathrm{Res}_H^G M, N) \tag{8.3.13}$$

for any G-module, M, and H-module, N.

Proof. Let

$$E = \mathrm{Hom}_G(M, \mathrm{Ind}_H^G N), \quad \text{and} \quad E' = \mathrm{Hom}_H(\mathrm{Res}_H^G M, N). \tag{8.3.14}$$

We need to define two functions

$$\Phi : E \to E', \quad \text{and} \quad \Phi' : E' \to E \tag{8.3.15}$$

which should then be inverses.
If $\alpha \in E$ then $\alpha : M \to \mathrm{Ind}_H^G N$ is a G-module homomorphism, and $\Phi(\alpha)$ should be an H-module homomorphism, $\Phi(\alpha) : \mathrm{Res}_H^G M \to N$. That is, $\Phi(\alpha)$ needs to take in an element of $\mathrm{Res}_H^G M$, which is just an element of M, and produce an element of N. The obvious way to do this is to simply evaluate α, which gives us an element of $\mathrm{Ind}_H^G N \cong \mathrm{Hom}_{\Bbbk H}(\Bbbk G, N)$, which we can then evaluate to produce an element of N. The only problem is what element of $\Bbbk G$ do we evaluate this map at? Fortunately since G is a group there's an obvious distinguished element, 1_G, at which to perform this evaluation. Thus, we define $\Phi(\alpha)$ by

$$\Phi(\alpha)(m) = \alpha(m)(1_G) \tag{8.3.16}$$

for $m \in \mathrm{Res}_M^G$ (which as a set is just M).
If $\beta \in E'$ then $\beta : \mathrm{Res}_H^G \to N$ is an H-module homomorphism, and $\Phi'(\beta)$ should be a G-module homomorphism, $\Phi'(\beta) : M \to \mathrm{Ind}_H^G N$. That is, $\Phi'(\beta)$ needs to take in an element of M and produce an element of $\mathrm{Ind}_H^G N \cong \mathrm{Hom}_{\Bbbk H}(\Bbbk G, N)$. The correct definition turns out to be

$$\Phi'(\beta)(m)(x) = \beta(xm) \tag{8.3.17}$$

where $m \in M$ and $x \in \Bbbk G$ so $xm \in M$ using the G-module structure of M, which is equal to $\mathrm{Res}_H^G M$ as a set, and so evaluating β at xm is a valid operation.

With these definitions we need to show that the resulting functions are well-defined. This comes down to the following three steps:

1. We need to show that $\Phi(\alpha)$ is an H-module homomorphism. That is, we need to show that $\Phi(\alpha)(h\,.\,m) = h\,.\,\Phi(\alpha)(m)$ for all $h \in H$ and $m \in M$. This is the case, as a direct calculation shows. First, using the definition of Φ we have
$$\Phi(\alpha)(h\,.\,m) = \alpha(h\,.\,m)(1_G). \tag{8.3.18}$$
Since α is a G-module homomorphism we have $\alpha(h\,.\,m) = h\,.\,\alpha(m)$, and so
$$\Phi(\alpha)(h\,.\,m) = (h\,.\,\alpha(m))(1_G). \tag{8.3.19}$$
Since $\alpha(m)$ is an H-module homomorphism the action of h on $\alpha(m)$ is to act on the right in the argument, which is just multiplication in this case:
$$\Phi(\alpha)(h\,.\,m) = \alpha(m)(1_G h). \tag{8.3.20}$$
Since $1_G h = h 1_G$ we can write this as
$$\Phi(\alpha)(h\,.\,m) = \alpha(m)(h 1_G). \tag{8.3.21}$$
We can then identify that acting on the left of the argument is the definition of the action of G on the G-module homomorphism α
$$\Phi(\alpha)(h\,.\,m) = h\,.\,(\alpha(m))(1_G) = h\,.\,(\Phi(\alpha)(m)). \tag{8.3.22}$$
2. Next, we need to show that $\Phi'(\beta)(m) \in \mathrm{Ind}_H^G N$. That is, we need to show that $\Phi'(\beta)(m)(hx) = h\,.\,\Phi'(\beta)(m)(x)$. This also follows from a direct calculation, we have
$$\Phi'(\beta)(m)(hx) = \beta(hxm) = h\,.\,\beta(xm) = h\,.\,\Phi'(\beta)(m)(x) \tag{8.3.23}$$
having used the fact that β is an H-module homomorphism.
3. Finally, we need to show that $\Phi'(\beta)$ is a G-module homomorphism. That is, we need to show that $\Phi'(\beta)(g.m) = g.\Phi'(\beta)(m)$. This follows since
$$\Phi'(\beta)(g\,.\,m)(x) = \beta(xg\,.\,m) = \Phi'(\beta)(m)(xg) = (g\,.\,\Phi'(\beta)(m))(x) \tag{8.3.24}$$
having used the fact that $\Phi'(\beta) \in \mathrm{Ind}_H^G M$ in the last step.

We now just have to show that Φ and Φ' are inverses, this follows from two calculations:
$$\Phi(\Phi'(\beta))(m) = \Phi'(\beta)(m)(1_G) = \beta(1_G m) = \beta(m), \tag{8.3.25}$$
so $\Phi \circ \Phi' = \mathrm{id}_{E'}$, and
$$\begin{aligned}\Phi'(\Phi(\alpha))(m)(x) &= \Phi(\alpha)(xm) = \alpha(xm)(1_G)\\ &= (x\,.\,\alpha)(m)(1_G) = \alpha(m)(1_G x) = \alpha(m)(1_G)\end{aligned} \tag{8.3.26}$$
so $\Phi' \circ \Phi = \mathrm{id}_E$. □

Part III

Symmetric Group Representations

Nine

Representation Theory of the Symmetric Group

9.1 Combinatoric Preliminaries

The representation theory of the symmetric group, S_n, is mostly controlled by the combinatorics of partitions. In this section we set up some of the important objects which allow for efficient computations with representations of S_n.

Definition 9.1.1 — Partition Let n be a nonnegative integer. A **partition** of n is a nonincreasing sequence of nonnegative integers, $\lambda = (\lambda_1, \lambda_2, \lambda_3, \dots)$, such that $\lambda_1 + \lambda_2 + \lambda_3 + \cdots = n$.

Notation 9.1.2 We write $\lambda \vdash n$ to say that λ is a partition of n.
We write $|\lambda|$ for n.
We write $\ell(\lambda)$ for the number of nonzero parts, that is λ_ℓ is the last nonzero term in the sequence λ.

Notice that since n is finite and λ is nonincreasing it must be that $\lambda_i = 0$ for i sufficiently large, so usually we'll just consider λ as a finite sequence. For example, there are 7 partitions of 5:

$$(5), \quad (4,1), \quad (3,2), \quad (3,1,1), \quad (2,2,1) \quad (2,1,1,1), \quad \text{and} \quad (1,1,1,1,1).$$

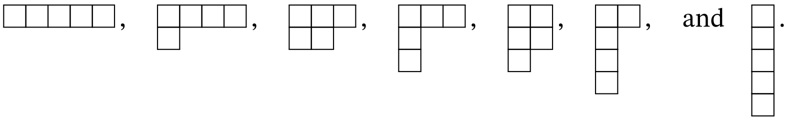
The number of partitions of n, often denoted $p(n)$, grows pretty quickly. For $n = 0, \dots, 15$ $p(n)$ is given by [OEIS A000041]

n	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
$p(n)$	1	1	2	3	5	7	11	15	22	30	42	56	77	101	135	176

Listing numbers makes it hard to spot patterns, and isn't very natural for some of the definitions we want to give. Most work with partitions is done with a graphical notation, known as Young diagrams.

Definition 9.1.3 — Young Diagrams For a partition, λ, of n, the corresponding **Young diagram**, also denoted λ, is made of n boxes arranged in a left-aligned grid with λ_i boxes in the ith row.

For example, the Young diagrams of the partitions of 5 listed above are



On their own Young diagrams are nice, but the real power comes when we start putting things in the boxes. In theory these could be anything, but the following definition gives the most useful case for us.

Definition 9.1.4 — Young Tableaux Let λ be a partition of n. A **Young tableau** (pl. tableaux) of shape λ is a filling of the boxes of λ with the numbers $1, \dots, n$. Write $Y(\lambda)$ for the set of boxes in λ, then a Young tableau of shape λ is precisely a function $T\colon Y(\lambda) \to \{1, \dots, n\}$.

> ! The lectures assume that T is a bijection, I think this is a bad assumption, since semistandard Young tableaux are pretty important.

Not all Young tableaux of a given shape are equally important when it comes to representation theory. The following definition gives the most common restrictions on Young tableaux.

It is useful to index the boxes by their position in the Young diagram. This is done with "matrix index" rules, we start at the top left corner with $(1,1)$, going one box right gives $(1,2)$, and one box down gives $(2,1)$. That is, we index with row number followed by column number.

Definition 9.1.5 Let λ be a partition, and T a Young tableau of shape λ. Then we say that T is **semistandard** if

$$T(i,j) \le T(i,j+1) \qquad \text{and} \qquad T(i,j) < T(i+1,j). \tag{9.1.6}$$

That is, a semistandard tableau has weakly increasing rows and strictly increasing columns. We say that T is **semistandard** if $T\colon Y(\lambda) \to \{1, \dots, n\}$ is a bijection and in addition

$$T(i,j) < T(i,j+1) \qquad \text{and} \qquad T(i,j) < T(i+1,j). \tag{9.1.7}$$

That is, a standard tableau has strictly increasing rows and strictly increasing columns and every number from 1 to n appears exactly once.

> R Some authors call any not-necessarily-bijective filling of a Young diagram with any alphabet a Young tableau, others assume that all Young tableau are at least semistandard, so you have to be careful about conventions.

Consider the partition $\lambda = (3,2)$. The following are all standard Young tableau of shape λ with labels in $\{1, \dots, 5\}$:

$$\begin{array}{|c|c|c|}\hline 1&2&3\\\hline 4&5\\\cline{1-2}\end{array}\,, \quad \begin{array}{|c|c|c|}\hline 1&2&4\\\hline 3&5\\\cline{1-2}\end{array}\,, \quad \begin{array}{|c|c|c|}\hline 1&2&5\\\hline 3&4\\\cline{1-2}\end{array}\,, \quad \text{and} \quad \begin{array}{|c|c|c|}\hline 1&3&4\\\hline 2&5\\\cline{1-2}\end{array}\,. \tag{9.1.8}$$

Notation 9.1.9 We write $\mathrm{SYT}(\lambda)$ for the set of all standard Young tableaux of shape λ.

The number of standard tableaux of any partition of n, that is

$$\sum_{\lambda \vdash n} |\mathrm{SYT}(\lambda)|, \tag{9.1.10}$$

for n from 0 to 15 is given by [OEIS A000085]

$$1, 1, 2, 4, 10, 26, 76, 232, 764, 2620, 9496, 35\,696,$$
$$140\,152, 568\,504, 2\,390\,480, 10\,349\,536.$$

This is also the number of involutions in S_n (Example 7.2.26) a fact that will follow from Theorem 7.2.8 once we have identified the relationship between standard Young tableaux and irreducible representations of S_n in the next section.

9.2 Constructing Simple S_n-Modules

9.2.1 Row and Column Groups

Fix some partition, λ, of n, and let the tableau $T\colon\, Y(\lambda) \to \{1, \dots, n\}$ be a bijective filling of the boxes of λ.

Definition 9.2.1 — Canonical Tableau The **canonical tableau** of shape λ is the filling given by assigning the numbers 1 through n in order going from left to right, top to bottom.

For example, the canonical tableau of shape $(3, 2)$ is

$$\begin{array}{|c|c|c|}\hline 1 & 2 & 3 \\ \hline 4 & 5 \\ \cline{1-2} \end{array}\,. \tag{9.2.2}$$

Definition 9.2.3 — Row and Column Groups There is a natural action of S_n on any bijective filling of boxes, simply permute the numbers as usual. That is, if $w \in S_n$ and $T(i, j) = k \in \{1, \dots, n\}$ then $w.T$ is the Young tableau of shape λ with $(w\,.\,T)(i, j) = w(k) = w(T(i, j))$.
The **row group** of a Young tableau, T, is the subgroup, R_T, of S_n which acts by permuting elements within rows without permuting elements between columns. Similarly, the **column group** of a Young tableau, T, is the subgroup, C_T, of S_n which acts by permuting elements within columns without permuting elements between rows.

R The lecture notes, and some other literature uses P_T and Q_T for the row and column group. I can't ever remember which is which, so I've changed it to R and C.

It is common to write R_λ and C_λ for the row and column group of the canonical tableau, T_0.

Explicitly, we have

$$R_T = \{w \in S_n \mid T^{-1}(w(T(i,j))) \text{ is in row } i\} \tag{9.2.4}$$

and

$$C_T = \{w \in S_n \mid T^{-1}(w(T(i,j))) \text{ is in column } j\}. \tag{9.2.5}$$

Since the action of the row group is always to permute rows for a Young tableau with ℓ rows we can identify that

$$R_n \cong S_{\lambda_1} \times S_{\lambda_2} \times \cdots \times S_{\lambda_\ell} =: S_\lambda \tag{9.2.6}$$

for $\lambda = (\lambda_1, \dots, \lambda_\ell)$. In this S_{λ_i} acts by permuting boxes in the ith row, which has, by definition, λ_i boxes. Before we can make a similar identification for C_T we need the notion of the transpose of a Young diagram.

> **Definition 9.2.7 — Transpose** Let λ be a Young diagram. Its **transpose**, λ', is the Young diagram given by reflecting along the main diagonal. This can be extended to Young tableau, simply transpose the underlying diagram and keep the corresponding numbering, so $T'(i,j) = T(j,i)$.

For example, if $\lambda = (3,2)$ then $\lambda' = (2,2,1)$, or in terms of Young diagrams,

$$\lambda = \ydiagram{3,2} \implies \lambda' = \ydiagram{2,2,1}. \tag{9.2.8}$$

Since the transpose swaps rows and columns of a Young diagram we can see that it swaps row and column groups, so $R_{T'} = C_T$ and $C_{T'} = R_T$. Thus, we can identify that

$$C_T = S_{\lambda'_1} \times S_{\lambda'_2} \cdots \times S_{\lambda'_\ell} = S_{\lambda'}. \tag{9.2.9}$$

9.2.2 Symmetrisers, Antisymmetrisers, and Projectors

> **Definition 9.2.10 — Symmetrisers, Antisymmetrisers and Projectors** Given a partition, λ, let T_0 be the corresponding canonical tableau. We define three elements of $\Bbbk S_n$:
>
> 1. The **Young symmetriser** is
>
> $$a_\lambda := \frac{1}{|R_{T_0}|} \sum_{w \in R_{T_0}} w. \tag{9.2.11}$$
>
> 2. The **Young antisymmetriser** is
>
> $$b_\lambda = \frac{1}{|C_{T_0}|} \sum_{w \in C_{T_0}} \operatorname{sgn}(w) w. \tag{9.2.12}$$
>
> 3. The **Young projector** is
>
> $$c_\lambda = a_\lambda b_\lambda. \tag{9.2.13}$$

For example, consider $\lambda = (2,1)$. The row group is $\{(), (1\,2)\}$, simply permuting the entries of the first row. The column group is also $\{(), (1\,3)\}$, permuting the entries of the first column. Thus,

$$a_\lambda = \frac{1}{2}[() + (1\,2)), \quad \text{and} \quad b_\lambda = \frac{1}{2}(() - (1\,3)]. \tag{9.2.14}$$

Then[1]

$$c_\lambda = \frac{1}{4}[() + (1\,2) - (1\,3) - (1\,3\,2)]. \tag{9.2.15}$$

[1] I'm making a decision here that permutations multiply by acting on something to their right, so they multiply to give a left action of S_n.

For any vector space, V, there is a natural action of S_n on $V^{\otimes n}$, permuting the factors. This is where the names above come from. For example, if $\lambda = (3) = \square\square\square$ then $R_{T_0} = S_3$ and

$$a_{(3)} = \frac{1}{6}[cycle + (1\,2) + (1\,3) + (2\,3) + (1\,2\,3) + (1\,3\,2)] \tag{9.2.16}$$

and the action on $v_1 \otimes v_2 \otimes v_3 \in V^{\otimes 3}$ is

$$\begin{aligned} v = a_{\square\square\square} \,.\, (v_1 \otimes v_2 \otimes v_3) = \frac{1}{6}(&v_1 \otimes v_2 \otimes v_3 + v_2 \otimes v_1 \otimes v_3 + v_3 \otimes v_2 \otimes v_1 \\ &+ v_1 \otimes v_3 \otimes v_2 + v_2 \otimes v_3 \otimes v_1 + v_3 \otimes v_1 \otimes v_1). \end{aligned} \tag{9.2.17}$$

This is then symmetric in the sense that $w \,.\, v = v$. Similarly, if

$$\lambda = (1,1,1) = \begin{array}{|c|}\hline \\ \hline \\ \hline \\ \hline \end{array} \tag{9.2.18}$$

then $C_{T_0} = S_3$,

$$b_{(1,1,1)} = \frac{1}{6}[() - (1\,2) - (1\,3) - (2\,3) + (1\,2\,3) + (1\,3\,2)] \tag{9.2.19}$$

and

$$\begin{aligned} v = b_{(1,1,1)} \,.\, (v_1 \otimes v_2 \otimes v_3) = \frac{1}{6}(&v_1 \otimes v_2 \otimes v_3 - v_2 \otimes v_1 \otimes v_3 - v_3 \otimes v_2 \otimes v_1 \\ &- v_1 \otimes v_3 \otimes v_2 + v_2 \otimes v_3 \otimes v_1 + v_3 \otimes v_1 \otimes v_1). \end{aligned} \tag{9.2.20}$$

This is then antisymmetric in the sense that $w \,.\, v = \operatorname{sgn}(w) v$.

One way of looking at this is that $a_{(3)}$ projects $V^{\otimes 3}$ to the subspace on which S_n acts trivially, whereas $b_{(1,1,1)}$ projects $V^{\otimes 3}$ onto the subspace where S_n acts by the sign representation. In general, we have $a_{(n)} V^{\otimes n} = S^n V$ and $b_{(1,\dots,1)} V^{\otimes n} = \Lambda^n V$.

9.2.3 Specht Modules

> **Definition 9.2.21 — Specht Module** For λ a partition of n we call the module $V_\lambda \coloneqq \Bbbk S_n c_\lambda$ the **Specht module**.

Remark 9.2.22 Our definition here is rather abstract. A more direct definition of the Specht modules is via **tabloids**, which are equivalence classes of Young tableau, $\{T\}$, under the action of the row group. That is, two Young tableau are equivalent if we can get from one to the other by permuting elements within a row. The column group acts on these tabloids by permuting elements between different rows, note that these elements now no longer need to be in the same column, since we can always move elements freely within rows of a tableau without leaving the equivalence class. Then for a tableau, T, we define the formal linear combination of equivalence classes

$$E_T = \sum_{w \in C_T} \operatorname{sgn}(w)[w \,.\, T]. \tag{9.2.23}$$

Doing this for all *standard* Young tableau of shape λ, we declare the resulting E_T to be a basis for some vector space, V. There is an action of the symmetric group on V, defined on this basis by

$$\sigma \,.\, E_T = \sum_{w \in C_T} \operatorname{sgn}(w)[\sigma w \,.\, T]. \tag{9.2.24}$$

For T of shape λ it turns out that V under this action is isomorphic to V_λ. The idea is that we are able to freely move about in a row, because we've symmetrised over rows in the Specht module, and we're able to move between rows at the cost of a sign, because we've antisymmetrised over columns in a Specht module, and here we have the sign appearing in the sum over the column group.

Elements of these modules are of the form xc_λ for some $x \in \Bbbk S_n$. The action of S_n on such an element is simply multiplication, $w \,.\, xc_\lambda = wxc_\lambda$. These are modules since the action is determined by the action on $\Bbbk S_n$, the c_λ is not involved since it is on the right.

Example 9.2.25 Consider S_3. There are three partitions, (3), $(1,1,1)$, and $(2,1)$. We've already seen $a_{(3)}$, $b_{((1,1,1))}$, $a_{(2,1)}$ and $b_{(2,1)}$. It's also clear that $a_{(1,1,1)} = ()$ and $b_{(3)} = ()$. Computing the projectors we have

$$c_{(3)} = \frac{1}{6}[() + (1\,2) + (1\,3) + (2\,3) + (1\,2\,3) + (1\,3\,2)] \tag{9.2.26}$$

$$c_{(1,1,1)} = \frac{1}{6}[() - (1\,2) - (1\,3) - (2\,3) + (1\,2\,3) + (1\,3\,2)] \tag{9.2.27}$$

$$c_{(2,1)} = \frac{1}{4}[() + (1\,2) - (1\,3) - (1\,3\,2)]. \tag{9.2.28}$$

We have the linear map $\mathbb{C}S_n \to \mathbb{C}S_n c_\lambda$ given by $x \mapsto xc_\lambda$. To compute the modules $\mathbb{C}S_n c_\lambda$ it is sufficient to look at the basis of $\mathbb{C}S_n$, which is of course just S_n. The image of the basis under $x \mapsto xc_\lambda$ is then a spanning set of $\mathbb{C}S_n c_\lambda$. Taking any maximal linearly independent subset of this spanning set then gives a basis of $\mathbb{C}S_n c_\lambda$.

Starting with $\lambda = (3)$ we can see that $wc_{(3)} = c_{(3)}$, thus $V_{(3)} = \operatorname{span}\{c_\lambda\}$ is a one-dimensional space. Since $wc_{(3)} = c_{(3)}$ for all $w \in S_3$ we can also see that S_n acts trivially on $V_{(3)}$ and so $V_{(3)}$ is the trivial representation. In general, $V_{(n)}$ is always the trivial representation of S_n.
Now consider $\lambda = (1,1,1)$. We have

$$()c_{(1,1,1)} = (1\,2\,3)c_{(1,1,1)} = (1\,3\,2)c_{(1,1,1)} \tag{9.2.29}$$

and

$$(1\,2)c_{(1,1,1)} = (1\,3)c_{(1,1,1)} = (2\,3)c_{(1,1,1)} = -c_{(1,1,1)}. \tag{9.2.30}$$

Thus, we have $V_{(1,1,1)} = \operatorname{span}\{c_{(1,1,1)}\}$, again a one-dimensional space. However, this time we have that $w \in S_3$ acts as its sign, since from the above we see that $wc_{(1,1,1)} = \operatorname{sgn}(w)c_{(1,1,1)}$. Thus, $V_{(1,1,1)}$ is the sign representation. In general, $V_{(1,\dots,1)}$ is always the sign representation of S_n.
Finally, consider $\lambda = (2,1)$. We then have

$$()c_{(2,1)} = (1\,2)c_{(2,1)} = c_{(2,1)}; \tag{9.2.31}$$

$$(1\,3)c_{(2,1)} = (1\,2\,3)c_{(2,1)} \tag{9.2.32}$$
$$= \frac{1}{4}[-() + (1\,3) - (2\,3) + (1\,2\,3)];$$

$$(2\,3)c_{(2,1)} = (1\,3\,2)c_{(2,1)} \tag{9.2.33}$$
$$= \frac{1}{4}[-(1\,2) + (2\,3) - (1\,2\,3) + (1\,3\,2)].$$

These are not all linearly independent, we have that

$$(2\,3)c_{(2,1)} = -c_{(2,1)} - (1\,3)c_{(2,1)}. \tag{9.2.34}$$

Thus, we have $V_{(2,1)} = \operatorname{span}\{c_{(2,1)}, (1\,3)c_{(2,1)}\}$, so this is a 2-dimensional representation. For the action of S_n on $V_{(2,1)}$ we can use relationships like Equation (9.2.34) and

$$(2\,3)(1\,3)c_{(2,1)} = (1\,2\,3)c_{(2,1)} = (1\,3)c_{(2,1)} \tag{9.2.35}$$

to compute that

$$\rho((2\,3)) = \begin{pmatrix} -1 & 0 \\ -1 & 1 \end{pmatrix} \tag{9.2.36}$$

when we use the ordered basis $\{c_{(2,1)} = (1,0)^{\mathsf{T}}, (1\,3)c_{(2,1)} = (0,1)^{\mathsf{T}}\}$. Note that the columns of the matrix are just the image of the basis vectors under the action of $(2\,3)$. Similar calculations give

$$\rho(()) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \rho((1\,2)) = \begin{pmatrix} 1 & -1 \\ 0 & -1 \end{pmatrix}, \quad \rho((1\,3)) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$
$$\rho((1\,2\,3)) = \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix}, \quad \text{and} \quad \rho((1\,3\,2)) = \begin{pmatrix} -1 & 1 \\ -1 & 0 \end{pmatrix}. \tag{9.2.37}$$

> It's the a straightforward calculation to check that this is indeed a representation of S_n, simply check that ρ defines a homomorphism $S_n \to \mathrm{GL}_2$. Of course, actually doing these calculations by hand for n much larger than 3 becomes very arduous pretty quickly, which is why a large chunk of my masters project was programming these calculations[a].
>
> [a] See https://github.com/WilloughbySeago/MphysReport for the report, and https://github.com/WilloughbySeago/MPhysProjectCode for the code.

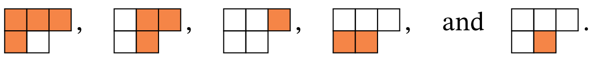
The bases of the Specht modules in the above example were $\{c_{(3)}\}$, $\{c_{(1,1,1)}\}$, and $\{c_{(2,1),(1\,3)c_{(2,1)}}\}$. It is actually possible to work out what these will be without having to do the calculations above. For a fixed shape, λ, there is always a basis of V_λ consisting of all wc_λ such that $w\,.\,T_0$ is a standard tableau. For the $n = 3$ case this corresponds to the only standard tableau being

$$\begin{array}{|c|c|c|}\hline 1&2&3\\ \hline\end{array}\,, \quad \begin{array}{|c|}\hline 1\\ \hline 2\\ \hline 3\\ \hline\end{array}, = , \quad \begin{array}{|c|c}\hline 1&\multicolumn{1}{c|}{2}\\ \hline 3 \\ \cline{1-1}\end{array}\,, \quad \text{and} \quad \begin{array}{|c|c}\hline 1&\multicolumn{1}{c|}{3}\\ \hline 2 \\ \cline{1-1}\end{array}\,. \tag{9.2.38}$$

The dimension of V_λ is thus the number of standard tableau of shape λ. That is,

$$f^\lambda := \dim V_\lambda = |\mathrm{SYT}(\lambda)|. \tag{9.2.39}$$

Fortunately, there is a nice rule for computing this number. First, we define the **hook length** of a box in a Young diagram to be the number of boxes to the right, plus the number of boxes below, plus one for the box itself. The idea is that this is the length of the "hooks" as depicted below for $\lambda = (3, 2)$:



(9.2.40)

The hook lengths of the corresponding boxes are then

$$\begin{array}{|c|c|c}\hline 4&3&\multicolumn{1}{c|}{1}\\ \hline 2&1 \\ \cline{1-2}\end{array}\,. \tag{9.2.41}$$

The **hook number** of a Young diagram, $\lambda⚳$, is then the product of the Hook lengths, so

$$\lambda⚳ = 4 \cdot 3 \cdot 2 \cdot 1 \cdot 1 = 24. \tag{9.2.42}$$

It is then a known fact that the number of semistandard tableaux of shape λ with n boxes, which is also the dimension of the S_n Specht module, V_λ, is given by the **hook length formula**

$$f^\lambda = \dim V_\lambda = |\mathrm{SYT}(\lambda)| = \frac{n!}{\lambda⚳} \tag{9.2.43}$$

9.2.4 Specht Modules are Simple

In this section we work over $\mathbb{C}$. Fix some positive integer, n, and partition, $\lambda \vdash n$.

We will show that the Specht modules, V_λ, are precisely the simple S_n-modules. The proof is pretty mechanical, and requires some lemmas and a bit more knowledge about Young diagrams first.

Lemma 9.2.44 For $g \in R_\lambda$ we have $a_\lambda g = g a_\lambda$, and for $g \in C_\lambda$ we have $b_\lambda g = \operatorname{sgn}(g) g b_\lambda$.

Lemma 9.2.45 For $x \in \mathbb{C}S_n$ we have $a_\lambda x b_\lambda = \ell_\lambda(x) c_\lambda$ where ℓ_λ is some linear function.

Proof. First note that if $g \in R_\lambda C_\lambda$ then $g = rc$ for some $r \in R_\lambda$ and $c \in C_\lambda$, and so $a_\lambda g b_\lambda = \operatorname{sgn}(c) c_\lambda$. To prove the statement we will show that if $g \notin R_\lambda C_\lambda$ then $a_\lambda g b_\lambda = 0$, since then we can take $\ell_\lambda(g) = \operatorname{sgn}(c)$ or $\ell_\lambda(g) = 0$ as appropriate, on the basis, S_n, to define $\ell_\lambda(x)$ on all of $\mathbb{C}S_n$.
To show that $a_\lambda g b_\lambda = 0$ for $g \notin R_\lambda C_\lambda$ it is sufficient to find some transposition, τ, such that $\tau \in R_\lambda$ and $g^{-1}\tau g \in C_\lambda$. Using the fact that a_λ is invariant under the action of R and C_λ acts on b_λ by the sign we have

$$a_\lambda g b_\lambda = a_\lambda \tau g b_\lambda = a_\lambda g g^{-1} \tau g b_\lambda = a_\lambda g (g^{-1}\tau g) b_\lambda = -a_\lambda g b_\lambda \tag{9.2.46}$$

which must mean that $a_\lambda g b_\lambda = 0$.
Finding such a transposition is equivalent to finding two elements in the same row of the tableau T, and in the same column of the tableau $g \,.\, T$. So, our goal is then equivalent to showing that if such a pair doesn't exist then $g \in R_\lambda C_\lambda$. That is, there exist some $r \in R$ and $c' \in C_{g.\lambda} = g C_\lambda g^{-1}$ such that $r \,.\, T = c' \,.\, (g \,.\, T)$, and then $g = rc^{-1}$ where $c = g^{-1} c' g \in C_\lambda$.
Any two elements of the first row of T are in different columns of $g \,.\, T$, so there exists some $c_1' \in C_{g.\lambda}$ such that all of these elements are in the first row. Thus, there is some $r_1 \in R_\lambda$ such that $r_1 \,.\, T$ and $c_1' \,.\, (g \,.\, T)$ have the same first row. Repeating this we can find $r_2 \in R_{r_1.\lambda}$ and $c_2' \in C_{c_1' g.T}$ such that $r_2 r_1 \,.\, T$ and $c_2' c_1' g \,.\, T$ have the same first two rows. Continuing on we will eventually construct the desired r and c', since this process will terminate eventually as the tableau has finitely many rows. ☐

Corollary 9.2.47 The Young projector, c_λ, is idempotent up to a scalar.

Proof. We have

$$c_\lambda^2 = a_\lambda b_\lambda a_\lambda b_\lambda = \ell_\lambda(b_\lambda a_\lambda) c_\lambda \tag{9.2.48}$$

for some scalar $\ell_\lambda(b_\lambda a_\lambda)$. ☐

Note that

$$\ell_\lambda(b_\lambda a_\lambda) = \frac{n!}{|R_\lambda||C_\lambda| \dim V_\lambda} = \frac{\lambda?}{|R_\lambda||C_\lambda|}. \tag{9.2.49}$$

Further, note that from c_λ we can construct an idempotent, $e = c_\lambda / \sqrt{\ell_\lambda(b_\lambda a_\lambda)}$, so long as $\ell_\lambda(b_\lambda a_\lambda) \neq 0$, which is true in this case.

Definition 9.2.50 — Lexicographic Ordering We define the **lexicographic order** on the set of partitions of n by declaring that $\lambda < \mu$ if for the smallest value of i such that $\lambda_i \neq \mu_i$ we have $\lambda_i < \mu_i$.

For example, consider the partitions of 5, under the lexicographic ordering we have

$$(1,1,1,1,1) < (2,1,1,1) < (2,2,1) < (3,1,1) < (3,2) < (4,1) < (5). \quad (9.2.51)$$

Note that this is the "dictionary order". When ordering two words we first compare their first two letters, if they're the same we move on to the second two letters, and so on. At the first pair of different letters we place first whichever word has the letter appearing earlier in the dictionary.

Lemma 9.2.52 If $\lambda > \mu$ in the lexicographic order then $a_\lambda \mathbb{C} S_n b_\mu = 0$.

Proof. Similarly to the previous lemma we just need to show that for any $g \in S_n$ there is some transposition, $\tau \in R_\lambda$ such that $g^{-1}\tau g \in C_\mu$. Let T be the canonical tableau of shape λ and T' the tableau we get if we act with g on the canonical tableau of shape μ. We claim that there are two entries in the same row of T and same column of T'. If $\lambda_1 > \mu_1$ this follows from the pigeonhole principle, there must be some element of the first row of T not in the first row of T', and thus we simply pick whatever element of the first row it sits below as our other element. If $\lambda_1 = \mu_1$ then as we did before we can find $r_1 \in R_\lambda$ and $c_1' \in Q_{g.\mu} = gQ_\mu g^{-1}$ such that $r_1 \,.\, T$ and $c_1' \,.\, T'$ have the same first row, then repeat the argument for the second row. Eventually, we will reach a row for which $\lambda_i > \mu_i$, since we have declared $\lambda > \mu$. □

Lemma 9.2.53 In any algebra, A, with an idempotent, e, any left A-module, M, is such that $\mathrm{Hom}_A(Ae, M) \cong eM$.

Proof. The desired isomorphism is $\varphi \colon eM \to \mathrm{Hom}_A(Ae, M)$, defined by $\varphi(m) = f_m \colon Ae \to M$ which is the morphism defined by $f_m(a) = a \,.\, m$. First note that elements of eM are of the form $e \,.\, m$ for some $m \in M$. Then eM is an A-module under the action $a \,.\, (e \,.\, m) = ae \,.\, m$. To see this first note that $ae =$
To show that this is well-defined we need to show that $f_m(a) = a \,.\, m$ really is an element of eM. That is, we need to show it is of the form $e \,.\, m'$ for some $m' \in M$. To do this we use the fact that $a \in Ae$, so $a = a'e$ for some $a' \in A$. Thus, we have $f_m(a) = f_m(a'e) = a'e \,.\, m$. This is the action of a' on $e \,.\, m$, and so
This is invertible, since given f_m we can recover m as $f_m(1) = 1 \,.\, m = m$.

This is an A-module homomorphism since

$$\begin{aligned}
\varphi(a\,.\,m)(a') &= f_{a.m}(a') && (9.2.54)\\
&= a'\,.\,(a\,.\,m) && (9.2.55)\\
&= a'a\,.\,m && (9.2.56)\\
&= f_m(a'a) && (9.2.57)\\
&= (a\,.\,f_m)(a') && (9.2.58)\\
&= (a\,.\,\varphi(m))(a'). && (9.2.59)
\end{aligned}$$

□

Theorem 9.2.60. The simple S_n-modules are precisely the Specht modules.

Proof. Corollary 9.2.47 tells us that c_λ is idempotent up to a scalar, so let e_λ be the idempotent we get by rescaling c_λ. Note then that $\mathbb{C}S_n c_\lambda = \mathbb{C}S_n e_\lambda$, since we can always absorb any scalar factor with the coefficients in $\mathbb{C}$.
Take two partitions, λ and μ, and without loss of generality suppose that $\lambda \geq \mu$ in the lexicographic order. We have that

$$\operatorname{Hom}_{S_n}(V_\lambda, V_\mu) = \operatorname{Hom}_{S_n}(\mathbb{C}S_n e_\lambda, \mathbb{C}S_n e_\mu) = e_\lambda \mathbb{C}S_n e_\mu \tag{9.2.61}$$

by Lemma 9.2.53 and its obvious left-analogue.
For $\lambda > \mu$ we have that

$$\dim(e_\lambda \mathbb{C}S_n e_\mu) = 0 \tag{9.2.62}$$

For $\lambda = \mu$ we have

$$\dim(e_\lambda \mathbb{C}S_n e_\lambda) = 1 \tag{9.2.63}$$

because tells us that $e_\lambda \mathbb{C}S_n e_\lambda$ is spanned by $c_\lambda g c_\lambda = a_\lambda b_\lambda g a_\lambda b_\lambda$, and by Lemma 9.2.45 we know that these elements are of the form $\ell_\lambda(b_\lambda g a_\lambda)c_\lambda$. We also have a flipped version of Lemma 9.2.45, which tells us that there is some linear function, ℓ'_λ such that $b_\lambda x a_\lambda = \ell'_\lambda(x) b_\lambda a_\lambda$. Applying this the spanning elements are all of the form

$$\ell_\lambda(b_\lambda g a_\lambda)c_\lambda = \ell_\lambda(\ell'_\lambda(g) b_\lambda a_\lambda)c_\lambda = \ell'_\lambda(g)\ell_\lambda(b_\lambda a_\lambda)c_\lambda. \tag{9.2.64}$$

Thus, all elements of $e_\lambda \mathbb{C}S_n e_\lambda$ are just a scalar multiple of c_λ, so this is a one-dimensional space.
From this we can apply Theorem 6.5.5, which tells us that

$$\langle \chi_\lambda, \chi_\mu \rangle = \delta_{\lambda\mu} \tag{9.2.65}$$

So, Theorem 6.5.5 tells us that $V_\lambda \ncong V_\mu$ for $\lambda \neq \mu$, and, Corollary 6.5.18 tells us that V_λ is simple.
To finish off the proof note that the number of simple modules is equal to the number of conjugacy classes (Corollary 6.2.9), and the conjugacy

classes of S_n are labelled by cycle type, which are themselves partitions of n. So, we have bijections

$$\{\text{Specht Modules}\} \overset{1:1}{\longleftrightarrow} \{\lambda \vdash n\} \overset{1:1}{\longleftrightarrow} \mathrm{Irr}(S_n). \tag{9.2.66}$$

Thus, we have exhausted the possible simple modules, so we know that all simple modules are isomorphic to some Specht module. $\square$

Ten

Branching Rules

10.1 Branching Rules

We have a natural embedding of symmetric groups

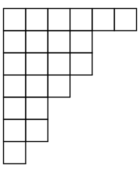
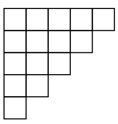
$$\{e\} = S_1 \hookrightarrow S_2 \hookrightarrow \dots \hookrightarrow S_{n-1} \hookrightarrow S_n \hookrightarrow \dots . \tag{10.1.1}$$

This allows us to view each symmetric group as a subgroup of any larger subgroup. Specifically, S_{n-1} can be viewed as the subgroup of S_n consisting of permutations of $\{1, \dots, n\}$ which leave n fixed[1].

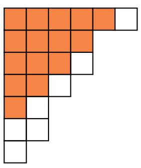
In terms of representations this means that any representation of S_{n-1} can be viewed as the restriction of some representation of S_n. Simply forget how any element that doesn't fix n acts. It turns out that the decomposition of such an S_{n-1}-module into simple S_{n-1}-modules is particularly simple (no pun intended). In a sense every "possible" simple S_{n-1}-module appears in the decomposition exactly once. What we mean by possible here is that when we take the Young diagram corresponding to the S_{n-1}-module it should fit inside the Young diagram corresponding to the S_n-module. We make this precise with the following definition.

> **Definition 10.1.2 — Skew Diagram** Let λ and μ be partitions such that $\mu_i \leq \lambda_i$ for all i. Then we may form the **skew diagram**, $\lambda \setminus \mu$, by placing both diagrams on top of each other and removing any boxes in the overlap.

For example, if we have

$\lambda =$  and $\mu =$  (10.1.3)

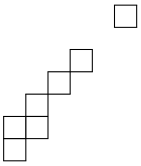
then overlapping these we have

 (10.1.4)

[1] We can fix any $k \in \{1, \dots, n\}$, the resulting subgroups are all conjugate, it's just that fixing n is the most "natural" choice.

so the corresponding skew diagram is

$$\lambda \setminus \mu = \tag{10.1.5}$$



Notice that it's possible to have entire rows missing, as we do here. This may include rows being cut off from the top or bottom of the diagram, but one should still imagine that they are there, they just have length zero.

If we're considering representations of S_n and S_{n-1} then λ must have n boxes and μ must have $n-1$ boxes, so $\lambda \setminus \mu$ (when it exits) must have 1 box.

Proposition 10.1.6 — Branching Rules Let V_λ be a simple S_n module, so $\lambda \vdash n$. Let[a] V_μ denote the simple S_{n-1}-modules, so $\mu \vdash n-1$. Then

$$\operatorname{Res}^{S_n}_{S_{n-1}} V_\lambda = \bigoplus_{\substack{\mu \vdash n-1 \\ |\lambda \setminus \mu| = 1}} V_\mu. \tag{10.1.7}$$

In particular, the restriction is multiplicity free.

[a] I think it's poor notation not to distinguish between S_n- and S_{n-1}-modules in a way that is immediately obvious, but V_λ is always an $S_{|\lambda|}$-module, so the notation is not ambiguous.

Proof. Step 1: Dimension Sum

An **inner corner** of a Young diagram is a box that we can remove and still have a (non-skew) Young diagram. Consider a standard Young tableau, T, of shape λ. Since T is standard n must appear in the right-most position of whichever row it is in. There must also be no box below the box containing n. This means that n is in an inner corner, and so we can remove it, to produce a Young tableau, T^-, with corresponding Young diagram λ^-. Further, T^- is still a standard Young tableau, now with $n-1$ boxes.

In reverse this process shows that any n-box standard Young tableau may be produced by starting with an $(n-1)$-box standard Young tableau and adding a single box labelled n. Thus, the number of n-box standard Young tableau of shape λ is precisely the sum of the number of standard Young tableau of shape λ^- as λ^- ranges over all Young diagrams we can produce by removing a single box from λ. That is,

$$f^\lambda = \sum_{\lambda^-} f^{\lambda^-}. \tag{10.1.8}$$

Another way of phrasing that λ^- is λ with a box removed is saying that we're considering all μ such that $\lambda \setminus \mu$ has precisely one box, so

$$f^\lambda = \sum_{\substack{\mu \vdash n-1 \\ |\lambda \setminus \mu| = 1}} . \tag{10.1.9}$$

Step 2: Module Sum

We now want to "categorify" this result. That is, we take the numerical sum,

$$f^\lambda = \sum_{\substack{\mu \vdash n-1 \\ |\lambda \setminus \mu| = 1}} f^\mu, \tag{10.1.10}$$

and we replace the f^λ with objects in some category and the sum with the coproduct. We've already seen that $f^\lambda = \dim V_\lambda$, so the correct choice of objects is the modules, V_λ, and the coproduct is then the direct sum. Making this replacement on the right we get

$$\bigoplus_{\substack{\mu \vdash n-1 \\ |\lambda \setminus \mu| = 1}} V_\mu. \tag{10.1.11}$$

We just have to show that this really does correspond to $\operatorname{Res}^{S_n}_{S_{n-1}} V_\lambda$.
To do this first let $r_1 < \cdots < r_k$ be the row numbers for the rows which end with an inner corner. Write λ^i for the Young diagram produced by removing the box at the end of row r_i. Similarly, if T is a standard Young tableau with n placed in the inner corner of row r_i then write T^i for the standard Young tableau given by removing this box.
We will construct a flag of vector spaces

$$0 = V_0 \subseteq V_1 \subseteq \cdots \subseteq V_k = V_\lambda. \tag{10.1.12}$$

It is not a coincidence that the maximum index chosen here, k, corresponds to the maximum index of the r_i before. We will do this in such a way that at each step we have $V_i / V_{i+1} \cong V_{\lambda^i}$ as S_{n-1}-modules. Then we will have that

$$\operatorname{Res}^{S_n}_{S_{n-1}} V_\lambda = V_k \cong V_{k-1} \oplus (V_k / V_{k-1}) \cong V_{k-1} \oplus V_{\lambda^k}. \tag{10.1.13}$$

Similarly, we'll have $V_{k-1} \cong V_{k-2} \oplus V_{\lambda^{k-1}}$, and so

$$\operatorname{Res}^{S_n}_{S_{n-1}} V_\lambda \cong V_{k-2} \oplus V_{\lambda^{k-1}} \oplus V_{\lambda^k}. \tag{10.1.14}$$

Continuing on, since the dimension is finite and so our flag has finite length, this process will eventually terminate, and we'll have the desired isomorphism.
All we have to do then is construct such a flag. Let M_λ denote the set of Young tabloids of shape λ. Define a map $\theta_i : M_\lambda \to M_{\lambda^i}$ to be removing n from row r_i if its present, and zero otherwise. So $\theta_i(\{T\}) = \{T^i\}$ if n is in row r_i, and $\theta_i(\{T\}) = 0$ otherwise. These are morphisms of S_{n-1}-modules, since S_{n-1} always fixes the box labelled n and thus the action of S_{n-1} commutes with removing the box labelled n.
Similarly, we can extend θ_i to a map $V_\lambda \to V_\lambda$ by defining $\theta_i(E_T) = E_{T^i}$ if n is in the row r_i, and $\theta_i(E_T) = 0$ if n appears in row r_j with $j < i$. We shall not need the case where $j > i$, so any definition will work there. This is well-defined since any column group action that moves n from the current row will result in a vanishing term in the expression of E_T. The

only permutations of the column group which don't send the tabloid to zero under θ_i are precisely those which fix the row of n, which means that this subgroup of the column group is precisely C_{T^i}.
Note that all standard tabloids, $E_{T^i} \in V_{\lambda^i}$ are in the image of θ_i. Further, all of these E_T have their n in row r_i, and thus we may define V_i to be the spae spanned by the E_{T^i}. Then $\theta_i(V_i) = V_{\lambda^i}$ as required. If T instead has its n above row r_i then $\theta_i(E_T) = 0$, and thus $V_{i-1} \subseteq \ker\theta_i$. This gives us the chain

$$0 = V_0 \subseteq V_1 \cap \ker\theta_1 \subseteq V_1 \subseteq \cdots \subseteq V_k \cap \ker\theta_k \subseteq V_k = V_\lambda. \tag{10.1.15}$$

We also have

$$\dim(V_i/(V_i \cap \ker\theta_i)) = \dim(\theta_i(V_i)) = \dim V_{\lambda^i} = f^{\lambda^i}. \tag{10.1.16}$$

Thus, the steps from $V_i \cap \ker\theta_i$ to V_i give us all the f^{λ^i} as we add up the dimensions. Thus, by we have accounted for all of $f^\lambda = \dim V_\lambda$ by Equation (10.1.9). Thus, the containment $V_{i-1} \subseteq V_i \cap \ker\theta_i$ is actually an equality, and so we have

$$\frac{V_i}{V_{i-1}} = \frac{V_i}{V_i \cap \ker\theta_i} \cong V_{\lambda^i} \tag{10.1.17}$$

as claimed. □

Corollary 10.1.18 With notation as in Proposition 10.1.6 we have

$$\mathrm{Ind}_{S_{n-1}}^{S_n} V_\mu = \bigoplus_{\substack{\lambda\vdash n\\ |\lambda\setminus\mu|=1}} V_\lambda. \tag{10.1.19}$$

Proof. By Frobenius reciprocity for an arbitrary irreducible character, χ_ν, of S_{n-1} we have

$$\langle \chi_\lambda \downarrow_{S_{n-1}}^{S_n}, \chi_\nu \rangle = \langle \chi_\lambda, \chi_\nu \uparrow_{S_{n-1}}^{S_n} \rangle. \tag{10.1.20}$$

This tells us that the multiplicity of V_ν in $\mathrm{Res}_{S_{n-1}}^{S_n} V_\lambda$ is the same as the multiplicity of V_λ in $\mathrm{Ind}_{S_{n-1}}^{S_n} V_\nu$, which is 1 if removing a box from λ gives ν and zero otherwise, and so the result follows. □

Example 10.1.21 Consider the S_3-module $V_{\ydiagram{2,1}}$ The branching rule tells us that

$$\mathrm{Res}_{S_2}^{S_3} V_{\ydiagram{2,1}} = V_{\ydiagram{1,1}} \oplus V_{\ydiagram{2}}. \tag{10.1.22}$$

Similarly, for the S_2-module $V_{\square\square}$ the branching rules tell us that

$$\operatorname{Ind}_{S_2}^{S_3} V_{\square\square} = V_{\square\square\square} \oplus V_{\substack{\square\square\\\square}}. \tag{10.1.23}$$

10.2 Gelfand–Zetlin Basis

Repeatedly applying the decomposition provided by the branching rules we can repeatedly restrict an S_n-module to an S_{n-1}-module, which we can restrict to an S_{n-2}-module, and so on, until we've restricted all the way down to an S_0-module, which is just a vector space.

At each step in the process we sum over all Young diagrams which can be obtained by removing just a single box. Reversing this, a fixed Young diagram, λ, can be thought of as being built up from single boxes. If we number the boxes in the order we add them, making sure that at each step we have a valid Young diagram, then we will end up with a Young tableau of shape λ labelled with the numbers 1 through n. Further, each row will be increasing, we cannot add to the end of a row before we have built up the start of the row, and so will each column for the same reason. Thus, the tableau we're left with will be standard.

This gives us a nice interpretation of standard Young tableau as paths in the **Young lattice**. This lattice has all Young diagrams as elements, and we declare $\lambda < \mu$ if $\mu_i \le \lambda_i$ for all i. That is, $\lambda < \mu$ if the Young diagram of μ fits entirely within the Young diagram of λ. Pictorially, this gives us Figure 10.1. Then a standard tableaux of shape λ corresponds to a path in this lattice from the diagram of λ to the empty partition[2], $\varnothing$, only travelling downwards.

[2] The empty partition, $\varnothing$, is the unique partition of 0, that is $(0, 0, \dots)$.

For example, two of the four standard tableaux of shape $(3, 2)$ correspond to the paths drawn in Figure 10.2.

This shows that in the repeated-restriction process above we get all standard tableau of shape λ appearing in the decomposition of the S_n-module restricted to an S_0-module. That is, we have as modules

$$\operatorname{Res}_{S_0}^{S_n} V_\lambda = \bigoplus_{T \in \mathrm{SYT}(\lambda)} V_T \tag{10.2.1}$$

where V_T are 1-dimensional vector spaces. Note that as vector spaces $\operatorname{Res}_{S_0}^{S_n} V_\lambda = V_\lambda$, which gives us the following result.

Definition 10.2.2 — Gelfand–Zetlin Basis The process detailed above defines, up to normalisation, a basis of V_λ, known as the **Gelfand–Zetlin basis**. Specifically, we let $V_T = \Bbbk v_T$ then $\{v_T \mid T \in \mathrm{SYT}(\lambda)\}$ is a basis of V_λ.

Suppose that $\operatorname{char} \Bbbk \nmid n!$, so that $\Bbbk S_n$ is semisimple. Then we have the decomposition

$$\Bbbk S_n \cong \bigoplus_{\lambda \vdash n} \operatorname{End}(V_\lambda) \cong \bigoplus_{\lambda \vdash n} \operatorname{Mat}_{\dim V_\lambda}(\mathbb{C}). \tag{10.2.3}$$

We can thus specify a subalgebra of $\Bbbk S_n$

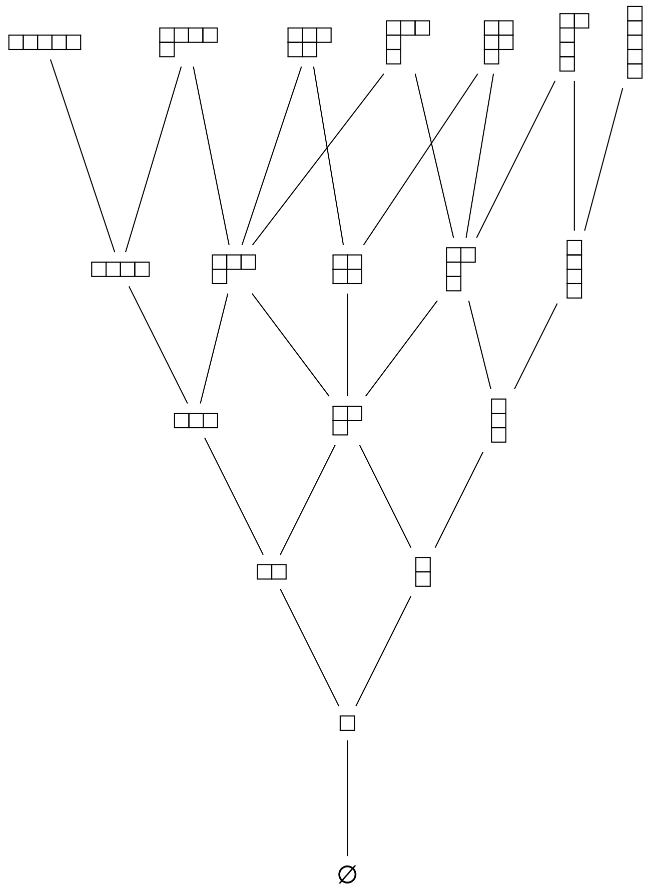


Figure 10.1: The Young lattice.

Definition 10.2.4 — Gelfand–Zetlin Subalgebra The **Gelfand–Zetlin subalgebra**, $A_n \subseteq \Bbbk S_n$, is the subalgebra consisting of elements whose action is diagonal in all irreducible representations.

That is, the Gelfand–Zetlin subalgebra consists of all elements of $\Bbbk S_n$ which correspond to a direct sum of diagonal matrices in the above decomposition.

Lemma 10.2.5 The Gelfand–Zetlin subalgebra is a maximal commutative subalgebra of $\Bbbk S_n$. Further, the Gelfand–Zetlin subalgebra is semisimple.

The Gelfand–Zetlin basis element, v_T, corresponds to the one-dimensional irreducible A_n-module, $V_T = \Bbbk v_T$.

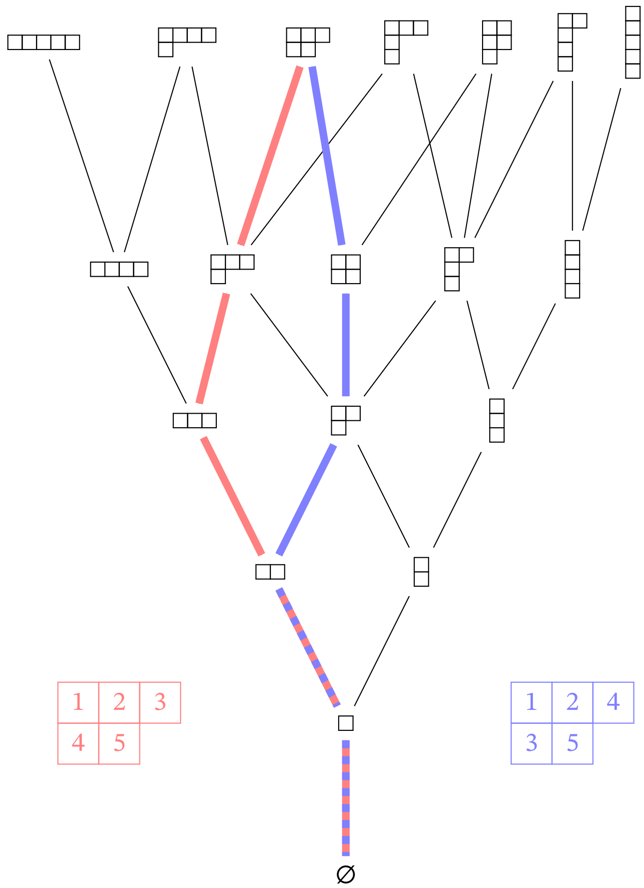


Figure 10.2: Paths in a Young lattice and the corresponding standard tableaux.

10.3 Jucys–Murphy Elements

Definition 10.3.1 — Jucys–Murphy Elements The jth **Jucys–Murphy element** of $\Bbbk S_n$ is

$$L_j := \sum_{1 \le i < j} (i\,j). \tag{10.3.2}$$

Note that $L_1 = 0$ is the empty sum.

Note that

$$L_n = (1\,n) + (2\,n) + \cdots + (n-1\,n) \tag{10.3.3}$$

commutes with all of $\Bbbk S_{n-1}$, since elements of $\Bbbk S_{n-1}$ fix n and so if $w \in S_{n-1}$ then $L_n w$ is just L_n with the order of the terms in the sum rearranged.

This means that the Jucys–Murphy elements generate a commutative subalgebra of $\Bbbk S_n$.

Lemma 10.3.4 The Gelfand–Zetlin subalgebra, A_n, is generated by either

- $Z_0, \dots, Z_n \subseteq \Bbbk S_n$ for $Z_i = Z(\Bbbk S_n)$; or
- $L_1, \dots, L_n$.

10.4 Young's Seminormal Form

For λ a partition of n fix some vector $v_{T_0} \in V_\lambda$ where T_0 is the canonical tableau of shape λ. Let T be some standard tableau of shape λ. Define $w_T \in S_n$ by $T = w_T.T_0$, where S_n acts on T by permuting the boxes according to their numbering. Then we may define $v_T = \pi_T(w_T \,.\, v_{T_0}) \in V_\lambda = V_T$ where

$$\pi_T : \bigoplus_{S \in \mathrm{SYT}(\lambda)} \twoheadrightarrow V_T \tag{10.4.1}$$

is projection onto the corresponding term of the direct sum.

Theorem 10.4.2. The simple transpositions, $s_i = (i\, i+1)$, act on V_λ in such a way that

$$s_i \,.\, v_T = \begin{cases} v_{s_i.T} & \text{if } s_i \,.\, T \text{ is a standard tableau;} \\ 0 & \text{else.} \end{cases} \tag{10.4.3}$$

Define $c_T(k) = j - i$ when $T(i,j) = k$. Then

$$s_i \,.\, v_T = \frac{1}{c_T(i+1) - c_T(i)} v_T + \left(1 + \frac{1}{c_T(i+1) - c_T(i)}\right) v_{s_i.T} \tag{10.4.4}$$

and

$$L_j \,.\, v_T = c_T(j) v_T. \tag{10.4.5}$$

Eleven

Symmetric Functions

11.1 Kostka Numbers

Recall that for a partition, $\lambda \vdash n$, the row group of λ is $S_\lambda \cong S_{\lambda_1} \times \cdots \times S_{\lambda_\ell}$ where S_{λ_1} acts on $\{1, \dots, \lambda_1\}$, S_{λ_2} acts on $\{\lambda_1 + 1, \dots, \lambda_1 + \lambda_2\}$, and so on. Consider the trivial representation of S_λ, $\mathbb{C}$. We can define an S_n-module by inducing this up:

$$M_\lambda \coloneqq \operatorname{Ind}_{S_\lambda}^{S_n} \mathbb{C}. \tag{11.1.1}$$

Lemma 11.1.2 With notation as above we have $M_\lambda \cong \mathbb{C}S_n a_\lambda$.

Recall that if e is an idempotent of the algebra A then

$$\operatorname{Hom}_A(Ae, M) \cong eM \tag{11.1.3}$$

for any left A-module, M. We then have

$$\operatorname{Hom}_{S_n}(M_\lambda, V_\mu) = \operatorname{Hom}_{S_n}(\mathbb{C}S_n a_\lambda, V_\mu) \cong a_\lambda V_\mu = a_\lambda \mathbb{C}S_n b_\mu a_\mu. \tag{11.1.4}$$

We also have that

$$\dim(a_\lambda \mathbb{C}S_n b_\mu a_\mu) = \begin{cases} 1 & \lambda = \mu, \\ 0 & \mu < \lambda. \end{cases} \tag{11.1.5}$$

Definition 11.1.6 — Weight Let λ be a partition of n. Let μ be a sequence of nonnegative integers, $\mu = (\mu_1, \mu_2, \dots)$ such that $\sum_i \mu_i = n$ (so a partition minus the requirement that the μ_i be weakly decreasing). We call such a μ a **composition** of n. We say that a semi-standard Young tableau of shape λ has weight μ if $i \in \{1, \dots, n\}$ appears in the labelling of boxes μ_i times.

Definition 11.1.7 — Kostka Numbers Let λ be a partition of n and μ a composition of n. The **Kostka numbers**, $K_{\lambda\mu}$, are the number of semistandard tableaux of shape λ and weight μ.

Writing $\mathrm{SSYT}(\lambda, \mu)$ for the set of semistandard Young tableaux of shape λ and weight μ we have that

$$K_{\lambda\mu} = |\mathrm{SSYT}|(\lambda, \mu). \tag{11.1.8}$$

Example 11.1.9 Suppose $\lambda = (3, 2)$ and $\mu = (1, 1, 2, 1)$. Then $K_{\lambda\mu}$ is the number of semistandard tableaux of shape $(3, 2)$ filled with one 1, one 2, two 3s, and one 3. It's not hard to check that the only options are

$$\begin{array}{|c|c|c|}\hline 1&2&3\\\hline 3&4\\\cline{1-2}\end{array}, \quad \begin{array}{|c|c|c|}\hline 1&2&4\\\hline 3&3\\\cline{1-2}\end{array}, \quad \text{and} \quad \begin{array}{|c|c|c|}\hline 1&3&3\\\hline 2&4\\\cline{1-2}\end{array}. \tag{11.1.10}$$

Thus, $K_{(3,2)(1,1,2,1)} = 3$.
Any partition, λ, is also a composition. In general, we have $K_{\lambda\lambda} = 1$, since the only way to fill a Young diagram of shape λ with λ_1 1s, λ_2 2s, and so on in such a way that the result is semistandard is to have the first row filled with 1s, the second row filled with 2s, and so on.
If $\mu = (1, 1, \ldots, 1)$ with n 1s then every number from 1 to n appears exactly once, and being semistandard is the same as being standard. Thus,

$$K_{\lambda(1,1,\ldots,1)} = f^{\lambda} = |\mathrm{SYT}(\lambda)| = \dim V_{\lambda} = \frac{n!}{\lambda?}. \tag{11.1.11}$$

Proposition 11.1.12 With notation as above we have that

$$M_{\lambda} = V_{\lambda} \oplus \bigoplus_{\mu > \lambda} K_{\mu\lambda} V_{\mu}. \tag{11.1.13}$$

11.2 Frobenius Character Formula for S_n

Consider the ring of polynomials in n-commuting indeterminates, $\mathbb{C}[x_1, \ldots, x_n]$. There is a natural action of the symmetric group, S_n, on this ring, specifically

$$(w\,.\,f)(x_1, \ldots, x_n) = f(x_{w^{-1}(1)}, \ldots, x_{w^{-1}(n)}). \tag{11.2.1}$$

Note that the action is defined in terms of w^{-1} simply because this is what gives us a left action.

Some polynomials in $\mathbb{C}[x_1, \ldots, x_n]$ are left invariant under this action. That is, if we permute the variables the polynomial doesn't change. The following are some examples of this in three variables:

$$xyz, \quad xy + xz + yz, \quad x + y + z, \quad x^2y + x^2z + y^2x + y^2z + z^2x + z^2y. \tag{11.2.2}$$

Notation 11.2.3 — Fixed Points Let X be a set with some specified action of a group, G. Write X^G for the set of fixed points of X under this action. That is,

$$X^G = \{x \in X \mid g\,.\,x = x \forall g \in G\}. \tag{11.2.4}$$

We will now study $\Lambda_n \coloneqq \mathbb{C}[x_1, \ldots, x_n]^{S_n}$, and the generalisation of this to infinitely many variables. We call elements of Λ_n **symmetric polynomials** in n variables. First, notice that the product of two such polynomials is once again symmetric, as is their sum. Thus, Λ_n is a ring. Further, if any complex multiple of a symmetric function is again symmetric, and thus Λ_n is a $\mathbb{C}$-algebra.

Definition 11.2.5 — Power Sums Let $r \in \mathbb{Z}_{\geq 0}$, and define the **power sum**, $p_r \in \mathbb{C}[x_1, \dots, x_n]$ by

$$p_r(x_1, \dots, x_n) \coloneqq \sum_{i=1}^{n} x_i^r. \tag{11.2.6}$$

First note that p_r is a symmetric polynomial. Permuting the variables just rearranges the order of the x_i^r in the sum, which doesn't change the polynomial.

It turns out that the power sums generate all of Λ_n.

Proposition 11.2.7 We have

$$\mathbb{C}[x_1, \dots, x_n]^{S_n} \cong \mathbb{C}[p_1, \dots, p_n]. \tag{11.2.8}$$

Note that we are not considering $\mathbb{C}[p_1, \dots, p_n]$ to be a polynomial ring. Instead it is simply all polynomials in the p_r subject to the relations that follow from the definition of the p_r in terms of the x_i.

Consider the examples of Equation (11.2.2). These are p_3, p_2, p_1, and $p_2 p_1 - p_3$ respectively.

Notation 11.2.9 Let λ be a partition of n, and suppose $N \geq \ell(\lambda)$ (which you'll recall is the number of nonzero terms in λ). Then we write

$$x^\lambda = x_1^{\lambda_1} x_2^{\lambda_2} \cdots x_N^{\lambda_N}. \tag{11.2.10}$$

Note that λ_i may be zero for some of these exponents. We also write

$$p_\lambda = p_{\lambda_1} p_{\lambda_2} \cdots p_{\lambda_{\ell(\lambda)}}. \tag{11.2.11}$$

Note that this is the same as if we carry on all the way to x_N, since $p_0 = 1$. We also define the sum of partitions in the obvious way, so $(\lambda+\rho)_i = \lambda_i+\rho_i$. The antisymmetric polynomial

$$\Delta(x) = \prod_{1 \leq i < j \leq N} (x_i - x_j) \tag{11.2.12}$$

is called the **van der Monde determinant**.

Proposition 11.2.13 Let λ be a partition of n. For $N \geq \ell(\lambda)$ we have the following relationship between characters and symmetric polynomials in N variables:

- $\chi_{M_\lambda}(\mu)$ is the coefficient of x^λ in p_μ; and
- $\chi_{V_\lambda}(\mu)$ is the coefficient of $x^{\lambda+\rho}$ in $\Delta(x) p_\mu$.

Note that $\chi_X(\mu)$ means the character of any element of the conjugacy class labelled by cycle-type μ in the representation X.

11.3 The Ring of Symmetric Functions

The previous result suggests that there is a close relationship between the representation theory of S_n and symmetric polynomials. One thing that gets in the way when we try to utilise this connection is that we always have to have "enough" variables. In the previous result this meant we had $N \geq \ell(\lambda)$. However, most things we can say about symmetric functions are fairly independent of the number of variables. The way we get around this is to consider an infinite number of variables. This takes a bit of care to set up properly, but then we can go back to thinking of the resulting elements as being symmetric polynomials in sufficiently many variables after we've put in the work upfront.

11.3.1 Construction

For this section we work with polynomials over $\mathbb{Z}$. This can then be extended to $\mathbb{C}$ by extension of scalars later.

Notice that

$$\Lambda_N = \mathbb{Z}[x_1, \dots, x_N]^{S_N} \tag{11.3.1}$$

is a graded ring, specifically,

$$\Lambda_N = \bigoplus_{d\geq 0} \Lambda_N^d \tag{11.3.2}$$

where Λ_N^d is the $\mathbb{Z}$-submodule of Λ_N consisting of homogeneous symmetric polynomials of degree d.

Let $\lambda = (\lambda_1 \geq \dots \geq \lambda_N \geq 0)$ be a partition of length at most N with $|\lambda| = d$. We define the **monomial symmetric polynomial** corresponding to λ to be

$$m_\lambda(x_1, \dots, x_N) = \sum_\alpha x^\alpha \tag{11.3.3}$$

where the sum is over all α which are given by permuting the first N terms of λ. For example, if $\lambda = (3, 2)$ and $N = 3$ then the permutations of the first three terms of λ are

$$(3,2,0), \quad (3,0,2), \quad (2,3,0), \quad (2,0,3), \quad (0,3,2), \quad \text{and} \quad (0,2,3). \tag{11.3.4}$$

Thus, we have

$$\begin{aligned} &m_{(3,2)}(x_1, x_2, x_3) \\ &= x_1^3x_2^2x_3^0 + x_1^3x_2^0x_3^2 + x_1^2x_2^3x_3^0 + x_1^2x_2^0x_3^3 + x_1^0x_2^3x_3^2 + x_1^0x_2^2x_3^3 \\ &= x_1^3x_2^2 + x_1^3x_3^2 + x_1^2x_2^3 + x_1^2x_3^3 + x_2^3x_3^2 + x_2^2x_3^3. \end{aligned} \tag{11.3.5}$$

Note that $m_{(r)} = p_r$. This definition makes sense so long as $N \geq \ell(\lambda)$, so if we want to consider all degree d polynomials then we should take $N \geq d$, which we'll assume from now on. Under these considerations the m_λ form a basis for Λ_N^d.

For $N' \geq N$ there is a surjection

$$\rho_{N',N}^d : \Lambda_{N'}^d \twoheadrightarrow \Lambda_N^d \tag{11.3.6}$$

defined by setting $x_{N+1} = \dots = x_{N'} = 0$. The action of this map on the monomial symmetric polynomials is

$$\rho_{N',N}^d(m_\lambda(x_1, \dots, x_{N'})) = \begin{cases} m_\lambda(x_1, \dots, x_N) & \ell(\lambda) \leq N; \\ 0 & \text{otherwise.} \end{cases} \tag{11.3.7}$$

Further, note that the map $\rho^d_{N',N}$ is bijective for $N' \geq N \geq d$.

We then have a sequence of bijections

$$\Lambda^d_1 \twoheadleftarrow \Lambda^d_2 \twoheadrightarrow \Lambda^d_3 \twoheadrightarrow \cdots . \tag{11.3.8}$$

This is an inverse system, by which we mean that $\rho^d_{i,k} = \rho^d_{i,j} \circ \rho^d_{j,k}$ for all $i, j, k \in \mathbb{Z}_{>0}$. Define the ring of homogeneous functions of degree d to be the inverse limit

$$\Lambda^d = \varprojlim_N \Lambda^d_N. \tag{11.3.9}$$

That is, elements of Λ^d are sequences, $(f_N)_{N\in\mathbb{Z}_{>0}}$, where each f_N is a homogeneous degree d polynomial in N variables. These sequences are (by definition of the inverse limit) such that

$$f_{N+1}(x_1, \dots , x_N, 0) = f_N(x_1, \dots , x_N). \tag{11.3.10}$$

The projections, sending such a sequence to its Nth term,

$$\mathrm{proj}^d_N : \Lambda^d \twoheadrightarrow \Lambda^d_N, \tag{11.3.11}$$

$$f = (f_N)_{N\in\mathbb{Z}_{>0}} \mapsto f_N, \tag{11.3.12}$$

are isomorphisms for $N \geq d$. This means that Λ^d is a free $\mathbb{Z}$-module with basis $\{m_\lambda \mid \lambda \vdash d\}$.

We define the **ring of symmetric functions** to be the graded ring

$$\Lambda = \bigoplus_{d\geq 0} \Lambda^d. \tag{11.3.13}$$

Remark 11.3.14 Note that *technically* elements of Λ are not polynomials, they are infinite sequences of polynomials. However, we can pretty much treat them as polynomials most of the time, just take some polynomial sufficiently far along in the sequence that there are enough variables to do whatever it is we're trying to do. To make this distinction we call elements of Λ "symmetric functions" instead of "symmetric polynomials", but we pretty much think of them as polynomials.

Let f_N and f_{N+1} be symmetric polynomials in N and $N + 1$ variables respectively such that

$$f_N(x_1, \dots , x_N) = f_{N+1}(x_1, \dots , x_N, 0). \tag{11.3.15}$$

If it's possible to make definitions of a family of polynomials in this way such that at each step adding a new variable and setting it to zero doesn't change anything then it makes sense to consider $(f_N)_{N\in\mathbb{Z}_{>0}}$ as an element of Λ. We call this the projective limit of f (where f is some label referring to this whole family of polynomials, which we really want to think of as all being the same symmetric function).

Remark 11.3.16 There are several constructions of Λ. The one we've given makes it an inverse limit in the category of graded rings. There is an alternative construction which makes it a direct limit in the category of rings of the direct system

$$\Lambda_1^d \hookrightarrow \Lambda_2^d \hookrightarrow \Lambda_3^d \hookrightarrow \tag{11.3.17}$$

where the inclusions are defined in terms of another basis of polynomials, called the elementary symmetric polynomials, e_r, and the maps defined by $e_r(x_1, \dots, x_n) \mapsto e_r(x_1, \dots, x_n, x_{n+1})$.
As categorical duals an inverse limit is some subset of the product, and the direct limit is some the disjoint union modulo some equivalence relation. There are benefits to both constructions. Elements of inverse limits are slightly easier to work with, because we don't have to keep track of the equivalence relation and worry if things are well-defined. Conversely, with the direct limit definition we can directly (no pun intended) identify (equivalence classes) of elements with elements of some object in the direct system.
Once we place a grading on the ring of symmetric functions as defined in terms of a direct limit it is isomorphic to the ring of symmetric functions as defined in terms of an inverse limit.
Since we won't have much reason to worry about the exact structure of elements of this ring we won't worry any more about exactly how it's defined.

Let Λ be the ring of symmetric functions with integer coefficients. Then for any ring, R, we can define $\Lambda_R = \Lambda \otimes_{\mathbb{Z}} R$, to be the ring of symmetric functions with coefficients in R. In particular, $\Lambda_{\mathbb{C}}$ is the ring of symmetric functions with coefficients in $\mathbb{C}$.

Proposition 11.3.18 We have that

$$\Lambda_{\mathbb{C}} \cong \mathbb{C}[p_1, p_2, \dots] \tag{11.3.19}$$

where the p_r are the projective limits of the power sums.

11.4 Schur Functions

Definition 11.4.1 — Schur Polynomial Let λ be a partition of length N. We define the corresponding **Schur polynomial** to be

$$s_\lambda(x) := \frac{\det(x_i^{\lambda_j + N - j})_{1 \le i,j \le N}}{\det(x_i^{N-j})_{1 \le i,j \le N}}. \tag{11.4.2}$$

Schur polynomials are stable, in the sense that if s_λ is the Schur polynomial in N variables, and $\hat{s}_\lambda$ is the Schur polynomial in $N+1$ variables then

$$s_\lambda(x_1, \dots, x_N) = \hat{s}_\lambda(x_1, \dots, x_N, 0). \tag{11.4.3}$$

In practice both of these are denoted s_λ with the number of variables distinguishing them. Since s_λ is unchanged by adding more variables the setting them to zero we can consider the projective limit of the Schur polynomials.

Definition 11.4.4 — Schur Function Let λ be a partition. The corresponding **Schur function** is the projective limit of the Schur polynomials s_λ as we increase the number of variables.

Theorem 11.4.5. The Schur functions form a $\mathbb{Z}$-basis, $\{s_\lambda \mid \lambda \vdash n, n \in \mathbb{Z}_{\geq 0}\}$, of Λ.

By a $\mathbb{Z}$-basis we mean that any symmetric function with integer coefficients can be expressed as a linear combination of Schur functions with integer coefficients. In other words, the s_λ form a basis of $\Lambda = \Lambda_{\mathbb{Z}}$. This is in contrast to the p_λ which form only a $\mathbb{Q}$-basis of $\Lambda_{\mathbb{Q}}$.

With these constructions we can restate the Frobenius formula as

$$p_\mu = \sum_\lambda \chi_{V_\lambda}(\mu) s_\lambda. \tag{11.4.6}$$

11.5 Macdonald's Characteristic Map

For $w \in S_n$ we denote by $\mu(w) = (\mu_1 \geq \mu_2 \geq \cdots \geq \mu_n \geq 0)$ the ordered cycle length of w. So, for example, in S_5 if $w = (1\,2\,3)(4\,5)$ then $\mu(w) = (3, 2)$. We can then define a map

$$\psi\colon S_n \to \Lambda \tag{11.5.1}$$

$$w \mapsto p_{\mu(w)} = p_{\mu_1} \cdots p_{\mu_n}. \tag{11.5.2}$$

Since this map is defined only by the cycle type of w we have that $\psi(w) = \psi(w')$ whenever w and w' are conjugate.

We have the obvious embedding $S_m \times S_n \hookrightarrow S_{m+n}$, in which $w \times w' \mapsto u$ defined by

$$u(i) = \begin{cases} w(i) & i \in \{1, \dots, m\}; \\ w'(i) & i \in \{m+1, \dots, m+n\}. \end{cases} \tag{11.5.3}$$

Then we have

$$\psi(w \times w') = \psi(w)\psi(w'). \tag{11.5.4}$$

Recall that $\mathcal{X}_n = \mathcal{X}_n(S_n)$ is the space of class functions, $S_n \to \mathbb{C}$, and that this is spanned by the irreducible characters $\{\chi_\lambda \mid \lambda \vdash n\}$.

We may then consider the (vector space) direct sum

$$\mathcal{X} = \bigoplus_{n \geq 0} \mathcal{X}_n \tag{11.5.5}$$

where $\mathcal{X}_0 = \mathbb{C}$. This graded vector space can be made into a graded ring by defining multiplication of homogeneous basis elements:

$$\chi_\lambda * \chi_\mu := (\chi_\lambda \times \chi_\mu) \uparrow_{S_m \times S_n}^{S_{m+n}}. \tag{11.5.6}$$

In words, we define the product of irreducible characters to be the induced character of the representation arising from the obvious embedding $S_m \times S_n \hookrightarrow S_{m+n}$.

Any $f, g \in \mathcal{X}$ may be expanded as a sum of $f_n, g_n \in \mathcal{X}_n$:

$$f = \sum_{n \geq 0} f_n, \quad \text{and} \quad g = \sum_{n \geq 0} g_n. \tag{11.5.7}$$

Recall that we've defined an inner product, $\langle -, - \rangle_{S_n} : \mathcal{X}_n \times \mathcal{X}_n \to \mathbb{C}$. We can extend this to an inner product, $\langle -, - \rangle : \mathcal{X} \times \mathcal{X} \to \mathbb{C}$, by defining

$$\langle f, g \rangle = \sum_n \langle f_n, g_n \rangle_{S_n}. \tag{11.5.8}$$

This is the obvious extension given by declaring that the different homogeneous subspaces are orthogonal.

Definition 11.5.9 — Macdonald's Characteristic Map **Macdonald's characteristic map** is the map $\mathrm{ch} : \mathcal{X} \to \Lambda_{\mathbb{C}}$ defined on homogenous $f \in \mathcal{X}_n$ by

$$\mathrm{ch}(f) = \langle f, \psi \rangle_{S_n} = \frac{1}{n!} \sum_{w \in S_n} f(w)\psi(w). \tag{11.5.10}$$

Lemma 11.5.11 With the notation as above

$$\mathrm{ch}(f) = \sum_{\lambda \vdash n} \frac{f_\mu}{z_\mu} p_\mu \tag{11.5.12}$$

where f_μ is the value of f on any element of the conjugacy class of cycle type μ, and z_μ is the size of the conjugacy class of cycle type μ.

Definition 11.5.13 — Hall Inner Product The **Hall inner product** on Λ is defined by

$$\langle p_\lambda, p_\mu \rangle = \delta_{\lambda\mu} z_\mu \tag{11.5.14}$$

with z_μ the size of the conjugacy class of cycle type μ.

Theorem 11.5.15. The ring $\mathcal{X}$ is isomorphic to Λ with the isomorphism given by $\mathrm{ch}(\chi_\lambda) = s_\lambda$. Further, this is an isometry with respect to the two inner products we've just defined. That is, $\langle \mathrm{ch}(f), \mathrm{ch}(g) \rangle = \langle f, g \rangle_{S_n}$ for homogeneous $f, g \in \mathcal{X}_n$.

Proof. To show that this is a ring homomorphism we have the following

$$\begin{aligned}
\mathrm{ch}(f * g) &= \langle \mathrm{Ind}_{S_m \times S_n}^{S_{m+n}} f \otimes g, \psi \rangle_{S_{m+n}} && (11.5.16)\\
&= \langle f \otimes g, \mathrm{Res}_{S_m \times S_n}^{S_{m+n}} \psi \rangle_{S_{m\times n}} && (11.5.17)\\
&= \langle f, \psi \rangle_{S_m} \langle g, \psi \rangle_{S_n} && (11.5.18)\\
&= \mathrm{ch}(f)\mathrm{ch}(g). && (11.5.19)
\end{aligned}$$

The Hall inner product is defined exactly such that this map is an isometry. Finally, since the χ_λ are a basis of $\mathcal{X}_n$ and the s_λ are a basis of Λ^n for $\lambda \vdash n$ then ch is an isomorphism. □

Theorem 11.5.20. Consider the following maps:

- $\mathrm{ch} : \mathcal{X} \to \Lambda_{\mathbb{C}}$;
- $\mathcal{X} \to Z = \bigoplus_{n\geq 0} Z(\mathbb{C}S_n)$ given by $\chi_\lambda \mapsto c_\lambda$;
- the Frobenius map $F : Z \to \Lambda_{\mathbb{C}}$ given by $F(c_\lambda) = p_\mu / z_\mu$.

These are isomorphisms, and the following diagram of these algebra isomorphisms commutes:

$$\begin{tikzcd}
\mathcal{X} \arrow[r] \arrow[dr] & Z \arrow[d] \\
& \Lambda_{\mathbb{C}}.
\end{tikzcd} \tag{11.5.21}$$

Corollary 11.5.22 If λ is a partition of n then

$$s_\lambda = \sum_{\mu \vdash n} \frac{\chi_\lambda(\mu)}{z_\mu} p_\mu. \tag{11.5.23}$$

Remark 11.5.24 The above diagram can be further extended via the boson–fermion correspondence to

$$\begin{tikzcd}
\mathcal{X} \arrow[r] \arrow[d] \arrow[dr] & Z \arrow[d] \\
\bigwedge^{\infty/2} V \arrow[r] & \Lambda_{\mathbb{C}}
\end{tikzcd} \tag{11.5.25}$$

where $V = \bigoplus_{n\in\mathbb{Z}} \mathbb{C}v_n$ and $\bigwedge^{\infty/2} V$ is defined to consist of all semi-infinite wedge products of the form

$$v_{i_1} \wedge v_{i_2} \wedge v_{i_3} \wedge \cdots . \tag{11.5.26}$$

The isomorphism $\bigwedge^{\infty/2} V \cong \Lambda_{\mathbb{C}}$ is called the boson-fermion correspondence. The details of this map are beyond the scope of this remark.

11.6 More Symmetric Functions

We have already seen three families of symmetric functions, p_λ, m_λ, and s_λ. Of these we've seen that s_λ are the images of irreducible characters under Macdonald's characteristic map. The m_λ are the images of the characters χ_{M_λ} under Macdonald's characteristic map. A corollary of this is that

$$m_\lambda = \sum_{\mu \vdash n} K_{\mu\lambda} s_\mu. \tag{11.6.1}$$

Note that this means that the expansion of m_λ in terms of Schur functions has only nonnegative coefficients, a property known as **Schur positivity**.

Characters of other representations likewise give us families of symmetric polynomials. In particular the **complete symmetric functions**,

$$h_n = \sum_{i_1 \leq \cdots \leq i_n} x_{i_1} \cdots x_{i_n}, \tag{11.6.2}$$

are the images of the trivial character, $\chi_{(n)}$, under this map. Note that this means that $h_n = s_{(n)}$. Similarly, the **elementary symmetric functions**,

$$e_n = \sum_{i_1 < \cdots < i_n} x_{i_1} \cdots x_{i_n}, \tag{11.6.3}$$

are the images of the sign character, $\chi_{(1,\ldots,1)}$, under the Macdonald characteristic map. Note that this means that $e_n = s_{(1,\ldots,1)}$.

We can use various relationships in the representation theory of the symmetric group to derive results about the symmetric polynomials. For example, the decomposition of M_λ as

$$V_\lambda \oplus \bigoplus_{\mu > \lambda} K_{\mu\lambda} V_\mu \tag{11.6.4}$$

factors through Macdonald's characteristic map to tell us that

$$h_\lambda = s_\lambda + \sum_{\mu > \lambda} K_{\mu\lambda} s_\mu. \tag{11.6.5}$$

Recalling that M_λ is defined by inducing the trivial representation of the row group, S_λ, we can also induce the sign representation of S_λ to get a similar decomposition which gives us

$$e_{\lambda'} = s_\lambda + \sum_{\mu < \lambda} K_{\mu'\lambda'} s_\mu. \tag{11.6.6}$$

The **Pieri rules** arise when we consider what happens if we induce the representation $\mathbb{C} \otimes V_\lambda$ or $\mathbb{C}_- \otimes V_\lambda$ (where $\mathbb{C}_-$ is the sign representation) of $S_m \times S_n$ up to S_{m+n}. The first gives

$$h_m s_\mu = \sum_{\substack{\lambda \vdash n \\ \lambda \backslash \mu \text{ horiz. strip}}} s_\lambda \tag{11.6.7}$$

and the second gives

$$e_m s_\mu = \sum_{\substack{\lambda \vdash n \\ \lambda\backslash\mu \text{ vert. strip}}} s_\lambda. \tag{11.6.8}$$

Note that $\lambda \setminus \mu$ is a **horizontal strip** if it has at most one box in each column, and a **vertical strip** if it has at most one box in each row.

11.7 Littlewood–Richardson Rule

Definition 11.7.1 — Littlewood–Richardson Coefficient Given a tableau we can form a word by concatenating the reversed rows from top to bottom. We say that the result of doing this is a **lattice word** if any prefix has at least as many 1s as it does 2s, at least as many 2s as it does 3s and so on. When the word of a tableau is a lattice word we call it a **Littlewood–Richardson tableau**.
The **Littlewood–Richardson coefficient**, $c^\nu_{\lambda\mu}$, is defined to be the number of of shape $\nu \setminus \lambda$ and weight μ.

Theorem 11.7.2 — Littlewood–Richardson Rule. Let λ and μ be partitions. Then

$$s_\lambda s_\mu = \sum_\nu c^\nu_{\lambda\mu} s_\nu. \tag{11.7.3}$$

The Littlewood–Richardson rule is a famously tricky result to prove, requiring some careful combinatorics. There are several related statements to the rule above.

One result which follows immediately is that if we have the simple $S_{|\lambda|}$ and $S_{|\mu|}$ modules V_λ and V_μ then we that, as $S_{|\nu|}$-modules, where $|\nu| = |\lambda| + |\mu|$, we have

$$\operatorname{Ind}^{S_{|\nu|}}_{S_{|\lambda|}\times S_{|\mu|}}(V_\lambda \otimes V_\mu) = \bigoplus_\nu c^\nu_{\lambda\mu} V_\nu \tag{11.7.4}$$

as S_{m+n}-modules. Conversely, we also have

$$\operatorname{Res}^{S_{|\nu|}}_{S_{|\lambda|}\times S_{|\mu|}} V_\nu = \bigoplus_{\lambda,\mu} c^\nu_{\lambda\mu} V_\lambda \otimes V_\mu \tag{11.7.5}$$

as $(S_m \times S_n)$-modules.

Another result, which is related to this one via Schur–Weyl duality, is that simple $\mathrm{SL}_n(\mathbb{C})$-modules can be labelled by partitions, call such a module E_λ, and then we have

$$E_\lambda \otimes E_\mu = \bigoplus_\nu c^\nu_{\lambda\mu} E_\nu. \tag{11.7.6}$$

In fact, it turns out that the Schur functions can be realised as the characters of these irreducible representations, and as such this is really a more general version of the Littlewood–Richardson rule as stated above, it's a sort of categorification of the rule.

11.8 Application: Intersection Cohomology of Grassmannians

Recall that the Grassmannian, $\mathrm{Gr}_k(\mathbb{C}^n)$, is defined to be the set of k-dimensional subspaces of $\mathbb{C}^n$. An element of $\mathrm{Gr}_k(\mathbb{C}^n)$ can be represented as a $k \times n$ matrix, specifically, it's the row space of this matrix. This doesn't give a unique representation of our subspace, but we can fix a unique representation by placing the matrix into reduced row echelon form, which doesn't change the row space. There are then k columns which are 0 apart from a single 1 (the pivot). The entries in the other $n-k$ columns determine exactly which k-dimensional subspace we're considering. These $k(n-k)$ entries can then be interpreted as coordinates, which makes $\mathrm{Gr}_k(\mathbb{C}^n)$ into a $k(n-k)$-dimensional complex manifold.

For example, consider $\mathrm{Gr}_4(\mathbb{C}^8)$, one particular subspace of this has pivots in columns 2, 3, 5, and 8, so it looks like

$$\begin{pmatrix} * & 1 & 0 & * & 0 & * & * & 0 \\ * & 0 & 1 & * & 0 & * & * & 0 \\ * & 0 & 0 & * & 1 & * & * & 0 \\ * & 0 & 0 & * & 0 & * & * & 1 \end{pmatrix}. \tag{11.8.1}$$

The $*$s represent values that we are free to vary. The above is the standard reduced row echelon form, but it will be more useful for us to use a slightly different convention, in which the left-most pivot appears lowest, so we would instead have

$$\begin{pmatrix} * & 0 & 0 & * & 0 & * & * & 1 \\ * & 0 & 0 & * & 0 & * & * & 0 \\ * & 0 & 1 & * & 0 & * & * & 0 \\ * & 1 & 0 & * & 0 & * & * & 0 \end{pmatrix}. \tag{11.8.2}$$

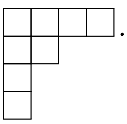
We can turn such a matrix into a Young diagram as follows. Take i_1 to be the left-most nonzero column, i_2 to be the left-most nonzero column linearly independent from i_1, and so on. Then we can use row operations to write the matrix so that there are pivots in columns i_j, each column before i_1 is just 0 (this is already the case) and each column between i_j and i_{j+1} starts with $k-j$ zeros. For example, with the (second) matrix form above, where the pivots appear in rows 2, 3, 5, and 8 we can basically set the first $k-j$ $*$s to be 0 up to column i_j:

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & * & * & 0 \\ 0 & 0 & 1 & * & 0 & * & * & 0 \\ 0 & 1 & 0 & * & 0 & * & * & 0 \end{pmatrix}. \tag{11.8.3}$$

Finally, delete the rows with the pivots, and we are left with

$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & * & * \\ 0 & * & * & * \\ 0 & * & * & * \end{pmatrix}. \tag{11.8.4}$$

Interpreting the 0s as boxes in our Young diagram gives



(11.8.5)

Conversely, given a partition, λ, we can consider the subset, $\Omega_\lambda^\circ \subseteq \mathrm{Gr}_k(\mathbb{C}^n)$, of all subspaces of $\mathbb{C}^n$ which have λ as their corresponding Young diagram. We call this a Schubert cell. We call the closure, Ω_λ, of one of these cells a Schubert variety.

It turns out that the cohomology of $\mathrm{Gr}_k(\mathbb{C}^n)$ is a freely generated abelian group on the classes $\sigma_\lambda = [\Omega_\lambda]$ as λ ranges over all Young diagrams with at most k rows and $n-k$ columns. This is a general fact about the cohomology of spaces admitting such a cellular decomposition.

Given a space, X, with a cohomology theory we can define the cohomology ring to be

$$H^\bullet(X) = \bigoplus_m H^m(X). \tag{11.8.6}$$

The product in this ring is given by the cup product, the details of which are beyond the scope of this course. However, in this case the product turns out to be given by

$$\sigma_\lambda \sigma_\mu = \sum_\nu c^\nu_{\lambda\mu} \sigma_\nu. \tag{11.8.7}$$

It turns out that the cohomology ring, $H^*(\mathrm{Gr}(k; \mathbb{C}^n))$, has the presentation

$$H^\bullet(\mathrm{Gr}(k; \mathbb{C}^n)) \cong \Lambda / I_{k,n} \tag{11.8.8}$$

where $I_{k,n}$ is the ideal generated by Schur functions s_λ where the Young diagram of λ has more than k rows or more than $n-k$ columns, so it doesn't fit in a $k \times (n-k)$ bounding box. The isomorphism is given by mapping a Schubert class to a Schur function, $\sigma_\lambda \mapsto s_\lambda$. Thus, the multiplication in the cohomology ring is nothing but the multiplication of Schur functions indexed by partitions fitting into a $k \times (n-k)$ bounding box. In this context the Littlewood–Richardson coefficients have an interpretation as the intersection numbers.

11.9 Hopf Algebra Structure

For more details on Hopf algebras see my notes from the Hopf algebras course (`https://github.com/WilloughbySeago/phd-courses-notes/tree/main/hopf-algebras`).

The ring of symmetric functions with coefficients in $\mathbb{C}$, $\Lambda = \Lambda_{\mathbb{C}}$, is a commutative and cocommutative Hopf algebra. First, note that we can identify $\Lambda \otimes \Lambda$ with Λ. The comultiplication is given by

$$\Delta(s_\lambda) = \sum_\mu s_{\lambda\backslash\mu} \otimes s_\mu \tag{11.9.1}$$

where $s_{\lambda\backslash\mu}$ is the skew Schur function defined by

$$s_{\lambda\backslash\mu} = \sum_\nu c^\lambda_{\mu\nu} s_\nu. \tag{11.9.2}$$

Thus,

$$\Delta(s_\lambda) = \sum_{\mu,\nu} c^\lambda_{\mu\nu} s_\nu \otimes s_\mu. \tag{11.9.3}$$

Note that this is cocommutative since $c^\lambda_{\mu\nu}$ is symmetric in μ and ν. In terms of power sums this comultiplication is given by

$$\Delta(p_r) = p_r \otimes 1 + 1 \otimes p_r. \tag{11.9.4}$$

The comultiplication on an arbitrary symmetric function, f, is

$$\Delta(f) = \sum_\mu s_\mu^\perp f \otimes s_\mu \tag{11.9.5}$$

where $s_\mu^\perp$ is the adjoint of s_μ with respect to the Hall inner product.

The counit is given by $\varepsilon(1) = 1$ and $\varepsilon(f) = 0$ for all homogeneous symmetric functions, f, of degree greater than zero. In other words, $\varepsilon(f)$ (for f not necessarily homogeneous) is simply the constant term of f.

The antipode is given by

$$\chi(s_\lambda) = (-1)^{|\lambda|} s_{\lambda'}. \tag{11.9.6}$$

The ring $\Lambda \otimes \Lambda$ inherits the inner product of Λ, namely

$$\langle f \otimes g, f' \otimes g' \rangle_{\Lambda\otimes\Lambda} = \langle f, f' \rangle_\Lambda \langle g, g' \rangle_\Lambda. \tag{11.9.7}$$

The ring $\mathcal{X}$, which we've shown to be isomorphic to Λ, inherits the Hopf algebra structure of Λ. Then, if we take class functions, $\varphi, \gamma, \eta \in \mathcal{X}$, such that $f = \mathrm{ch}(\varphi)$, $g = \mathrm{ch}(\gamma)$, and $h = \mathrm{ch}(\eta)$ Frobenius reciprocity tells us that

$$\langle \mathrm{Res}^{S_{m+n}}_{S_m \times S_n} \varphi, \gamma \otimes \eta \rangle_{S_m\times S_n} = \langle \varphi, \mathrm{Ind}^{S_{m+n}}_{S_m\times S_n} (\gamma \otimes \eta) \rangle_{S_{m+n}}. \tag{11.9.8}$$

Going back to Λ this tells us that

$$\langle \Delta(f), g \otimes h \rangle_{\Lambda\otimes\Lambda} = \langle f, gh \rangle_\Lambda. \tag{11.9.9}$$

From this fact it follows that the Hopf algebra structure of Λ is self-dual, and in particular that

$$\langle \Delta(s_\lambda), s_\mu \otimes s_\nu \rangle_{\Lambda\otimes\Lambda} = \langle s_\lambda, s_\mu s_\nu \rangle = c^\lambda_{\mu\nu}, \tag{11.9.10}$$

giving yet another interpretation of the Littlewood–Richardson coefficients.

11.10 Another Product

We induced a product structure on the class functions by defining a product on the symmetric functions. We can go the other way around, and use the product on homogeneous class functions, φ and ψ, to induce a product, $\cdot$, on the corresponding symmetric functions. Essentially, we require that ch is a homomorphism with respect to this product, so

$$\mathrm{ch}(\varphi) \cdot \mathrm{ch}(\psi) = \mathrm{ch}\varphi\psi \tag{11.10.1}$$

for $\varphi, \psi \in \mathcal{X}_n$. Then, taking all partitions to be partitions of n, we have

$$s_\lambda \cdot s_\mu = \sum_\nu \gamma^\nu_{\lambda\mu} s_\nu \tag{11.10.2}$$

where

$$\gamma^{\nu}_{\lambda\mu} = \langle \chi_\nu, \chi_\lambda \chi_\mu \rangle_{S_n}. \tag{11.10.3}$$

The power sums are unnormalised idempotents with respect to this product, that is

$$p_\lambda \cdot p_\mu = \delta_{\lambda\mu} z_\mu p_\mu \tag{11.10.4}$$

where z_μ is the size of the conjugacy class of cycle type μ.

This product is related to the Hall inner product, which turns out to be the result of evaluating this product at zero:

$$\langle f, g \rangle_\Lambda = (f \cdot g)(0, 0, \dots). \tag{11.10.5}$$

Twelve

Schur–Weyl Duality

12.1 Double Centraliser Theorem

Remark 12.1.1 The following result is commonly called the double centraliser theorem in representation theory. In functional analysis there is a version of this result replacing E with a (not-necessarily finite-dimensional) Hilbert space, H, and $\operatorname{End} E$ with the set space of bounded linear operators on H. Then the equivalent result (which holds for the closure of A) is often called the bicommutant theorem.

Definition 12.1.2 — Centraliser Let X be an algebra and A a subalgebra. Then the centraliser of A in X is

$$C_X(A) = \{x \in X \mid xa = ax \forall a \in A\}. \tag{12.1.3}$$

In the special case where $X = \operatorname{End} E$ for some finite-dimensional vector space, E, we have

$$C_{\operatorname{End} E}(A) = \{\varphi \in \operatorname{End} E \mid \varphi \circ f = f \circ \varphi \forall f \in A\}. \tag{12.1.4}$$

From this we see that this is exactly the condition for φ to be an intertwiner of f, viewed as a representation map of $\operatorname{End} E$. Thus, we have

$$C_{\operatorname{End} E}(A) = \operatorname{End}_A E. \tag{12.1.5}$$

Theorem 12.1.6 — Double Centraliser Theorem. Let E be a finite dimensional vector space, and let $A \subseteq \operatorname{End} E$ be a subalgebra. Let $B = \operatorname{End}_A E$. Then

- $A = \operatorname{End}_B E$;
- B is semisimple; and
- $E = \bigoplus_{i \in I} V_i \otimes W_i$ where V_i and W_i are simple modules of A and B respectively, in particular there is some common indexing set, I, for the corresponding simple modules.

R Note that we do not in general have a bijection between simple A-modules and simple B-modules. Instead, the common index set, I, may repeat some simple modules.

Proof. First, note that $\operatorname{End} E$ is a matrix algebra, since E is finite dimensional. Thus, $\operatorname{End} E$ is semisimple (Proposition 4.2.34). Then $A \subseteq \operatorname{End} E$ must be semisimple.
This tells us that, as A-modules, we have

$$E \cong \bigoplus_i V_i \otimes \operatorname{Hom}_A(V_i, E). \tag{12.1.7}$$

The right-hand-side inherits the action of A on V_i, that is $a.(v \otimes f) = av \otimes f$. Next, define the space $W_i = \operatorname{Hom}_A(V_i, E)$. Then we have

$$A \cong \bigoplus_i \operatorname{End} V_i \tag{12.1.8}$$

as algebras, and we have the chain of isomorphisms

$$\begin{aligned}
B &= \operatorname{End} AE && (12.1.9)\\
&= \operatorname{Hom}_A(V, V) && (12.1.10)\\
&\cong \operatorname{Hom}_A\Big(\bigoplus_i V_i \otimes W_i, E\Big) && (12.1.11)\\
&\cong \bigoplus_i \operatorname{Hom}_A(V_i \otimes W_i, E) && (12.1.12)\\
&\cong \bigoplus_i \operatorname{Hom}_A(W_i \otimes V_i, E) && (12.1.13)\\
&\cong \bigoplus_i \operatorname{Hom}(W_i, \operatorname{Hom}_A(V_i, E)) && (12.1.14)\\
&= \bigoplus_i \operatorname{Hom}(W_i, W_i) && (12.1.15)\\
&= \bigoplus_i \operatorname{End} W_i. && (12.1.16)
\end{aligned}$$

From this we know that *if* the W_i are simple B-modules then B is semisimple and we have *all* simple B-modules in this decomposition. We can check that the W_i are simple by checking that B acts transitively on the nonzero mps in $\operatorname{Hom}_A(V, E)$ where V is any simple A-module. Fix some nonzero $v \in V$. Since V is simple any map, $f \in \operatorname{Hom}_A(V, E)$, is determined by where it takes v as Av is a nonzero submodule of V and so by simplicity $Av = V$. Take $f, \tilde{f} \in \operatorname{Hom}_A(V, E)$ with $f(v) = e$ and $\tilde{f}(v) = \tilde{e}$. Since Ae is an invariant subspace of E we have the decomposition $E = Ae \oplus W$ for some submodule W. Define $T \colon E \to E$ by $T(ae) = ae'$ for $ae \in Ae$, and $T(w) = w$ for $w \in W$. This is a homomorphism of A-modules, and $T \circ f = \tilde{f}$. Thus, this defines a transitive action on the nonzero maps, and so the W_i really are simple B-modules.
We can now consider the original decomposition,

$$E \cong \bigoplus_i V_i \otimes \operatorname{Hom}_A(V_i, E) \tag{12.1.17}$$

as a decomposition of B-modules,

$$E \cong \bigoplus V_i \otimes W_i \tag{12.1.18}$$

where on the right $b \,.\, (v \otimes w) = v \otimes bw$.
Finally, since $V_i \cong \mathrm{Hom}_B(W_i, E)$ we get the same result if we start with $B \subseteq \mathrm{End}\, E$ and $A = \mathrm{End}_B\, E$. □

12.2 Schur–Weyl Duality for $\mathfrak{gl}_m$

Let $\Bbbk$ be an algebraically closed field of characteristic 0 (so basically $\mathbb{C}$). Let V be an m-dimensional $\Bbbk$-vector space. Take $E = V^{\otimes n}$, which is an mn-dimensional vector space.

Then $\mathrm{End}\, E$ naturally contains a copy of $\Bbbk S_n$, call this copy A. It acts by permuting factors in the tensor product. That is, if $\sigma \in S_n$ then

$$\sigma \,.\, (v_1 \otimes \cdots \otimes v_n) = v_{\sigma^{-1}(1)} \otimes \cdots \otimes v_{\sigma^{-1}(n)}. \tag{12.2.1}$$

Note that the inverse is used in the definition so that we get a left action. It is also possible to just use σ on the right, in which case we get a right action, but none of the following results are significantly effected by this choice.

We claim that $B = \mathrm{End}_A\, E$ is the image of[1] $U(\mathfrak{gl}_m)$ in $\mathrm{End}\, E$. The action of $x \in \mathfrak{gl}_m$ on $V^{\otimes n}$ is given by a generalisation of the Hopf algebra structure[2] of $U(\mathfrak{gl}_m)$, specifically, x acts as

$$\Delta(x) = x \otimes 1 \otimes \cdots \otimes 1 + 1 \otimes x \otimes 1 \otimes \cdots \otimes 1 + \cdots + 1 \otimes \cdots \otimes 1 \otimes x. \tag{12.2.2}$$

For example, if $n = 3$ then

$$x \,.\, (v_1 \otimes v_2 \otimes v_3) = xv_1 \otimes v_2 \otimes v_3 + v_1 \otimes xv_2 \otimes v_3 + v_1 \otimes v_2 \otimes xv_3 \tag{12.2.3}$$

where $x \in \mathfrak{gl}_m$ acts on $v_i \in V \cong \Bbbk^m$ in the obvious way.

[1] $U(\mathfrak{g})$ is the universal enveloping algebra of $\mathfrak{g}$, and $\mathfrak{gl}_m$ is nothing but the set of $m \times m$ matrices with coefficients in $\Bbbk$, so $U(\mathfrak{gl}_m)$ is exactly $\mathrm{Mat}_m(\Bbbk)$.

[2] See https://github.com/WilloughbySeago/phd-courses-notes/tree/main/hopf-algebras.

Proposition 12.2.4 — Shcur–Weyl Duality With notation as above $B = \mathrm{End}_A\, E$ is the image of $U(\mathfrak{gl}_m)$ in $\mathrm{End}\, E$ where $x \in \mathfrak{gl}_m$ acts by $\Delta(x)$.

Proof. First note that the actions of A and B on E commute. If we act first with A we permute the order of terms in the tensor product, then acting with B we sum in a symmetric way over all terms. Instead, acting first with B we get a symmetric sum of terms, and acting with A then permutes the tensor product in each term, but the result is just the same as we achieved first acting with A and then B.
This shows that the image of $U(\mathfrak{gl}_m)$ in $\mathrm{End}\, E$ is certainly a subalgebra of $B = \mathrm{End}_A\, E$, as commuting with A is exactly what is needed for an element of $\mathrm{End}\, E$ to be in $\mathrm{End}_A\, E$.
So, all that we need to do is show that B is contained in the image of $U(\mathfrak{gl}_m)$. This follows from the fact that we can identify $B = S^n(\mathrm{End}\, V)$, as this is by definition the subspace of $\mathrm{End}(V^{\otimes n})$ which is invariant under the action

of A. We can then apply the second part of Lemma 12.2.5, which tells us that B is generated by $\Delta(x)$ for $x \in U(\mathfrak{gl}_m)$, and thus we have containment in both directions. □

Lemma 12.2.5 Let $\Bbbk$ be a field of characteristic zero.

1. For any finite dimensional $\Bbbk$-vector space, U, the space $S^n U$ is spanned by elements of the form $u \otimes \cdots \otimes u$ for $u \in U$.
2. For any algebra, A, over $\Bbbk$, the algebra $S^n A$ is generated by $\Delta(a)$ for $a \in A$ with Δ as defined above.

Proof. 1. The space $S^n U$ is a simple $\mathrm{GL}(U)$-module, and the space spanned by $u \otimes \cdots \otimes u$ is nonzero, and is also a $\mathrm{GL}(U)$-module, so it must be all of $S^n U$.

2. Consider the symmetric polynomial $x_1 \cdots x_m$. We know that the ring of symmetric functions is generated by the power sums, p_r, meaning that there is a polynomial, P, such that

$$P(p_1(x), \dots, p_n(x)) = x_1 \cdots x_n. \tag{12.2.6}$$

We can take this polynomial, viewed as a formal expression, and evaluate it on elements of $S^n A$, replacing multiplication with the tensor product, and identifying x_r with $1 \otimes \cdots \otimes 1 \otimes a \otimes 1 \otimes \cdots \otimes 1$ where the a appears in the rth position. Then, for example with $n = 3$, we have

$$p_2(x) = x_1^2 + x_2^2 + x_3^2 \tag{12.2.7}$$

which we can identify with

$$\Delta(a^2) = a^2 \otimes 1 \otimes 1 + 1 \otimes a^2 \otimes 1 + 1 \otimes 1 \otimes a^2. \tag{12.2.8}$$

In general, we may identify $p_r(x)$ with $\Delta(a^r)$. Then with P as defined above we must have

$$P(\Delta(a), \Delta(a^2), \dots, \Delta(a^n)) = a \otimes \cdots \otimes a, \tag{12.2.9}$$

so we can generate the elements $a \otimes \cdots \otimes a$ for $a \in A$, and we know from the first part that these generate all of $S^n A$.

□

12.3 Schur–Weyl Duality for GL_m

Let V be a finite dimensional vector space, and let $E = V^{\otimes n}$. Then a copy of S_n is contained within $\operatorname{End} E$, with the copy of S_n acting by permuting factors in the tensor product. There is also a copy of GL_m contained in $\operatorname{End} E$, in which $g \in \mathrm{GL}_m$ acts by[3] $\Delta(g) = g \otimes \cdots \otimes g$, that is

[3] Note that this is once again the comultiplication of the Hopf algebra $\Bbbk\,\mathrm{GL}_m$.

$$g\,.\,(v_1 \otimes \cdots \otimes v_n) = gv_1 \otimes \cdots \otimes gv_n. \tag{12.3.1}$$

Theorem 12.3.2 — Schur–Weyl Duality. With notation as above the image of GL_m spans $\mathrm{End}_{S_n} E$.

Proof. First, note that the image of GL_m (denote this by B) is spanned by $g^{\otimes n}$ for $g \in \mathrm{End}\, V$. For $g \in \mathrm{GL}_m$ denote the span of $g^{\otimes n}$ by B'. Let $b \in \mathrm{End}\, V$ be arbitrary.
We claim that $b^{\otimes n} \in B'$. Note that for all but finitely many values of $t \in \mathbb{C}$ the matrix $b + tI$ is invertible (i.e., it's invertible when t is not an eigenvalue of b, of which there are only finitely many). Then $(b + tI)^{\otimes n}$ defines a one-parameter subset of B, and the fact that this element is invertible for all but finitely many terms means it's actually in B' by continuity (*I think*). In particular, for $t = 0$ we have that $b^{\otimes n} \in B'$. Thus, $B' = B$, and we are done. □

Corollary 12.3.3 With notation as above we can consider $E = V^{\otimes n}$ as a $(S_n \times \mathrm{GL}_m)$-module, and it decomposes as

$$\bigoplus_{\lambda} V_\lambda \otimes L_\lambda \tag{12.3.4}$$

where λ ranges over all partitions of n and

$$L_\lambda = \mathrm{Hom}_{S_n}(V_\lambda, E) \tag{12.3.5}$$

are distinct simple GL_m-modules or zero.

Example 12.3.6 For $\lambda = (n)$ we have

$$L_{(n)} = \mathrm{Hom}_{\Bbbk S_n}(V_{(n)}, V^{\otimes n}). \tag{12.3.7}$$

We know that as a vector space $V_{(n)}$ is one-dimensional, and $\sigma \in S_n$ acts trivially on $V_{(n)}$. Thus, maps $V_{(n)} \to V^{\otimes n}$ preserving this action are precisely maps $f : V_{(n)} \to V^{\otimes n}$ such that $\sigma\,.\,f(v) = f(\sigma\,.\,v) = f(v)$ for $v \in V_{(n)}$ and $\sigma \in S_n$. Such a map can be identified with the image of the single basis vector, $v \in V_{(n)}$, providing a bijection $\mathrm{Hom}_{\Bbbk S_n}(V_{(n)}, V^{\otimes n}) \to V^{\otimes n}$ by $f \mapsto f(v)$. Since $f(v)$ is invariant under the action of S_n we know that $f(v) \in S^n V \subseteq V^{\otimes n}$. Thus, we can identify $\mathrm{Hom}_{\Bbbk S_n}(V_{(n)}, V^{\otimes n})$ with an S_n-submodule of $S^n V$, and since $S^n V$ is a simple S_n-module it must be that $\mathrm{Hom}_{\Bbbk S_n}(V_{(n)}, V^{\otimes n}) \cong S^n V$.
We can do something similar for $\lambda = (1^n)$, we have

$$L_{(1^n)} = \mathrm{Hom}_{\Bbbk S_n}(V_{(1^n)}, V^{\otimes n}). \tag{12.3.8}$$

We know that as a vector space $V_{(1^n)}$ is one-dimensional, and $\sigma \in S_n$ acts by a sign. Thus, maps $V_{(1^n)} \to V^{\otimes n}$ preserving this action are precisely maps

$f : V_{(1^n)} \to V^{\otimes n}$ such that $\sigma \,.\, f(v) = f(\sigma \,.\, v) f((\operatorname{sgn}\sigma)v) = (\operatorname{sgn}\sigma) f(v)$ for $v \in V_{(1^n)}$ and $\sigma \in S_n$. Again, we can identify such a map with $f(v)$ for some fixed $v \in V_{(1^n)}$. Since S_n acts on $f(v)$ by a sign we know that $f(v) \in \Lambda^n V$, and since $\Lambda^n V$ is a simple S_n-module (for $n \le \dim V$) we must have $\operatorname{Hom}_{\Bbbk S_n}(V_{(1^n)}, V^{\otimes n}) \cong \Lambda^n V$.

12.3.1 Finite Dimensional $\mathrm{GL}_m(\Bbbk)$-Modules

Let V be a finite-dimensional $\mathrm{GL}_m(\Bbbk)$-module. Then we have a representation map

$$\rho : \mathrm{GL}_m(\Bbbk) \to \mathrm{GL}(V). \tag{12.3.9}$$

Since we're working with finite-dimensional spaces we can pick a basis and identify $\mathrm{GL}(V) = \mathrm{GL}_n(\Bbbk)$. Then this is a map taking in an invertible $m \times m$ matrix, g, and outputting an invertible $n \times n$ matrix, $\rho(g)$.

Definition 12.3.10 — Regular and Polynomial Representations Let V be a finite dimensional $\mathrm{GL}_m(\Bbbk)$-module with representation map ρ. We call this representation **regular** if the matrix elements, $\rho(g)_{kl}$, are polynomial in g_{ij} and $(\det g)^{-1}$. If there is no dependence on $(\det g)^{-1}$ then we call the representation **polynomial**.

For non-finite fields $\mathrm{GL}_m(\mathbb{C})$ is not finite, so there's no guarantee that any of our results about representations of finite groups hold. However, in many cases the regular or polynomial representations turn out to be nice enough that many of our results still hold. For example, it's possible to classify these subclasses of representations.

Now consider $\Bbbk = \mathbb{C}$. The Lie algebra $\mathfrak{gl}_m(\mathbb{C})$ acts on V by

$$x \,.\, v = \frac{\mathrm{d}}{\mathrm{d}t} \mathrm{e}^{tx} \,.\, v \bigg|_{t=0}. \tag{12.3.11}$$

The action on the right is that of $\mathrm{GL}_m(\mathbb{C})$ on V, which is to say $\mathrm{e}^{tx} \,.\, v = \rho(\mathrm{e}^{tx})v$.

It is a fact that $\mathrm{GL}_m(\mathbb{C})$ contains a compact subgroup, U_m, consisting of only the unitary matrices. This is a *real* Lie group with real Lie algebra, $\mathfrak{u}_m$, consisting of skew-Hermitian matrices. Then we can recover all of $\mathfrak{gl}_m(\mathbb{C})$ as the complexification

$$\mathfrak{gl}_m(\mathbb{C}) = \mathfrak{u}_m \oplus i\mathfrak{u}_m. \tag{12.3.12}$$

This is simply saying that every complex matrix can be written as a sum of a skew-Hermitian matrix and a Hermitian matrix, which can be seen immediately by realising that $A + A^*$ is Hermitian and $A - A^*$ is skew-Hermitian, and then $A = (A + A^*)/2 + (A - A^*)/2$.

Proposition 12.3.13 — Weyl's Unitarity Trick Let V be a GL_m-module. Then V is a simple GL_m-module if and only if it is a simple U_m-module.

Proof. The details of this are beyond the scope of this course, needing the notion of a Haar measure. The idea is the same as that of Theorem 6.5.33, we can make any representation of GL_m unitary by defining a new inner product on V by

$$(v, w) = \int_{\mathrm{U}_m} \langle gv, gw \rangle \, \mathrm{d}\mu(g) \tag{12.3.14}$$

where μ is the Haar measure. □

Theorem 12.3.15 — Polynomial Representations. The irreducible polynomial representations of $\mathrm{GL}_m(\mathbb{C})$ are precisely the L_λ for which λ is a partition (of an arbitrary nonnegative integer) of length at most m. Further, the character of $g \in \mathrm{GL}_m(\mathbb{C})$ in this representation is precisely the Schur polynomial $s_\lambda(x_1, \dots, x_m)$ evaluated at the m eigenvalues $x_1, \dots, x_m$ of g.

If V is an m-dimensional vector space then we can consider the polynomial representation L_λ as a submodule of $V^{\otimes n}$. Specifically, it's the image of V under the Schur functor $S_\lambda : \mathrm{GL}_m\text{-Mod} \to \mathrm{GL}_m\text{-Mod}$, defined by setting $S_\lambda(V)$ to be the result of acting on $V^{\otimes n}$ with the corresponding Young projector of λ. For example, $S_{(n)}V = S^n V$ and $S_{(1^n)}(V) = \Lambda^n V$. For $n = 3$ $S_{(2,1)}V$ is the subspace of $V^{\otimes n}$ which is symmetric under exchange of the first two factors, and antisymmetric under exchange of any factor with the third factor.

Non-polynomial representations can also be indexed by decreasing sequences of integers, known as weights, but there is no positivity requirement, so they aren't (necessarily) partitions.

Let $g \in \mathrm{GL}_m(\mathbb{C})$ have eigenvalues $x_1, \dots, x_n \in \mathbb{C}$. Consider the setup of Schur–Weyl duality, that is $V = \mathbb{C}^m$, $E = V^{\otimes n}$, considered as an $(S_n \times \mathrm{GL}_m(\mathbb{C}))$-module. Then for $w \in S_n$ of cycle type μ we can take the trace in this representation. On the one hand, we have

$$\operatorname{tr}_E((w, g^{\otimes n})) = p_\mu(x_1, \dots, x_m), \tag{12.3.16}$$

and on the other we have

$$\operatorname{tr}_E((w, g^{\otimes n})) = \operatorname{tr}_{\bigoplus_\lambda V_\lambda \otimes L_\lambda}(w g^{\otimes n}) = \sum_\lambda \chi_\lambda(\mu) s_\lambda(x_1, \dots, x_m). \tag{12.3.17}$$

Thus, we have

$$p_\mu(x_1, \dots, x_m) = \sum_\lambda \chi_\lambda(\mu) s_\lambda(x_1, \dots, x_m) \tag{12.3.18}$$

where λ runs over all partitions of n, μ is some fixed partition of n, and χ_λ is the character of the corresponding irreducible S_n-module.

Theorem 12.3.19 — Peter–Weyl Theorem. Let R be the algebra of polynomial functions on $\mathrm{GL}(V)$. Then this is a $(\mathrm{GL}(V) \times \mathrm{GL}(V))$-module, with the action $((g,h).\varphi)(x) = \varphi(g^{-1}xh)$ for $g,h,x \in \mathrm{GL}(V)$ and $\varphi \in R$. Then R decomposes as

$$R = \bigoplus_{\lambda} L_\lambda^* \otimes L_\lambda \tag{12.3.20}$$

where λ runs over all partitions.

12.4 Howe Duality

Schur–Weyl duality is concerned with $(S_n \times \mathrm{GL}_m)$-modules. Howe duality, on the other hand, is concerned with $(\mathrm{GL}_m \times \mathrm{GL}_n)$-modules.

We'll work over the complex numbers in this section. There is a natural action of $\mathrm{GL}_m \times \mathrm{GL}_n$ on $\mathbb{C}^m \otimes \mathbb{C}^n$, namely $(g,g').v \otimes v' = gv \otimes g'v'$. By the same arguments as applied to Schur–Weyl duality this action commutes with the action of $S_k \times S_\ell$ on $(\mathbb{C}^m)^{\otimes k} \otimes (\mathbb{C}^n)^{\otimes \ell}$. Thus, we can consider $S(\mathbb{C}^m \otimes \mathbb{C}^n)$ and $\Lambda(\mathbb{C}^m \otimes \mathbb{C}^n)$. Howe duality is then a statement as to how these decompose as $(\mathrm{GL}_m \times \mathrm{GL}_n)$-modules.

Theorem 12.4.1 — Howe Duality. We have

- $S(\mathbb{C}^m \otimes \mathbb{C}^n) \cong \bigoplus_{\lambda : \ell(\lambda) \le \min\{m,n\}} L_\lambda^m \otimes L_\lambda^n$;
- $\Lambda(\mathbb{C}^m \otimes \mathbb{C}^n) \cong \bigoplus_{\lambda \subseteq \square_{n}^{m}} L_\lambda^m \otimes L_{\lambda'}^n$, where $\lambda \subseteq \square_{n}^{m}$ means that the Young diagram of λ fits in an $m \times n$ bounding box, that is, λ has at most m rows and n columns.

Note that $S(\mathbb{C}^m \otimes \mathbb{C}^n)$ is infinite dimensional (for $m, n \neq 0$). This means that the character of this representation is not well defined. We fix this with the graded character, which encodes the character of each homogeneous component. Specifically, we have

$$S(\mathbb{C}^m \otimes \mathbb{C}^n) = \bigoplus_{k \ge 0} S^k(\mathbb{C}^m \otimes \mathbb{C}^n), \tag{12.4.2}$$

and we can define the graded character to be the formal power series

$$\chi_S = \sum_{k \ge 0} z^k \chi_{S^k}. \tag{12.4.3}$$

Here χ_{S^k} is the character in the representation $S^k(\mathbb{C}^m \otimes \mathbb{C}^n)$, which is well defined as the trace of an operator on a finite-dimensional space. Note that when we evaluate χ_S on $(g,g') \in \mathrm{GL}_m \times \mathrm{GL}_n$ we get a power series in z, and so χ_S is a power-series valued linear function, $\chi_S \in \mathrm{Hom}(\mathrm{GL}_m \times \mathrm{GL}_n, \mathbb{C}[\![z]\!])$. Note that taking the graded trace of (I_m, I_n) gives us the graded dimension,

$$\sum_{k \ge 0} z^k \dim(S^k(\mathbb{C}^m \otimes \mathbb{C}^n)). \tag{12.4.4}$$

This definition of the graded trace and dimension can be extended to any graded representation.

Proposition 12.4.5 — Cauchy Identities The following hold

- $\displaystyle\prod_{i=1}^{m}\prod_{j=1}^{n}\frac{1}{1-zx_iy_j} = \sum_{\lambda:\ell(\lambda)\leq\min\{m,n\}} z^{|\lambda|}s_\lambda(x)s_\lambda(y);$
- $\displaystyle\prod_{i=1}^{m}\prod_{j=1}^{n}(1+zx_iy_j) = \sum_{\lambda\subseteq \square^{m}_{n}} z^{|\lambda|}s_\lambda(x)s_{\lambda'}(y).$

The Cauchy identities can be proven from Howe duality by taking characters, or they can be proven purely from the theory of symmetric functions and a result known as the RSK correspondence. We can then use the Cauchy identities to prove Howe duality.

Proposition 12.4.6 — Pieri Rules Let μ be a partition of n. Then as GL_m-modules we have the decompositions

- $L_\mu \otimes S^r(\mathbb{C}^m) \cong \bigoplus_\lambda L_\lambda$ where the sum is over partitions, λ, of $n+r$ such that i) $\lambda\backslash\mu$ is a horizontal strip (at most one box in each column) and ii) λ has at most m rows;
- $L_\mu \otimes \Lambda^r(\mathbb{C}^m) \cong \bigoplus_\lambda L_\lambda$ where the sum is over partitions, λ, of $n+r$ such that i) $\lambda \setminus \mu$ is a vertical strip (at most one box in each row) ii) λ has at most m rows.

Notice that $g \in \mathrm{GL}_m$ acts on the rth tensor power by acting on each term with g. This means that g acts like g^r. If g is diagonal with eigenvalues $\{x_1, \dots, x_m\}$ then g^r is diagonal with eigenvalues x_i^r, and thus taking the trace of this action we find that the character is $x_1^r + \dots + x_m^r = p_r(x)$. For $S^r(\mathbb{C}^m)$ we have a similar result, except that we are symmetrising everything, which means that we get all possible degree r monomials in the x_i, not just x_i^r, we also get, for example, $x_1x_2^{r-1}$ and $x_1x_2^3x_7^{r-4}$. Thus, the character of the representation $S^r(\mathbb{C}^m)$ is $h_r(x)$. For $\Lambda^r(\mathbb{C}^m)$ we are antisymmetrising everything, and this means we get all possible degree r monomials in the x_i with the additional restriction that no element can be repeated. For example, if $r=3$ then we get $x_1x_2x_3$ and $x_1x_2x_7$, but not $x_1^2x_2$. Thus, the character of the representation $\Lambda^r(\mathbb{C}^m)$ is $e_r(x)$.

Taking characters of the results in the previous proposition thus gives us the following corollary.

Corollary 12.4.7 With the same notation as above

- $s_\mu h_r = \sum_\lambda s_\lambda$ where the sum is over partitions, λ, of $n+r$ such that i) $\lambda \setminus \mu$ is a horizontal strip (at most one box in each column) ii) λ has at most m rows
- $s_\mu e_r = \sum_\lambda s_\lambda$ where the sum is over partitions, λ, of $n+r$ such that i) $\lambda \setminus \mu$ is a vertical strip (at most one box in each row) ii) λ has at most m rows.

We can check this for a small example, taking $m = 2$, $n = 3$, $r = 2$, and $\mu = (2, 1)$. We then have

$$s_{\ydiagram{2,1}} h_2 = (x_1^2 x_2 + x_1 x_2^2)(x_1^2 + x_1 x_2 + x_2^2) \tag{12.4.8}$$

$$= x_1^4 x_2 + 2x_1^3 x_2^2 + 2x_1^2 x_2^3 + x_1 x_2^4. \tag{12.4.9}$$

The 5 box Young diagrams with at most 2 rows are

$$\ydiagram{5}\,, \quad \ydiagram[*(white)]{4,1}*[*(orange)]{2+2}\,, \quad \text{and} \quad \ydiagram[*(white)]{3,2}*[*(orange)]{2+1,1+1}\,. \tag{12.4.10}$$

Of these, the first cannot be achieved by adding boxes to $\mu = (2, 1)$, the other two can, with the added boxes highlighted above. Note that no column contains more than one highlighted box, and therefore both are given by adding a horizontal strip to $\mu = (2, 1)$, so $\lambda \setminus \mu$ is always a horizontal strip. Thus, if the result above holds we should have

$$s_{\ydiagram{2,1}} h_2 = s_{\ydiagram{4,1}} + s_{\ydiagram{3,2}}, \tag{12.4.11}$$

and indeed this is the case, as one can check:

$$s_{\ydiagram{4,1}} + s_{\ydiagram{3,2}} = (x_1^4 x_2 + x_1^3 x_2^2 + x_1^2 x_2^3 + x_1 x_2^4) + (x_1^3 x_2^2 + x_1^2 x_2^3) \tag{12.4.12}$$

which gives the same result as above.

If instead $m \geq 5$ then we have to consider all 5 box Young diagrams which are generated by adding a two box horizontal strip to $\mu = (2, 1)$. These are

$$\ydiagram[*(white)]{2,2,1}*[*(orange)]{0,1+1,1} \quad \ydiagram[*(white)]{3,2}*[*(orange)]{2+1,1+1} \quad \ydiagram[*(white)]{3,1,1}*[*(orange)]{2+1,0,1} \quad \text{and} \quad \ydiagram[*(white)]{4,1}*[*(orange)]{2+2}\,. \tag{12.4.13}$$

One can then check that

$$h_2 s_{\ydiagram{2,1}} = s_{\ydiagram{2,2,1}} + s_{\ydiagram{3,2}} + s_{\ydiagram{3,1,1}} + s_{\ydiagram{4,1}}. \tag{12.4.14}$$

For example, the following code does this in Mathematica.

Code 12.4.15

```
SchurS = ResourceFunction["SchurS"];
Module[{h, s, vars}
vars = {x1, x2, x3, x4, x5};
h = SchurS[{2}, vars];
s[λ_] = SchurS[λ, vars];
h s[{2,1}] == s[{2,2,1}] + s[{3,2}]
+ s[{3,1,1}] + s[{4,1}] // Simplify
]
```

Note that these results are special cases of the Littlewood–Richardson rule. In particular, we've taken $h_r = s_{(r)}$ and $e_r = s_{(1^r)}$.

12.5 $\mathrm{GL}_m(\mathbb{C})$ Branching Rules

Let λ and μ be partitions. We say that μ **interleaves** λ if

$$\lambda_1 \geq \mu_1 \geq \lambda_2 \geq \mu_2 \geq \cdots \geq \mu_{m-1} \geq \lambda_m. \tag{12.5.1}$$

Consider the inclusion

$$\begin{aligned} \mathrm{GL}_{m-1} &\hookrightarrow \mathrm{GL}_m \\ g &\mapsto \begin{pmatrix} g & 0 \\ 0 & 1 \end{pmatrix}. \end{aligned} \tag{12.5.2}$$

Proposition 12.5.3 With the inclusion above we have

$$\mathrm{Res}^{\mathrm{GL}_m}_{\mathrm{GL}_{m-1}} L^{\mathrm{GL}_m}_\lambda = \bigoplus_\mu L^{\mathrm{GL}_{m-1}}_\mu \tag{12.5.4}$$

where μ runs over all partitions which interleave λ, and we write L^G_μ for the irreducible G-modules of $G = \mathrm{GL}_m, \mathrm{GL}_{m-1}$.

Note in particular that the decomposition above is multiplicity free.

We can chain together inclusions like the above:

$$\mathbb{C}^\times \cong \mathrm{GL}_1 \hookrightarrow \mathrm{GL}_2 \hookrightarrow \ldots \hookrightarrow \mathrm{GL}_{m-n} \hookrightarrow GL_m. \tag{12.5.5}$$

Corollary 12.5.6 With the chain of inclusions above we have

$$\mathrm{Res}^{\mathrm{GL}_m}_{\mathrm{GL}_1} L^{\mathrm{GL}_m}_\lambda = \bigoplus_\Lambda \mathbb{C} v_\Lambda \tag{12.5.7}$$

where Λ runs over all Gelfand–Zetlin patterns (defined after this result) starting with λ.

Let λ be a partition with m parts. A **Gelfand–Zetlin pattern** is an upside-down triangle of rows of numbers, Λ_{ij}, where i is the row, and j the position in the row. The Gelfand–Zetlin pattern corresponding to λ starts with the row

$$\Lambda_{m1} \qquad \Lambda_{m2} \qquad \Lambda_{m3} \qquad \ldots \qquad \Lambda_{mm} \tag{12.5.8}$$

where $\Lambda_{mk} = \lambda_k$. For a valid Gelfand–Zetlin pattern the row below this must satisfy $\Lambda_{m\ell} \geq \Lambda_{m-1,\ell} \geq \Lambda_{m-1,\ell+1}$. The second row has $m-2$ entries, and interpreted as a partition, μ with $\mu_k = \Lambda_{m-1,k}$ this construction is such that μ interleaves λ. Thus, we have two rows

$$\begin{array}{ccccccccc} \Lambda_{m1} & & \Lambda_{m2} & & \Lambda_{m3} & & \cdots & & \Lambda_{mm} \\ & \Lambda_{m-1,1} & & \Lambda_{m-1,2} & & \Lambda_{m-1,3} & \cdots & \Lambda_{m-1,m-1} & \end{array} \tag{12.5.9}$$

The next row is defined similarly, and so on, we always have $\Lambda_{m-k,\ell} \geq \Lambda_{m-k-1,\ell} \geq \Lambda_{m-k,\ell+1}$, and the kth row has k entries. So, a full Gelfand–Zetlin pattern looks like

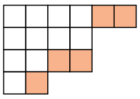
$$
\begin{array}{ccccccccc}
\Lambda_{m1} & & \Lambda_{m2} & & \Lambda_{m3} & & \cdots & & \Lambda_{mm} \\
& \Lambda_{m-1,1} & & \Lambda_{m-1,2} & & \Lambda_{m-1,3} & \cdots & \Lambda_{m-1,m-1} & \\
& \ddots & & \ddots & & \iddots & & \iddots & \\
& & & \Lambda_{21} & & \Lambda_{22} & & & \\
& & & & \Lambda_{11} & & & &
\end{array}
\tag{12.5.10}
$$

where each entry is bounded between the two entries above it to either side.

Since we have a decomposition into one-dimensional spaces, $\mathbb{C}v_\Lambda$, indexed by Gelfand–Zetlin patterns we see that the Gelfand–Zetlin patterns provide a basis for

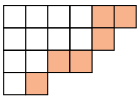
$$L_\lambda \cong \mathrm{Hom}_{S_n}(V_\lambda, (\mathbb{C}^m)^{\otimes n}). \tag{12.5.11}$$

Let λ and μ be partitions with μ interleaving λ. In terms of Young diagrams this means that $\lambda \setminus \mu$ must be a horizontal strip. We can see this from the following example:



(12.5.12)

where the highlighted boxes are $\lambda \setminus \mu$ (so the white boxes are μ and λ is the whole diagram). If instead we had



(12.5.13)

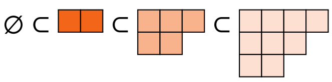
then the extra box means that $\lambda_2 = 5 > \mu_1 = 4$, which isn't allowed if μ interleaves λ.

From this we can see that the Gelfand–Zetlin patterns are in bijection with the semistandard Young tableaux of shape λ, since we can consider such a tableau to be built up in horizontal strips in the order of labelling the boxes. Recall also that the number of semistandard Young tableau of shape λ with weight μ is given by the Kostka numbers, $K_{\lambda\mu}$.

Start with the semistandard Young tableau of shape $\lambda = (4, 3, 2)$ given by

$$T = \begin{array}{|c|c|c|c|} \hline 1 & 1 & 2 & 3 \\ \hline 2 & 2 & 3 \\ \cline{1-3} 3 & 3 \\ \cline{1-2} \end{array}\,. \tag{12.5.14}$$

We then have the inclusions



(12.5.15)

The corresponding Gelfand–Zetlin pattern is given by taking each row to be one of these partitions:

$$
\begin{array}{ccccc}
4 & & 3 & & 2 \\
& 3 & & 2 & \\
& & 2 & &
\end{array}
\tag{12.5.16}
$$

This gives us a bijection between semistandard Young tableaux of shape λ and Gelfand–Zetlin patterns.

Taking the character of the module

$$\operatorname{Res}^{\mathrm{GL}_m}_{\mathrm{GL}_1} L_\lambda^{\mathrm{GL}_m} = \bigoplus_{\Lambda} \mathbb{C} v_\Lambda \tag{12.5.17}$$

we get

$$s_\lambda(x_1, \dots, x_m) = \sum_T x^T \tag{12.5.18}$$

where T is a semistandard tableau of shape λ and

$$x^T := x_1^{|T^{-1}(1)|} \cdots x_m^{|T^{-1}(m)|} \tag{12.5.19}$$

where $|T^{-1}(i)|$ is the number of boxes filled with an i.

This result shows that the s_λ really are polynomials, our initial definition only has them as rational functions. It also shows that the coefficients of the L_λ characters are manifestly positive.

In order for the Gelfand–Zetlin basis to be useful we need to understand how GL_m acts on it. It turns out to actually be easier to consider how $\mathfrak{gl}_m$ acts on this basis. Let E_{ij} be the elementary $m \times m$ matrix with a 1 in position (i, j) and zero everywhere else. These matrices form a basis of $\mathfrak{gl}_m$, and $[E_{ij}, E_{k\ell}] = \delta_{jk} E_{i\ell} - \delta_{i\ell} E_{kj}$ is the Lie bracket in this basis.

The matrices E_{kk} form the standard Cartan subalgebra of diagonal matrices. The entirety of $\mathfrak{gl}_m$ is generated by E_{kk}, $E_{k,k+1}$ and $E_{k+1,k}$.

The basis $\{v_\Lambda\}$ for L_λ is then such that

$$E_{kk} v_\Lambda = \left(\sum_{i=1}^{k} \Lambda_{ki} - \sum_{i=1}^{k-1} \Lambda_{k-1,i} \right) v_\Lambda \tag{12.5.20}$$

$$E_{k,k+1} v_\Lambda = -\sum_{i=1}^{k} \frac{(l_{ki} - l_{k+1,1}) \cdots (l_{ki} - l_{k+1,k+1})}{(l_{ki} - l_{k1}) \cdots (\widehat{l_{ki} - l_{ki}}) \cdots (l_{ki} - l_{kk})} v_{\Lambda + \delta_{ki}} \tag{12.5.21}$$

$$E_{k+1,k} v_\Lambda = \sum_{i=1}^{k} \frac{(l_{ki} - l_{k-1,1}) \cdots (l_{ki} - l_{k-1,k-1})}{(l_{ki} - l_{k1}) \cdots (\widehat{l_{ki} - l_{ki}}) \cdots (l_{ki} - l_{kk})} v_{\Lambda - \delta_{ki}} \tag{12.5.22}$$

where $l_{ki} = \Lambda_{ki} - i + 1$ and $\hat{x}$ denotes that x is omitted from the product and $\Lambda \pm \delta_{ki}$ is given by replacing Λ_{ki} with $\Lambda_{ki} \pm 1$. If the result is not a Gelfand–Zetlin pattern then we set $v_{\Lambda \pm \delta_{ki}} = 0$.

For example, for $\mathfrak{gl}_2$ take $\lambda = (2, 1)$. The semistandard Young tableaux of shape λ are then

$$\begin{array}{|c|c|}\hline 1 & 1 \\ \hline 2 & \\ \cline{1-1} \end{array} \quad \text{and} \quad \begin{array}{|c|c|}\hline 1 & 2 \\ \hline 2 & \\ \cline{1-1} \end{array}. \tag{12.5.23}$$

The corresponding inclusions chains are

$$\varnothing \subset \begin{array}{|c|c|}\hline & \\ \hline \end{array} \subset \begin{array}{|c|c|}\hline & \\ \hline & \\ \cline{1-1} \end{array}, \quad \text{and} \quad \varnothing \subset \begin{array}{|c|}\hline \\ \hline \\ \hline \end{array} \subset \begin{array}{|c|c|}\hline & \\ \hline & \\ \cline{1-1} \end{array}. \tag{12.5.24}$$

The corresponding Gelfand–Zetlin patterns are

$$\begin{array}{ccc} 2 & & 1 \\ & 2 & \end{array}, \quad \text{and} \quad \begin{array}{ccc} 2 & & 1 \\ & 1 & \end{array}. \tag{12.5.25}$$

Let these be patterns Λ^1 and Λ^2 respectively, and call the corresponding basis vectors v_1 and v_2. Then we have the action of the Cartan subalgebra given by

$$E_{11}v_1 = \left(\sum_{i=1}^{1} \Lambda^1_{1i} - \sum_{i=1}^{1-1} \Lambda^1_{1-1,i}\right) v_1 = \Lambda^1_{11} v_1 = 2v_1, \tag{12.5.26}$$

$$E_{11}v_2 = \left(\sum_{i=1}^{1} \Lambda^2_{1i} - \sum_{i=1}^{1-1} \Lambda^2_{1-1,i}\right) v_2 = \Lambda^2_{21} v_2 = v_2, \tag{12.5.27}$$

$$E_{22}v_1 = \left(\sum_{i=1}^{2} \Lambda^1_{2i} - \sum_{i=1}^{2-1} \Lambda^1_{2-1,i}\right) v_1 = (\Lambda^1_{21} + \Lambda^1_{22} - \Lambda^1_{11}) v_1 \tag{12.5.28}$$

$$= (2 + 1 - 2)v_1 = v_1, \tag{12.5.29}$$

$$E_{22}v_2 = \left(\sum_{i=1}^{2} \Lambda^2_{2i} - \sum_{i=1}^{2-1} \Lambda^2_{2-1,i}\right) v_2 = (\Lambda^2_{21} + \Lambda^2_{22} - \Lambda^1_{11}) v_2 \tag{12.5.30}$$

$$= (2 + 1 - 1)v_2 = 2v_2. \tag{12.5.31}$$

The action of the off diagonal matrices can also be computed with a bit of work, for example, we have

$$E_{12}v_1 = -\sum_{i=1}^{1} \frac{(l_{1i} - l_{1+1,1}) \cdots (l_{1i} - l_{1+1,1+1})}{(l_{1i} - l_{11}) \cdots (\widehat{l_{1i} - l_{1i}}) \cdots (l_{1i} - l_{11})} v_{\Lambda^1 - \delta_{1i}} \tag{12.5.32}$$

$$= -(l_{11} - l_{21})(l_{11} - l_{22})v_1 \tag{12.5.33}$$

$$= -(\Lambda^1_{11} - 1 + 1 - \Lambda^1_{21} + 1 - 1)(\Lambda^1_{11} - 1 + 1 - \Lambda^1_{22} + 2 - 1)v_1 \tag{12.5.34}$$

$$= -(2 - 1 + 1 - 2 + 1 - 1)(2 - 1 + 1 - 1 - 2 + 1)v_1 \tag{12.5.35}$$

$$= 0. \tag{12.5.36}$$

Part IV

Other Topics in Representation Theory

Thirteen

Lie Algebras

In this section we give a rapid, relatively proof free, tour of the representation theory of Lie algebras. We refer the reader to other sources for details, such as my lecture notes `https://github.com/WilloughbySeago/phd-courses-notes/tree/main/lie-theory`.

13.1 Lie Algebras

Definition 13.1.1 — Lie Algebra A **Lie algebra**, $\mathfrak{g}$, is a $\Bbbk$-vector space equipped with a linear map $\mathfrak{g} \otimes \mathfrak{g} \to \mathfrak{g}$ called the **Lie bracket** subject to the following:

- **alternativity**: $[x, x] = 0$ for all $x \in \mathfrak{g}$;
- **Jacobi identity**: $[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0$ for all $x, y, z \in \mathfrak{g}$.

Note that more commonly the definition is given as a bilinear map $\mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$. The universal property of the tensor product means that these are equivalent. For fields of characteristic other than 2 the first relation is usually replaced with antisymmetry, $[x, y] = -[y, x]$ for all $x, y \in \mathfrak{g}$. With our definition using the tensor product we can pass to the quotient $\Lambda^2\mathfrak{g}$ and we see that $[-, -]$ induces a map $[-, -] : \Lambda^2\mathfrak{g} \to \mathfrak{g}$ which trivially is such that $[x, x] = 0$ since $x \otimes x$ maps to zero in $\Lambda^2\mathfrak{g}$.

Definition 13.1.2 Let $\mathfrak{g}$ and $\mathfrak{g}'$ be Lie algebras over the same field, $\Bbbk$. A morphism of Lie algebras, $\varphi : \mathfrak{g} \to \mathfrak{g}'$ is a linear map which preserves the Lie bracket, that is

$$\varphi([x, y]) = [\varphi(x), \varphi(y)] \tag{13.1.3}$$

where the bracket on the left is that of $\mathfrak{g}$ and on the right it's that of $\mathfrak{g}'$.

Example 13.1.4

- Let A be an associative algebra, then we can make this into a Lie algebra by defining the bracket $[a,b] = ab - ba$. A special case of this is $A = \operatorname{End} V$ for some vector space, V. Then we call the corresponding Lie algebra $\mathfrak{gl}(V)$, or if $\dim V = n$ we call it $\mathfrak{gl}_n$ (note that as vector spaces $\mathfrak{gl}(V)$ is exactly $A = \operatorname{End} V$, the name change just reflects a shifting view point from associative algebras to Lie algebras).

- Any vector space, V, can be made into a Lie algebra by defining $[x,y] = 0$ for all $x,y \in V$. Such a Lie algebra is called **abelian**. The idea is that the commutator vanishing means that multiplication is commutative, an idea that only makes sense if $[-,-]$ really is the commutator, like in the previous example.

Definition 13.1.5 — Lie Subalgebra Let $\mathfrak{g}$ be a Lie algebra over $\Bbbk$. A Lie subalgebra, $\mathfrak{h}$, is a Lie algebra over $\Bbbk$ equipped with an injective Lie algebra morphism $\mathfrak{h} \hookrightarrow \mathfrak{g}$.

An almost identical definition is that a Lie subalgebra is a subspace, $\mathfrak{h} \subseteq \mathfrak{g}$ such that $\mathfrak{h}$ is a Lie algebra in its own right (with the same bracket as $\mathfrak{g}$). One can then show that this is true so long as the $\mathfrak{h}$ is closed under the Lie bracket. That is, $[\mathfrak{h},\mathfrak{h}]$ is a subset of $\mathfrak{h}$. Note that in general if U and V are subspaces of $\mathfrak{g}$ then $[U,V]$ is defined to be the span of all $[u,v]$ with $u \in U$ and $v \in V$. Similarly, if $x \in \mathfrak{g}$ then $[x,U]$ is the span of all $[x,y]$ with $y \in \mathfrak{g}$.

The only subtle difference between these two definitions is that the existence of a monomorphism $\mathfrak{h} \hookrightarrow \mathfrak{g}$ only implies that $\mathfrak{h}$ is isomorphic to a subalgebra of $\mathfrak{g}$ with the second definition, but we'll only consider things up to isomorphism most the time so this is really the definition we want.

Example 13.1.6

- Let $\mathfrak{g}$ be any Lie algebra. Any one-dimensional subspace, $\mathfrak{l}$, is an abelian subalgebra, since if $l,l' \in \mathfrak{l}$ then $l = \lambda l'$ for some $\lambda \in \Bbbk$, and so $[l,l'] = [kl',l'] = k[l',l'] = 0$ and $0 \in \mathfrak{l}$.
- The **centre** of a Lie algebra, $\mathfrak{g}$, is the abelian subalgebra
$$\mathfrak{z}(\mathfrak{g}) \coloneqq \{x \in \mathfrak{g} \mid [x,\mathfrak{g}] = 0\} \subseteq \mathfrak{g}. \tag{13.1.7}$$
- For V a finite-dimensional vector space of dimension n we know that $\mathfrak{gl}_n = \operatorname{End} V$ is a Lie algebra. Fixing a basis the elements of $\mathfrak{gl}_n$ are just all $n \times n$ matrices with entries in $\Bbbk$. There is a subalgebra, $\mathfrak{sl}_n \subset \mathfrak{gl}_n$, consisting of only the matrices with zero trace. This follows because we have
$$\operatorname{tr}([x,y]) = \operatorname{tr}(xy) - \operatorname{tr}(yx) = 0. \tag{13.1.8}$$
This holds for all $x,y \in \mathfrak{gl}_n$, not just for the traceless case, and so this turns out to be a special case of another construction, called the derived subalgebra, $\mathfrak{g}' \coloneqq [\mathfrak{g},\mathfrak{g}]$.

Definition 13.1.9 — Ideal Let $\mathfrak{g}$ be a Lie algebra. A Lie subalgebra, $\mathfrak{i} \subseteq \mathfrak{g}$, is an **ideal** if $[\mathfrak{i}, \mathfrak{g}] \subseteq \mathfrak{i}$.

Compare this to the definition of a subalgebra, which only requires that $[\mathfrak{i}, \mathfrak{i}] \subseteq \mathfrak{i}$. Compare this also to the notion of an ideal, I, of a ring, R, which is a subgroup of the additive group such that $IR \subseteq I$.

The idea is that ideals are to Lie algebras as ideals are to rings, or as normal subgroups are to groups. In particular, we have a correspondence between ideals, $\mathfrak{i} \subseteq g$ and Lie algebra morphisms, $\varphi : \mathfrak{g} \to \mathfrak{h}$ given by $\mathfrak{i} \leftrightarrow \ker \varphi$ (where the kernel is defined as it is for any linear map). We also have that $\mathfrak{g}/\mathfrak{i}$ is a well defined quotient and a Lie algebra. Note that the quotient of any vector space by a subspace is again a vector space, but it's only a Lie algebra again if we quotient by an ideal. The bracket of this quotient is defined by $[x + \mathfrak{i}, y + \mathfrak{i}] = [x, y] + \mathfrak{i}$.

Definition 13.1.10 — Derived Subalgebra Let $\mathfrak{g}$ be a Lie algebra, then $\mathfrak{g}' = [\mathfrak{g}, \mathfrak{g}]$ is the **derived subalgebra**.

Definition 13.1.11 — Solvable Lie Algebra A Lie algebra, $\mathfrak{g}$, is solvable if the series

$$\mathfrak{g} \supseteq \mathfrak{g}' \supseteq \mathfrak{g}'' \supseteq \cdots \tag{13.1.12}$$

terminates.

Definition 13.1.13 — Nilpotent Lie ALgebra A Lie algebra, $\mathfrak{g}$, is solvable if the series

$$\mathfrak{g} \supseteq [\mathfrak{g}, \mathfrak{g}] \supseteq [\mathfrak{g}, [\mathfrak{g}, \mathfrak{g}]] \supseteq \cdots \tag{13.1.14}$$

terminates.

The difference between these two is subtle, one nests brackets on both sides, and the other only on the other side. More concretely, the upper triangular matrices form a solvable subalgebra of $\mathfrak{gl}_n$ (in fact, this is a maximal solvable subalgebra, also known as a **Borel subalgebra**), and the *strictly* upper triangular matrices form a (maximal) nilpotent subalgebra of $\mathfrak{gl}_n$.

Definition 13.1.15 The maximal solvable *ideal* of $\mathfrak{g}$ is called its **radical**, $\operatorname{Rad} \mathfrak{g}$.

Definition 13.1.16 A Lie algebra, $\mathfrak{g}$, is **semisimple** if $\operatorname{Rad} \mathfrak{g} = 0$, that is, if $\mathfrak{g}$ has no proper solvable ideals. Similarly, $\mathfrak{g}$ is **simple** if it has no proper ideals (solvable or not).

Definition 13.1.17 — Linear Lie Algebra A **linear Lie algebra** is any Lie algebra which is isomorphic to a Lie subalgebra of some $\mathfrak{gl}(V)$ for V a finite-dimensional vector space.

Ado's theorem tells us that (over a field of characteristic zero) every finite-dimensional Lie algebra is linear.

Theorem 13.1.18 — Ado's Theorem. Let $\mathfrak{g}$ be a finite-dimensional Lie algebra over a field of characteristic zero. Then $\mathfrak{g}$ admits a faithful representation $\mathfrak{g} \hookrightarrow \mathfrak{gl}(V)$ for some finite-dimensional vector space, V. Further, one can choose this representation such that the maximal nilpotent ideal, $\mathfrak{n} \subseteq \mathfrak{g}$ acts nilpotently on V.

There are some special linear Lie algebras. Over $\mathbb{C}$ these are

- $\mathfrak{gl}_n = \{x \in \mathrm{Mat}_n(\mathbb{C}\}$ (real dimension $2n^2$)
- $\mathfrak{sl}_n = \{x \in \mathrm{Mat}_n(\mathbb{C}) \mid \operatorname{tr} x = 0\}$ (real dimension $2(n^2 - 1)$);
- $\mathfrak{so}_n = \{x \in \mathrm{Mat}_n(\mathbb{C}) \mid x^\top + x^\top = 0\}$ (real dimension $n(n-1)$);
- $\mathfrak{sp}_{2n} = \{x \in \mathrm{Mat}_{2n}(\mathbb{C}) \mid Jx + x^\top J = 0\}$ where $J = \begin{pmatrix} 0 & -I_n \\ I_n & 0 \end{pmatrix}$ with $I_n \in \mathrm{Mat}_n(\mathbb{C})$ the identity matrix (real dimension $2\binom{2n+1}{2}$).

Over $\mathbb{R}$ these are

- $\mathfrak{gl}_n = \{x \in \mathrm{Mat}_n(\mathbb{R})\}$ (real dimension n^2);
- $\mathfrak{so}_n = \{x \in \mathrm{Mat}_n(\mathbb{R}) \mid \operatorname{tr} x = 0\}$ (real dimension $n^2 - 1$);
- $\mathfrak{u}_n = \{x \in \mathrm{Mat}_n(\mathbb{C}) \mid x + x^* = 0\}$ (real dimension n^2);
- $\mathfrak{su}_n = \{x \in \mathrm{Mat}_n(\mathbb{C}) \mid x + x^* = 0 \text{ and } \operatorname{tr} x = 0\}$ (real dimension $n^2 - 1$);
- $\mathfrak{sp}_{2n} = \{x \in \mathrm{Mat}_n(\mathbb{H}) \mid x + x^* = 0\}$ (real dimension $2n^2 + n$).

13.2 Representation Theory of Lie Algebras

Definition 13.2.1 — Representation A **representation**, $\mathfrak{g}$ (over $\Bbbk$), is a $\Bbbk$-vector space, V, equipped with a Lie algebra morphism

$$\rho \colon \mathfrak{g} \to \mathfrak{gl}(V). \tag{13.2.2}$$

Equivalently, a $\mathfrak{g}$-**module**, V, is a vector space equipped with a (left) Lie algebra action of $\mathfrak{g}$, that is, a map $\mathfrak{g} \times V \to V$, $(x, v) \mapsto x \,.\, v$ subject to the following:

- Linearity in the first argument: $(\alpha x + \beta y) \,.\, v = \alpha(x \,.\, v) + \beta(y \,.\, v)$ for all $\alpha, \beta \in \Bbbk$, $x, y \in \mathfrak{g}$ and $v \in V$;
- Linearity in the second argument: $x \,.\, (\alpha v + \beta w) = \alpha(x \,.\, v) + \beta(x \,.\, w)$ for all $\alpha, \beta \in \Bbbk$, $x \in \mathfrak{g}$ and $v, w \in V$;

- Respects the bracket: $[x, y] . v = x . (y . v) - y . (x . v)$ for all $x, y \in \mathfrak{g}$ and $v \in V$.

As with groups and associative algebras the $\mathfrak{g}$-module and representation of $\mathfrak{g}$ carry exactly the same information, and as such which we use is a matter of preference.

Definition 13.2.3 — Adjoint Representation Every Lie algebra, $\mathfrak{g}$, is a $\mathfrak{g}$-module in a canonical way, known as the **adjoint representation**

$$\begin{aligned} \mathrm{ad} : \mathfrak{g} &\to \mathfrak{gl}(\mathfrak{g}) \\ x &\mapsto \mathrm{ad}_x \end{aligned} \tag{13.2.4}$$

where $\mathrm{ad}_x : \mathfrak{g} \to \mathfrak{g}$ is defined by $\mathrm{ad}_x(y) = [x, y]$ for all $x, y \in \mathfrak{g}$.

For the adjoint representation to be a representation we need ad to be a Lie algebra morphism. That is, we need to have $\mathrm{ad}_{[x,y]} = [\mathrm{ad}_x, \mathrm{ad}_y]$ for $x, y \in \mathfrak{g}$. It turns out that this is true precisely because the this statement, upon applying both sides of the above to $z \in \mathfrak{g}$, expands to the Jacobi identity:

$$\mathrm{ad}_{[x,y]}(z) = [[x, y], z] \tag{13.2.5}$$

$$[\mathrm{ad}_x, \mathrm{ad}_y](z) = (\mathrm{ad}_x \circ \mathrm{ad}_y - \mathrm{ad}_y \circ \mathrm{ad}_x)(z) = [x, [y, z]] - [y, [x, z]]. \tag{13.2.6}$$

Equality between the two lines above is, after applying the antisymmetry property, exactly the Jacobi identity.

Definition 13.2.7 Given $\mathfrak{g}$-modules V and W we can define

- the **direct sum**, $V \oplus W$, which has the action $x . (v + w) = x . v + x . w$;
- the **tensor product**, $V \otimes W$, which has the action $x . (v \otimes w) = (x . v) \otimes w + v \otimes (x . w)$;
- the **dual representation**, V^*, which has the action $\rho_{V^*}(x) = -\rho_V(x)^*$

all for $x \in \mathfrak{g}$, $v \in V$, and $w \in W$.

13.3 Universal Enveloping Algebra

Definition 13.3.1 — Universal Enveloping Algebra Let $\mathfrak{g}$ be a Lie algebra. An enveloping algebra, (E, i), is an associative unital algebra, E, and an inclusion of vector spaces $i : \mathfrak{g} \hookrightarrow E$ such that

$$i([x, y]) = i(x)i(y) - i(y)i(x). \tag{13.3.2}$$

The **universal enveloping algebra** is the[a] enveloping algebra $(U(\mathfrak{g}), \iota)$

such that for any other enveloping algebra, (E, i), there is a unique morphism of associative unital algebras, $\varphi\colon\ U(\mathfrak{g}) \to E$ such that $i = \varphi \circ \iota$.

[a] turns out that the universal enveloping algebra both exists, and is unique up to unique isomorphism

The definition is a bit terse, the idea is that $U(\mathfrak{g})$ (dropping ι from the notation) is the smallest associative unital algebra containing $\mathfrak{g}$ in such a way that the bracket of $\mathfrak{g}$ in $U(\mathfrak{g})$ really is just the commutator. For example, the universal enveloping algebra of $\mathfrak{gl}(V)$ is simply $\mathrm{End}(V)$, which is just $\mathfrak{gl}(V)$ but viewed as an associative algebra.

Theorem 13.3.3. The universal enveloping algebra exists. An explicit construction is as follows. Let $U(\mathfrak{g}) = T(\mathfrak{g})/I$, where I is the ideal of the tensor algebra, $T(\mathfrak{g})$, generated by elements of the form

$$[x, y] - x \otimes y + y \otimes x \tag{13.3.4}$$

for $x, y \in \mathfrak{g}$.

The universal property of the universal enveloping algebra can be characterised as the statement that there is an isomorphism

$$\mathrm{Hom}_{\Bbbk\text{-}\mathsf{Lie}}(\mathfrak{g}, L(A)) \cong \mathrm{Hom}_{\Bbbk\text{-}\mathsf{Alg}}(U(\mathfrak{g}), A) \tag{13.3.5}$$

where

- $\Bbbk$-Lie is the category of Lie algebras and Lie algebra homomorphisms;
- $\mathfrak{g}$ is a Lie algebra
- A is an unital associative algebra;
- $L(A)$ is the Lie algebra given by equipping A with the commutator;
- $\Bbbk$-Alg is the category of unital associative algebras and their homomorphisms.

Simply send the Lie algebra homomorphism $\varphi\colon\ \mathfrak{g} \to L(A)$ to the associative algebra homomorphism $\tilde{\varphi}\colon\ U(\mathfrak{g}) \to A$ defined by $\tilde{\varphi}(x) = \varphi(x)$ for $x \in \mathfrak{g}$ and extended by linearity and the requirement that $\tilde{\varphi}$ preserves multiplication. This works precisely because of the universal property. For the inverse, send $\psi\colon\ U(\mathfrak{g}) \to A$ to the restriction $\psi|_{\mathfrak{g}}$.

It turns out that $L\colon\ \Bbbk\text{-}\mathsf{Alg} \to \Bbbk\text{-}\mathsf{Lie}$ is a functor, if $f\colon\ A \to B$ is a morphism of associative algebras then we can define $L(f)\colon\ L(A) \to L(B)$ by defining $L(f)([x, y]) = [f(x), f(y)] = f(x)f(y) - f(y)f(x)$ for $x, y \in A$. That is, we just require that $L(f)$ is a Lie algebra homomorphism. Similarly, $U\colon\ \Bbbk\text{-}\mathsf{Lie} \to \Bbbk\text{-}\mathsf{Alg}$ is a functor, if $f\colon\ \mathfrak{g} \to \mathfrak{h}$ is a morphism of Lie algebras then we can define $U(f)\colon\ U(\mathfrak{g}) \to U(\mathfrak{h})$ by defining $U(f)(xy) = U(f)(x)U(f)(y)$ for $x, y \in \mathfrak{g}$ and similarly for products of more than two elements, and extended by linearity to all of $U(\mathfrak{g})$. That is, we just require that $U(f)$ respects the multiplication of the associative algebra. Then the above isomorphism happens to be natural, and we thus have that L is right adjoint to U.

The important thing here is that if we take $A = \operatorname{End} V$ then we have

$$\operatorname{Hom}_{\Bbbk\text{-Lie}}(\mathfrak{g}, \mathfrak{gl}(V)) \cong \operatorname{Hom}_{\Bbbk\text{-Alg}}(U(\mathfrak{g}), \operatorname{End} V). \tag{13.3.6}$$

This means that a map $\mathfrak{g} \to \mathfrak{gl}(V)$ carries the same data as a map $U(\mathfrak{g}) \to \operatorname{End} V$. We can identify a map of the first type as a Lie algebra representation of $\mathfrak{g}$, and a map of the second type as a unital associative algebra representation of $U(\mathfrak{g})$. That is, representations of $\mathfrak{g}$ are "the same" as representations of $U(\mathfrak{g})$.

Another way of thinking about this is that $U(\mathfrak{g})$ is to $\mathfrak{g}$ as $\Bbbk G$ is to G for a finite group, G. We can study the representation theory of $\mathfrak{g}$ or G just by studying the representation theory of the universal enveloping algebra or group algebra.

Proposition 13.3.7 The universal enveloping algebra, $U(\mathfrak{g})$, is a Hopf algebra with the comultiplication

$$\Delta(x) = x \otimes 1 + 1 \otimes x, \tag{13.3.8}$$

counit

$$\varepsilon(x) = 0, \tag{13.3.9}$$

and antipode

$$\chi(x) = -x. \tag{13.3.10}$$

Compare and contrast this to the group algebra, $\Bbbk G$, which is a Hopf algebra with

$$\Delta(g) = g \otimes g, \quad \varepsilon(g) = 1, \quad \text{and} \quad \chi(g) = g^{-1}. \tag{13.3.11}$$

These are, in some ways, two opposite ends of the scale for how a Hopf algebra can behave.

Definition 13.3.12 — Filtred Algebra Let A be an associative algebra. We say that A is $\mathbb{Z}_{\geq 0}$-**filtred** if we have a chain of subspaces

$$0 = F_{-1}A \subseteq F_0 A \subseteq F_1 A \subseteq \cdots \subseteq F_n A \subseteq \cdots \tag{13.3.13}$$

such that $1 \in F_0 A$,

$$\bigcup_{n=0}^{\infty} F_n A = A, \tag{13.3.14}$$

and $F_i A \cdot F_j A \subseteq F_{i+j} A$.

Definition 13.3.15 — Degree Filtration If A is an associative algebra generated by $\{x_\alpha\}$ then we can define a filtration on A by declaring all x_α to be of degree 1, and defining $F_n A := (F_1 A)^n$ to be formed of all terms of degree at most n (note that the degree of $x_\alpha x_{\alpha'}$ is 2, as is the degree of x_α^2, and so on).

Definition 13.3.16 — Associated Graded Algebra Given a filtred algebra, A, we define the **associated graded algebra** to be

$$\operatorname{gr}(A) := \bigoplus_{n=0}^{\infty} F_n(A)/F_{n-1}(A). \tag{13.3.17}$$

For the degree filtration the associated graded algebra is

$$\operatorname{gr}(A) = \bigoplus_{n=0}^{\infty} A_n \tag{13.3.18}$$

where A_n is the span of all words of degree exactly n.

If $\mathfrak{g}$ is a Lie algebra then we can define a degree filtration on $U(\mathfrak{g})$ by setting the degree of any $x \in \mathfrak{g}$ to be 1. Then $F_nU(\mathfrak{g})$ is the image of $\bigoplus_{k=0}^{n} \mathfrak{g}^{\otimes k} \subset T(\mathfrak{g})$ under the quotient map $T(\mathfrak{g}) \twoheadrightarrow T(\mathfrak{g})/I$. Since in $U(\mathfrak{g})$ we have $xy - yx = [x, y]$ for $x \in \mathfrak{g}$ and $y \in U(\mathfrak{g})$ it follows that $[F_iU(\mathfrak{g}), F_jU(\mathfrak{g})] \subseteq F_{i+j-1}U(\mathfrak{g})$. It then follows that when we take $F_nU(\mathfrak{g})/F_{n-1}U(\mathfrak{g})$ in $\operatorname{gr}(U(\mathfrak{g}))$ we are quotenting by (among other things) all commutators of elements of degree less than n. This makes $\operatorname{gr}(U(\mathfrak{g}))$ commutative. This in turn means that there is an epimorphism of associative algebras

$$S(\mathfrak{g}) \twoheadrightarrow \operatorname{gr}(U(\mathfrak{g})). \tag{13.3.19}$$

This is a statement that $S(A)$ is universal amongst commutative subalgebras of $T(A)$, i.e., that any such subalgebra can be recognised by taking $S(A)$ and applying some quotient to identify certain terms.

Definition 13.3.20 — PBW Theorem The homomorphism $S(\mathfrak{g}) \to \operatorname{gr}(U(\mathfrak{g}))$ is an isomorphism.

Corollary 13.3.21 If $\{x_i\}$ is a basis of $\mathfrak{g}$ we can fix an order on the basis. Then $U(\mathfrak{g})$ is spanned by ordered monomials $\prod_i x_i^{n_i}$ with $n_i \in \mathbb{Z}_{\geq 0}$.

Theorem 13.3.22 — PBW Theorem. The ordered monomials described above are actually linearly independent, and thus form a basis for $U(\mathfrak{g})$.

Example 13.3.23 Consider $\mathfrak{sl}_2(\mathbb{C})$. This is a three-dimensional Lie algebra with generators $\{e, h, f\}$. If we order them so that $e < h < f$ then a basis for $U(\mathfrak{sl}_2(\mathbb{C}))$ is $e^a h^b f^c$ with $a, b, c \in \mathbb{Z}_{\geq 0}$.

13.4 Representation Theory of $\mathfrak{sl}_2(\mathbb{C})$

The representation theory of all finite dimensional semisimple Lie algebras over $\mathbb{C}$ is almost entirely controlled by the representation theory of $\mathfrak{sl}_2$. For this reason we'll now devote some time to the study of $\mathfrak{sl}_2$.

Recall that $\mathfrak{sl}_2$ (working over $\mathbb{C}$) is defined to consist of all traceless 2×2 complex matrices. There is a basis for these given by

$$e = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad h = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \text{and} \quad f = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}. \tag{13.4.1}$$

One can check that these satisfy the commutation relations

$$[h, e] = 2e, \quad [h, f] = -2f, \quad \text{and} \quad [e, f] = h. \tag{13.4.2}$$

We can then abstract the definition of $\mathfrak{sl}_2$ to be $\operatorname{span}_{\mathbb{C}}\{e, h, f\}$ subject to the above commutation relations, without needing an explicit matrix form.

Lemma 13.4.3 Let V be a finite-dimensional representation of $\mathfrak{sl}_2$. Then we have the decomposition

$$V \cong \bigoplus_{\alpha \in \mathbb{C}} V_\alpha \tag{13.4.4}$$

where V_α is the **weight space**, defined to be the eigenspace

$$V_\alpha = \{v \in V \mid h \,.\, v = \alpha v\}. \tag{13.4.5}$$

Proof. It is a fact that finite-dimensional $\mathfrak{sl}_2$-representations are completely reducible. Thus, we may assume without loss of generality that V is irreducible, since if it isn't we can decompose it into a sum of irreducibles and then treat each of these separately.

Let W be the subspace of eigenvectors of h. It is then sufficient to show that $W = V$. To do this we show that W is a subrepresentation, that is, it's closed under h, e, and f. Then irreducibility will imply that $W = V$.

By definition h acts as a scalar on W, so W is closed under h. For e let $v \in W$ be an eigenvector of h, that is $hv = \alpha v$. Then a direct computation gives

$$\begin{aligned} he \,.\, v &= ([h, e] + eh) \,.\, v && (13.4.6)\\ &= (2e + eh) \,.\, v && (13.4.7)\\ &= 2e \,.\, v + eh \,.\, v && (13.4.8)\\ &= 2e \,.\, v + \alpha e \,.\, v && (13.4.9)\\ &= (\alpha + 2)e \,.\, v. && (13.4.10) \end{aligned}$$

Thus, $e \,.\, v$ is again an eigenvector of h, with eigenvalue $\alpha + 2$. Similarly, one can show that $f \,.\, v$ is an eigenvector of h with eigenvalue $\alpha - 2$.

Thus, W is closed under the action of e, h, and f, and thus is a subrepresentation, and so by irreducibility $W = V$. Thus, if V is not irreducible is a direct sum of irreducibles, each of which is an eigenspace of h with some given eigenvalue α. We may as well sum over all possible eigenvalues, $\alpha \in \mathbb{C}$, and simply have $V_\alpha = 0$ for many terms. $\square$

Example 13.4.11 The definition of $\mathfrak{sl}_2$ in terms of 2×2 matrices gives us a natural action of $\mathfrak{sl}_2$ on $\mathbb{C}^2$. Let $\{e_1, e_2\}$ be the standard basis of $\mathbb{C}^2$. We have $he_1 = e_1$ and $he_2 = -e_2$, so we have two eigenvectors, and the corresponding eigenspaces $V_1 = \mathbb{C}e_1$ and $V_{-1} = \mathbb{C}e_2$. Then we have the following picture:

$$\begin{tikzcd}
V_1 \arrow[loop right, "h"] \arrow[d, bend left, "f"] \\
V_{-1} \arrow[loop right, "h"] \arrow[u, bend left, "e"]
\end{tikzcd} \tag{13.4.12}$$

The interpretation of this picture is that e and f act to shift the eigenvalue up and down by 2. Note that applying e to e_1 gives $ee_1 = 0$, and likewise, $fe_2 = 0$. Thus, we can add 0 to the top and bottom of this picture:

$$\begin{tikzcd}
0 \arrow[d, bend left, "f"] \\
V_1 \arrow[loop right, "h"] \arrow[d, bend left, "f"] \arrow[u, bend left, "e"] \\
V_{-1} \arrow[loop right, "h"] \arrow[u, bend left, "e"] \arrow[d, bend left, "f"] \\
0 \arrow[u, bend left, "e"]
\end{tikzcd} \tag{13.4.13}$$

The picture above actually generalises to any finite dimensional representation, we can always draw a picture like the following:

$$
\begin{tikzcd}
0 \arrow[d, bend left, "f"] \\
V_{\alpha+2k} \arrow[u, "e"] \arrow[loop right, "h"] \arrow[d, bend left, "f"] \\
\vdots \arrow[u, bend left, "e"] \arrow[d, bend left, "f"] \\
V_{\alpha+2} \arrow[u, bend left, "e"] \arrow[loop right, "h"] \arrow[d, bend left, "f"] \\
V_{\alpha} \arrow[u, bend left, "e"] \arrow[loop right, "h"] \arrow[d, bend left, "f"] \\
V_{\alpha-2} \arrow[u, bend left, "e"] \arrow[loop right, "h"] \arrow[d, bend left, "f"] \\
\vdots \arrow[u, bend left, "e"] \arrow[d, bend left, "f"] \\
V_{\alpha-2\ell} \arrow[u, bend left, "e"] \arrow[loop right, "h"] \arrow[d, bend left, "f"] \\
0 \arrow[u, bend left, "e"]
\end{tikzcd}
\tag{13.4.14}
$$

The fact that we must always eventually get to 0 going either up or down is simply due to the fact that V is finite-dimensional.

Example 13.4.15 Consider the vector space $S^k(\mathbb{C}^2)$. We may identify this with the space of degree k homogenous polynomials (with coefficients in $\mathbb{C}$). For example, for $S^3(\mathbb{C}^2)$ we identify $e_1 \otimes e_1 \otimes e_1$ with x^3, $e_1 \otimes e_1 \otimes e_2 = e_1 \otimes e_2 \otimes e_1 = e_2 \otimes e_1 \otimes e_1$ with $x^2 y$, and so on. Basically, send e_1 to x, e_2 to y, and remember that all tensor products are symmetrised. Note then that we can identify $S(\mathbb{C}^2)$ and $\mathbb{C}[x, y]$ (more generally, $S(\mathbb{C}^m)$ and $\mathbb{C}[x_1, \dots, x_m]$), an important identification in algebraic geometry.
There is a representation of $\mathfrak{sl}_2$ on $\mathbb{C}[x, y]$ given by

$$e = -y\partial_x, \quad h = -x\partial_x + y\partial_y, \quad \text{and} \quad f = -x\partial_y. \tag{13.4.16}$$

Note that each operator preserves the total degree of any polynomial (so long as it doesn't send it to zero). Thus, we can identify submodules of degree k-polynomials. More generally, the above identification defines an action of $\mathfrak{sl}_2$ on smooth functions $\mathbb{C}^2 \to \mathbb{C}$, of which the $S^k(\mathbb{C}^2)$ are submodules.
Consider $S^k(\mathbb{C}^2)$, which we now identify with the space of degree k poly-

nomials in x and y. A basis for this space consists of vectors

$$v_r = \binom{k}{r} x^r y^{k-r}. \tag{13.4.17}$$

Acting on this with h we have

$$\begin{aligned} hv_r &= (-x\partial_x + y\partial_y)\binom{k}{r} x^r y^{k-r} \\ &= -r\binom{k}{r} x^r y^{k-r} - (k-r)\binom{k}{r} x^r y^{k-r}) = (k-2r)v_r, \end{aligned} \tag{13.4.18}$$

so v_r has h-eigenvalue $\alpha = k - 2r$. We also have

$$ev_r = -y\partial_x \binom{k}{r} x^r y^{k-r} = -r\binom{k}{r} x^{r-1} y^{k-r+1} = (r-k-1)v_{r-1} \tag{13.4.19}$$

and the h-eigenvalue of v_{r-1} is $k-2(r-1) = k-2r+2 = \alpha+2$. Similarly, we have

$$fv_r = -x\partial_y \binom{k}{r} x^r y^{k-r} = -(k-r)\binom{k}{r} x^{r+1} y^{k-r-1} = -(1+r)v_{r+1} \tag{13.4.20}$$

and the h-eigenvalue of v_{r+1} is $k-2(r+1) = k-2r-2 = \alpha-2$. Then letting $V_{k-2r} = \mathbb{C}v_r$ we have

$$\begin{tikzcd} V_{k-2r+2} \arrow[loop right, "h\sim k-2r"] \arrow[d, bend left, "f\sim -(1+r)"] \\ V_{k-2r} \arrow[loop right, "h\sim k-2r"] \arrow[u, bend left, "e\sim r-k-1"] \end{tikzcd} \tag{13.4.21}$$

Here $a \sim \lambda$ we mean that a acts by sending the basis vector of one space to the basis vector of the next multiplied by λ.
Let $V(k) = S^k(\mathbb{C}^2)$ be this $\mathfrak{sl}_2$-module. This is an irreducible module. Given any basis vector it lives in one of the V_α, and if we continuously act with e we eventually get v_0. Then v_0 generates this entire module by acting with f and scalar multiplication. Note that $\dim V(k) = k+1$, since we have the basis $\{v_0, \dots, v_k\}$.

The previous example actually captures all irreducible modules of $\mathfrak{sl}_2$, as the following proves. The argument basically mirrors the argument above without reference to an explicit structure of polynomials.

Proposition 13.4.22 — Classification of Finite Dimensional Irreducible $\mathfrak{sl}_2$-Modules Let V be a $(k+1)$-dimensional $\mathfrak{sl}_2$-module. Then $V \cong V(k)$ with $V(k)$ as defined in Example 13.4.15.

Proof. By the same argument as in the proof of Lemma 13.4.3 we know that the eigenvectors of h span V (which we're assuming is irreducible). Since V is finite-dimensional h has a finite number of eigenvalues, so there must be some h-eigenvector, v_0, for which we have $hv_0 = 0$. Consider $f^k v_0$, as we have a finite-dimensional space, and thus finitely many eigenvectors of h, we must have for some N that $f^N v_0 = 0$, and suppose N is the smallest such value. If we take $B = \{v_0, fv_0, \dots, f^{N-1}v_0\}$ then this is a submodule of V, and thus is all of V. Thus, knowing that V has dimension $k+1$ we know that $N = k+1$. In particular, $f^{N-1}v_0 = f^k v_0$ is the last element of this basis.
For what follows it's useful to absorb some scale factor into the basis, define $v_r = f^r v_0/r!$ for $r = 0, \dots, k$. Then $\{v_r\}$ is a basis of V.
All that remains is to show that the action of e and f on this basis is fully determined. Starting with e we use the fact that $hv_r = (\alpha_0 - 2r)v_r$ where α_0 is the h-eigenvalue of v_0. We then have

$$\begin{aligned} ev_0 &= 0 && (13.4.23)\\ ev_1 &= efv_0 = [e,f]v_0 + fev_0 = hv_0 + 0 = \alpha_0 v_0 && (13.4.24)\\ ev_2 &= efv_1/2 = [e,f]v_1/2 + fev_1/2 = hv_1/2 + \alpha_0 fv_0/2 && (13.4.25)\\ &= (\alpha_0 - 2)v_1/2 + \alpha_0 v_1/2 = (\alpha_0 - 1)v_1. && (13.4.26) \end{aligned}$$

We thus make the induction hypothesis that

$$ev_n = (\alpha_0 - n + 1)v_{n-1}. \tag{13.4.27}$$

Assuming the equivalent statement for v_{n-1} holds we then have

$$\begin{aligned} ev_n &= efv_{n-1}/n = [e,f]v_{n-1}/n + fev_{n-1}/n && (13.4.28)\\ &= hv_{n-1}/n + fev_{n-1}/n && (13.4.29)\\ &= (\alpha_0 - 2n + 2)v_{n-2} + (\alpha_0 - n + 2)fv_{n-2}/n && (13.4.30)\\ &= (\alpha_0 - 2n + 2)v_{n-1}/n + (n-1)(\alpha_0 - n + 2)v_{n-1}/n && (13.4.31)\\ &= (\alpha_0 - n + 1)v_{n-1}. && (13.4.32) \end{aligned}$$

This shows that the structure of V is entirely determined by α_0, we now show that α_0 is fixed. We know that $fv_k = 0$, and we have

$$\begin{aligned} efv_k &= [e,f]v_k + fev_k = hv_k + (\alpha_0 - k + 1)fv_{k-1} && (13.4.33)\\ &= (\alpha_0 - 2k)v_k + (\alpha_0 - k + 1)kv_k && (13.4.34)\\ &= (k+1)(\alpha_0 - k)v_{k-1}. && (13.4.35) \end{aligned}$$

For this to vanish, given that $k+1$, the dimension, is positive (for $k+1 = 0$ clearly all zero dimensional $\mathfrak{sl}_2$-modules are isomorphic), and thus $\alpha_0 = k$ is fixed, and so as soon as we know the dimension of a finite-dimensional irreducible $\mathfrak{sl}_2$-module we know everything about it. $\square$

Definition 13.4.36 — Weight Vectors Let V be an $\mathfrak{sl}_2$-module. We call eigenvectors of h **weight vectors**, and the eigenvalue is called its weight. If v is a weight vector and $ev = 0$ we call v a **highest weight vector**, similarly, if $fv = 0$ we call v a **lowest weight vector**.

The above proposition then says that any finite-dimensional irreducible $\mathfrak{sl}_2$-module is generated by a highest weight vector, v_0.

13.5 Classification of Semisimple Lie Algebras Over $\mathbb{C}$

The steps followed for classifying irreducible finite-dimensional irreducible $\mathfrak{sl}_2$-modules actually generalise remarkably well to classifying not just representations of other Lie algebras, but classifying a whole type of algebra, just by studying the adjoint representations in which these algebras act on themselves.

There were three steps we followed with $\mathfrak{sl}_2$. First, decompose V into eigenspaces of h. Second, use the commutation relations to determine how e and f act on these eigenspaces. Finally, use the irreducibility of the module to show that it is generated by a single highest weight vector.

In order to apply this method to other Lie algebras we'll need to generalise some things. The main one is that instead of just a single operator, h, we end up with a whole subalgebra of operators, $\mathfrak{h}$. Before we get to this we need a few definitions.

Definition 13.5.1 — Semisimple and Nilpotent Elements Let $\mathfrak{g}$ be a Lie algebra. We say that $x \in \mathfrak{g}$ is **semisimple** if ad_x is diagonalisable, and **nilpotent** if ad_x is nilpotent.

For example, in $\mathfrak{sl}_2$ h is semisimple, since in the adjoint representation, with the ordered basis $\{e, h, f\}$, we have

$$\mathrm{ad}_h = \begin{pmatrix} 2 & & \\ & 0 & \\ & & -2 \end{pmatrix}. \tag{13.5.2}$$

On the other hand, e and f are nilpotent, since in the adjoint representation

$$\mathrm{ad}_e = \begin{pmatrix} 0 & -2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad \text{and} \quad \mathrm{ad}_f = \begin{pmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 2 & 0 \end{pmatrix}, \tag{13.5.3}$$

both of which have vanishing third power.

An abelian subalgebra, $\mathfrak{h} \subseteq \mathfrak{g}$ is called **toral**[1] if it consists of only semisimple elements. For any toral subalgebra we have the following decomposition:

$$\mathfrak{g} = \bigoplus_{\alpha \in \mathfrak{h}^*} \mathfrak{g}_\alpha \tag{13.5.4}$$

where

$$\mathfrak{g}_\alpha = \{x \in \mathfrak{g}_\alpha \mid \mathrm{ad}_h(x) = [h, x] = \alpha(h)x \text{ for } h \in \mathfrak{h}\}. \tag{13.5.5}$$

[1]This name comes from the fact that if G is a Lie group with Lie algebra $\mathfrak{g}$ then any toral subgroup, H, will have a Lie algebra isomorphic to $\mathfrak{h}$. In turn, a toral subgroup is a Lie subgroup of G which is isomorphic to a torus.

This is simply the weight space decomposition of $\mathfrak{g}$ viewed as an $\mathfrak{h}$-module through (restricted) adjoint action.

One can show that

$$[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subseteq \mathfrak{g}_{\alpha+\beta}. \tag{13.5.6}$$

In particular, $\mathfrak{g}_0$ is a Lie subalgebra, since $[\mathfrak{g}_0, \mathfrak{g}_0] \subseteq \mathfrak{g}_0$, and $\mathfrak{h} \subseteq \mathfrak{g}_0$.

Definition 13.5.7 — Cartan Subalgebra If $\mathfrak{g}$ is a Lie algebra with toral subalgebra, $\mathfrak{h}$, such that, with the notation above, we have $\mathfrak{g}_0 = \mathfrak{h}$ then we call $\mathfrak{h}$ a **Cartan subalgebra** of $\mathfrak{g}$.

Note that while Cartan subalgebras aren't unique they are all conjugate, so we typically speak of *the* Cartan subalgebra, when it exists.

When we have a Cartan subalgebra we can change the decomposition to

$$\mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{\alpha\in\Delta} \mathfrak{g}_\alpha \tag{13.5.8}$$

where $\Delta = \{\alpha \in \mathfrak{h}^* \setminus 0 \mid \mathfrak{g}_\alpha \neq 0\}$ is the subset of $\mathfrak{h}^*$ for which $\alpha \neq 0$ and $\mathfrak{g}_\alpha$ is nontrivial. We call Δ a set of **simple roots**.

For example, for $\mathfrak{sl}_2$ we have the Cartan subalgebra $\mathfrak{h} = \mathbb{C}h$. In this case we have $\mathfrak{g}_2 = \mathbb{C}e$ and $\mathfrak{g}_{-2} = \mathbb{C}f$, and we get the decomposition

$$\mathfrak{sl}_2 = \mathbb{C}h \oplus \mathbb{C}e \oplus \mathbb{C}f. \tag{13.5.9}$$

13.5.1 Root Systems

Definition 13.5.10 — Reflection Let E be a Euclidean space with inner product $(-,-)\colon E \otimes E \to \mathbb{R}$. A **reflection** is a linear map $s\colon E \to E$ such that there exists some $v \in E$ such that $s(v) = -v$ and the hyperplane $(\mathbb{R}v)^\perp$ is fixed pointwise by s. Then we call s a reflection along v.

Note that given v the following formula gives a reflection along v:

$$s_v(w) = w - 2\frac{(v,w)}{(v,v)}v. \tag{13.5.11}$$

Definition 13.5.12 — Root System Let E be a real Euclidean space with inner product $(-,-)$. A **root system**, Φ, in E is a finite set of nonzero vectors or **roots** such that

1. $\operatorname{span}_{\mathbb{R}} \Phi = E$;
2. if $\alpha \in \Phi$ then $c\alpha \in \Phi$ only for $c = \pm 1$;
3. $s_\alpha(\Phi) = \Phi$ for $\alpha \in \Phi$;
4. $2(\alpha,\beta)/(\alpha,\alpha) \in \mathbb{Z}$.

Sometimes the second condition isn't required, root systems for which the second condition holds are known as **reduced root systems**.

$$-\alpha \longleftarrow \bullet \longrightarrow \alpha$$

Figure 13.1: The A_1 root system, $\Phi = \{\alpha, -\alpha\}$, with chosen positive roots, $\Pi = \{\alpha\}$, and simple roots, $\Delta = \{\alpha\}$.

Table 13.1: Information on the root systems of rank at most 2. Notice that $\Phi = \Pi \sqcup (-\Pi)$ and in all cases we have chosen our naming of roots such that $\Delta = \{\alpha, \beta\}$. Notice that the positive roots, Π, are always found in the cone between the simple roots.

	Φ	Π	Δ
A_1	$\pm\alpha$	α	α
$\mathrm{A}_1 \oplus \mathrm{A}_1$	$\pm\alpha, \pm\beta$	α, β	α, β
A_2	$\pm\alpha, \pm\beta, \pm(\alpha+\beta)$	$\alpha, \beta, \alpha+\beta$	α, β
B_2	$\pm\alpha, \pm\beta, \pm(\alpha+\beta), \pm(2\alpha+\beta)$	$\alpha, \beta, \alpha+\beta, 2\alpha+\beta$	α, β
G_2	$\pm\alpha, \pm\beta, \alpha+\beta, \pm(2\alpha+\beta), \pm(3\alpha+\beta)$	$\alpha, \beta, \alpha+\beta, 2\alpha+\beta, 3\alpha+\beta$	α, β

The **rank** of the root system is $\dim_{\mathbb{R}} E$.

Definition 13.5.13 — Positive and Simple Roots Given a root system we can make arbitrary choice of a hyperplane containing none of the roots. We then choose one side of this hyperplane, again, arbitrarily, and declare roots in this half to be **positive**. The **simple roots** are the positive roots which cannot be written as a sum, $\alpha + \beta$, of two elements of the positive roots, α and β, alternatively, the simple roots are precisely the subset of the positive roots which generate the positive roots through linear combinations with positive integral coefficients.

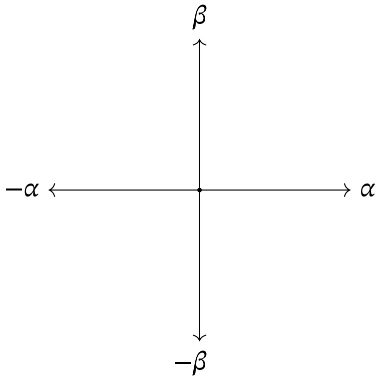
Notation 13.5.14 Notation varies here, but we'll call Φ the set of roots, Π the set of positive roots and Δ the set of simple roots.

It turns out that root systems actually turn up in many different areas of mathematics, but we'll focus on how they're relevant to Lie algebras.

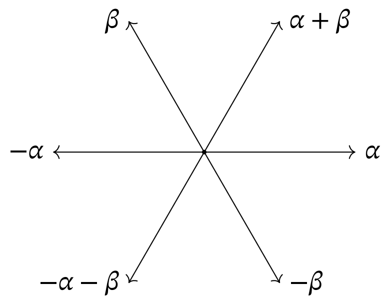
It turns out that, up to scaling, there is only one rank 1 root system. For reasons we'll get into later this root system is known as A_1. This root system is depicted in Figure 13.1. There are also only four rank 2 root systems, known as $\mathrm{A}_1 \oplus \mathrm{A}_1$ (being two orthogonal copies of A_1), A_2, B_2 (or C_2) and G_2. These are depicted in Figure 13.2. Table 13.1 lists the roots, Φ, positive roots, Π, and simple roots, Δ. In all cases we've chosen to label our roots by expressing them in terms of two chosen simple roots, α and β.

13.5.2 Connection to Semisimple Lie Algebras

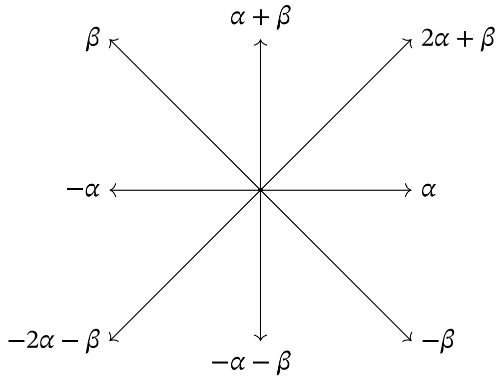
The reason that these root systems, as abstract subsets of some Euclidean space, are relevant is that given a semisimple Lie algebra the set of simple roots, Δ, (that is $\alpha \in \mathfrak{h}^*$ such that $\mathfrak{g}_\alpha \neq 0$) is actually the set of simple roots of a corresponding root system.



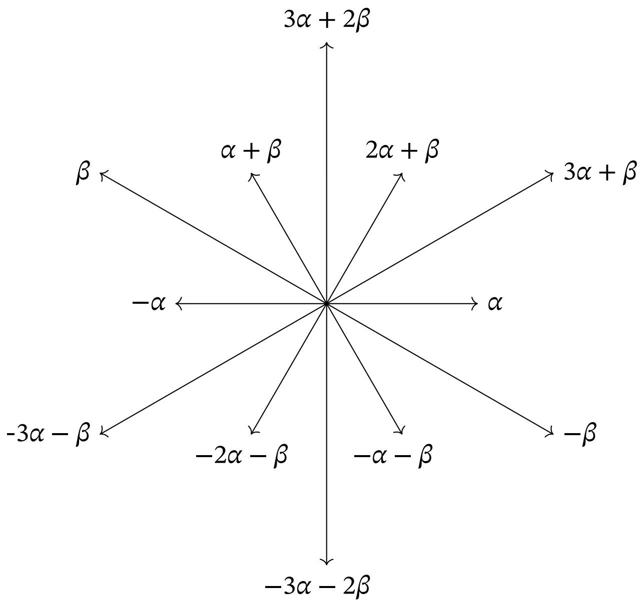
(a) The $A_1 \oplus A_1$ root system.



(b) The A_2 root system.



(c) The B_2 root system.



(d) The G_2 root system.

Figure 13.2: The rank 2 root systems.

Theorem 13.5.15. Let $\mathfrak{g}$ be a semisimple Lie algebra over $\mathbb{C}$, with Cartan subalgebra $\mathfrak{h}$. Let E be a Euclidean space such that the complexification of E is $\mathfrak{h}^*$. Then

- Δ forms a reduced root system in E;
- Eigenspaces are one-dimensional, $\mathfrak{g}_\alpha \cong \mathbb{C}$ for $\alpha \in \Delta$;
- $[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] = \mathfrak{g}_{\alpha+\beta}$.

It turns out that these properties are exactly as is required in order for the following result to hold.

Theorem 13.5.16. There is a bijection between semisimple Lie algebras over $\mathbb{C}$ and reduced root systems.

We've constructed the root system from a semisimple Lie algebra. Since these objects are in bijection we can construct a semisimple Lie algebra in a unique way from a given root system. The process is unfortunately not that insightful, and basically reduces to imposing a bunch of relations on a free Lie algebra according to information encoded in the root system. The nice thing about this result is that it turns out to be much simpler to classify all of the finite-rank root systems.

Definition 13.5.17 — Cartan Matrix A (finite-type) **Cartan matrix** is an $n\times n$ matrix, $A = (a_{ij})_{1\le i,j\le n}$ such that

- $a_{ii} = 2$ and $a_{ij} \in \mathbb{Z}_{\le 0}$ for $i \ne j$;
- A is symmetrisable (there exists some diagonal matrix, D, such that DA is a symmetric matrix);
- A is positive (all principle minors of A are positive).

We consider two Cartan matrices to be the same if they are equal up to a simultaneous permutation of the rows and columns. That is, A and B are the same if $a_{i,j} = b_{\sigma(i),\sigma(j)}$ for some $\sigma \in S_n$.

Lemma 13.5.18 Let Φ be a root system with chosen simple roots, $\Delta = \{\alpha_1, \dots, \alpha_n\}$. Define a matrix $A = (a_{ij})_{1\le i,j\le n}$ by

$$a_{ij} := \frac{2(\alpha_i, \alpha_j)}{(\alpha_i, \alpha_i)}. \tag{13.5.19}$$

This is a Cartan matrix, and is uniquely determined by the root system (up to permutation of the labels of our simple roots). Conversely, given a Cartan matrix one can construct a root system with that Cartan matrix.

The above result means that classifying Cartan matrices classifies root systems, which in turn classifies semisimple Lie algebras.

We're now ready to state the reverse process, for going from a root system or Cartan matrix to the corresponding semisimple Lie algebra.

Proposition 13.5.20 Let $A = (a_{ij})$ be an $n \times n$ Cartan matrix. Let $\mathfrak{g}$ be the Lie algebra generated by $\{e_i, h_i, f_i \mid 1 \le i \le n\}$ subject to the relations

- $[h_i, e_j] = a_{ij} e_j$;
- $[h_i, f_j] = -a_{ij} f_j$;
- $[e_i, f_j] = \delta_{ij} h_i$;
- $[h_i, h_j] = 0$;
- $(\mathrm{ad}_{e_i})^{1-a_{ij}} e_j = 0$;
- $(\mathrm{ad}_{f_i})^{1-a_{ij}} f_i = 0$.

Then this is a semisimple Lie algebra over $\mathbb{C}$ and is uniquely determined by A.

The last two relations above are called the **Serre relations**.

Note that in the above $1-a_{ij}$ is always positive, and $(\mathrm{ad}_{e_i})^k$ means the k-nested bracket with e_i, for example, $(\mathrm{ad}_{e_i})^3(x) = [e_i, [e_i, [e_i, x]]]$.

Example 13.5.21 — $\mathfrak{sl}_2$ Consider $\mathfrak{sl}_2$. We will demonstrate here that $\mathfrak{sl}_2$ is precisely the semisimple Lie algebra corresponding to A_1.
To do so we start with finding the Cartan matrix of A_1. Since $\Phi = \{\pm\alpha\}$ and $\Delta = \{\alpha\}$ this Cartan matrix is just 1×1, with the single entry being

$$a_{11} = \frac{2(\alpha, \alpha)}{(\alpha, \alpha)} = 2. \tag{13.5.22}$$

So, $A = (2)$, of course the diagonal of the Cartan matrix is, by definition, always 2s, so we didn't actually need this calculation.
Then we can take $\mathfrak{g}$ to be the Lie algebra generated by $\{e_1, h_1, f_1\}$ subject to the relations

- $[h_1, e_1] = a_{11} e_1 = 2e_2$;
- $[h_1, f_1] = -a_{11} e_1 = -2f_1$;
- $[e_1, f_1] = \delta_{11} h_1 = h_1$;
- $[h_1, h_1] = 0$.

The last of these is always true, the first three are exactly the relations on $\{e, h, f\}$ which we impose on $\mathfrak{sl}_2$, so $\mathfrak{g} \cong \mathfrak{sl}_2$.
More generally, if we construct a Lie algebra from an arbitrary root system and take the subalgebra generated by e_i, h_i and f_i for fixed i then, since $a_{ii} = 2$ we always get a copy of $\mathfrak{sl}_2$.

Example 13.5.23 — $\mathfrak{sl}_3$ Let's go one dimension up and consider A_2. This root system has $\Phi = \{\pm\alpha, \pm\beta, \pm(\alpha+\beta)\}$ and $\Delta = \{\alpha, \beta\}$. Let $\alpha_1 = \alpha$ and $\alpha_2 = \beta$ in what follows. Then the Cartan matrix has diagonals 2. Looking at the root diagram in Figure 13.2b the angle between α and β is $2\pi/3$, and both roots are the same length. Thus, $(\alpha, \beta) = (\alpha_1, \alpha_2) = \cos(2\pi/3) = -1/2$, and thus

$$a_{12} = \frac{2(\alpha_1, \alpha_1)}{(\alpha_1, \alpha_1)} = -1, \quad \text{and} \quad a_{21} = \frac{2(\alpha_2, \alpha_1)}{(\alpha_2, \alpha_2)} = 1 \tag{13.5.24}$$

having chosen a normalisation such that $(\alpha_1, \alpha_1) = (\alpha_2, \alpha_2) = 1$. The Cartan matrix of A_2 is thus

$$A = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}. \tag{13.5.25}$$

The corresponding semisimple Lie algebra is generated by $\{e_1, e_2, h_1, h_2, f_1, f_2\}$ subject to

- $[h_1, e_1] = 2e_1$, $[h_1, e_2] = -e_2$, $[h_2, e_1] = -e_1$, $[h_2, e_2] = 2e_2$;
- $[h_1, f_1] = -2e_1$, $[h_1, f_2] = f_2$, $[h_2, f_1] = f_1$, $[h_2, f_2] = -2f_2$;
- $[e_1, f_1] = h_1$, $[e_2, f_2] = h_2$, $[e_1, f_2] = [e_2, f_1] = 0$;
- $[h_1, h_2] = 0$;
- $(\mathrm{ad}_{e_1})^{1-a_{12}} e_2 = (\mathrm{ad}_{e_1})^2 e_2 = [e_1, [e_1, e_2]] = 0$, $[e_2, [e_2, e_1]] = 0$;
- $[f_1, [f_1, f_2]] = [f_2, [f_2, f_1]] = 0$.

This algebra is isomorphic to $\mathfrak{sl}_3$.

Example 13.5.26 — $\mathfrak{so}_5$ Consider the root system B_3, which has $\Delta = \{\alpha_1, \alpha_2\}$. Looking at the root diagram, Figure 13.2c, we see that if we choose $\alpha = \alpha_1$ to have length 1 then $\alpha_2 = \beta$ has length $\sqrt{2}$, and the angle between α and β is $3\pi/4$, and $\cos(3\pi/4) = -\sqrt{2}/2$. Thus,

$$a_{12} = \frac{2(\alpha_1, \alpha_2)}{(\alpha_1, \alpha_1)} = \frac{2\|\alpha_1\|\|\alpha_2\|\cos(3\pi/4)}{\|\alpha_1\|^2} = \frac{2 \cdot 1 \cdot \sqrt{2} \cdot (-\sqrt{2}/2)}{1} = -2,$$
$$a_{21} = \frac{2(\alpha_2, \alpha_1)}{(\alpha_2, \alpha_2)} = \frac{2\|\alpha_2\|\|\alpha_1\|\cos(3\pi/4)}{\|\alpha_2\|^2} = \frac{2 \cdot \sqrt{2} \cdot 1 \cdot (-\sqrt{2}/2)}{(\sqrt{2})^2} = -1.$$

So, the Cartan matrix of B_3 is

$$A = \begin{pmatrix} 2 & -2 \\ -1 & 2 \end{pmatrix}. \tag{13.5.27}$$

Note that this is symmetrisable:

$$D = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \implies DA = \begin{pmatrix} 2 & -2 \\ -2 & 4 \end{pmatrix}. \qquad (13.5.28)$$

The corresponding Lie algebra is generated by $\{e_1, e_2, h_1, h_2, f_1, f_2\}$, subject to the relations that

- $[h_1, e_1] = 2e_1$, $[h_2, e_2] = 2e_2$, $[h_1, e_2] = -2e_2$, $[h_2, e_1] = -e_1$;
- $[h_1, f_1] = -2f_1$, $[h_2, f_2] = -2f_2$, $[h_1, f_2] = 2f_2$, $[h_2, f_1] = f_1$;
- $[e_1, f_1] = h_1$, $[e_2, f_2] = h_2$, $[e_1, f_2] = [e_2, f_1] = 0$;
- $[h_i, h_j] = 0$ for $i, j \in \{1, 2\}$;
- $(\mathrm{ad}_{e_1})^{1-a_{12}} e_2 = (\mathrm{ad}_{e_1})^3 e_2 = [e_1, [e_1, [e_1, e_2]]] = 0$, $[e_2, [e_2, e_1]] = 0$;
- $[f_1, [f_1, [f_1, f_2]]] = [f_2, [f_2, f_1]] = 0$.

This Lie algebra is isomorphic to that of $\mathfrak{so}_5$.

Notice that in all of these examples, and more generally by inspecting the relations defining $\mathfrak{g}$, we always have that $\{e_i, h_i, f_i\}$ (for fixed i) generates a copy of $\mathfrak{sl}_2$. These copies of $\mathfrak{sl}_2$ are such that the e_is and f_js of distinct copies don't "interact" (i.e., they commute). The interaction only occurs when h_is are involved. The h_is themselves form a subalgebra, which is exactly the Cartan subalgebra, which we can see from these relations is always abelian.

13.5.3 Classification of Cartan Matrices

The final part to classifying all finite-dimensional semisimple Lie algebras over $\mathbb{C}$ is to classify all finite-type Cartan matrices. This has been done. The tidiest way to frame this classification is to encode the information of a root system into a labelled graph, and then it turns out that all of the corresponding graphs either fall into one of four families of graphs, or one of five exceptional cases.

First, given an $n \times n$ Cartan matrix, A, or the corresponding root system, (Φ, Π, Δ), we can construct a labelled graph as follows:

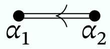
- The nodes are the simple roots, $\alpha_i \in \Delta$;
- Draw $a_{ij}a_{ji}$ edges between α_i and α_j ($i \neq j$);
- If α_i is longer than α_j draw an arrow on the edge pointing towards the shorter root.

The graph that we get is called the **Dynkin diagram** of the root system/Cartan matrix.

Example 13.5.29 Consider A_2, this has two simple roots, α_1 and α_2. We have $a_{12}a_{21} = (-1)(-11) = 1$, and so the corresponding Dynkin diagram is

(13.5.30)

Now consider B_2, this has two simple roots, α_1 and α_2. We have $a_{12}a_{21} = (-2)(-1) = 2$, and α_2 is longer than α_1, so the corresponding Dynkin diagram is

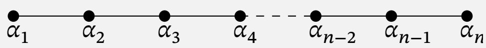


(13.5.31)

This process is invertible, since the Dynkin diagram fully encodes the angles between roots and their relative lengths (well, it encodes which is longer, the actual relative length can then be computed by requiring that the Cartan matrix have integral entries).

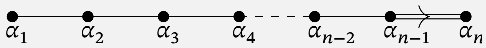
Theorem 13.5.32 — Classification of Root Systems. Every (finite-type) $n \times n$ Cartan matrix and its corresponding root system has a Dynkin diagram which is in one of the following infinite families (all with n vertices),

A_n



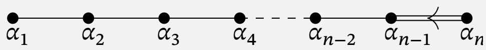
(13.5.33)

B_n



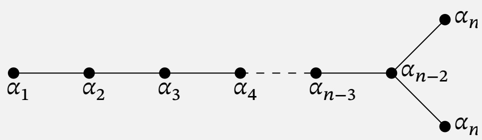
(13.5.34)

C_n



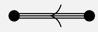
(13.5.35)

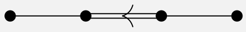
D_n

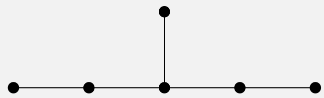


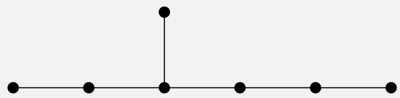
(13.5.36)

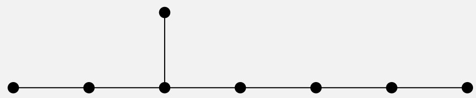
or is one of the following exceptional cases,

G_2  (13.5.37)

F_4  (13.5.38)

E_6  (13.5.39)

E_7  (13.5.40)

E_8  (13.5.41)

There is much more to be said about Dynkin diagrams and the things that they classify, but this is all we have time for here.

13.6 Verma Modules

We can use this classification to say something about the representation theory of semisimple Lie algebras over $\mathbb{C}$. To start with, when $\mathfrak{g}$ is defined from a root system in terms of the generators e_i, h_i, and f_i we can make the following definition.

Definition 13.6.1 — Verma Module Let $\mathfrak{g}$ be a semisimple Lie algebra over $\mathbb{C}$ with Cartan subalgebra $\mathfrak{h}$, and let $\lambda \in \mathfrak{h}^*$ be a weight. Let $I_\lambda \subseteq U(\mathfrak{g})$ be the left ideal generated by the elements $h - \lambda(h)1$ for $h \in \mathfrak{h}$ and e_i for $i = 1, \dots, r$. The **Verma module**, M_λ, is $U(\mathfrak{g})/I_\lambda$.

The idea of this definition is that M_λ is the largest (with respect to inclusion) highest weight representation with highest weight λ. Recall that by "highest weight representation" we mean that M_λ is generated (as a $U(\mathfrak{g})$-module) by some highest weight vector, v, which is such that $h \,.\, v = \lambda(h)v$ and $e_i \,.\, v = 0$. Thus, M_λ consists of linear combinations of elements of the form $f_{i_1} \cdots f_{i_k} \,.\, v$. The only relations imposed amongst these elements are those that are enforced by the commutation relations of the f_is. As a consequence f_i need not act nilpotently, and thus M_λ is infinite dimensional.

Let $\mathfrak{n}_+$ ($\mathfrak{n}_-$) denote the subalgebra of $\mathfrak{g}$ generated by the e_i (f_i). Then one can show that the Verma module, M_λ, is isomorphic to $U(\mathfrak{g}) \otimes_{U(\mathfrak{h}\oplus\mathfrak{n}_+)} \mathbb{C}_\lambda$ where $\mathbb{C}_\lambda$ is the one-dimensional representation of $\mathfrak{h} \oplus \mathfrak{n}_+$ in which $h \in \mathfrak{h}$ acts as $h \,.\, v = \lambda(h)v$ and $e \in \mathfrak{n}_+$ acts as $e \,.\, v = 0$ (define $\lambda_+ : \mathfrak{h} \oplus \mathfrak{n}_+ \to \mathbb{C}$ by $\lambda_+(h) = h$ and $\lambda_+(e) = 0$ and then this is the "obvious" one-dimensional representation). We can identify this construction as inducing $\mathbb{C}_\lambda$ up to all of $\mathfrak{g}$, so

$$M_\lambda \cong \operatorname{Ind}_{U(\mathfrak{h}\oplus\mathfrak{n}_+)}^{U(\mathfrak{g})} \mathbb{C}_\lambda. \tag{13.6.2}$$

This makes sense, the Verma module is such that $\mathfrak{h} \oplus \mathfrak{n}_+$ acts by highest weight, which is what $\mathbb{C}_\lambda$ captures, and then $\mathfrak{n}_-$ acts freely imposing only the required commutation relations, which is captured by inducing up to $\mathfrak{g} \cong \mathfrak{n}_- \oplus \mathfrak{h} \oplus \mathfrak{n}_+$.

The Verma module is infinite dimensional, but nevertheless it is still important in the theory of finite dimensional representations of $\mathfrak{g}$.

Proposition 13.6.3 Let $\mathfrak{g}$ be a semisimple Lie algebra with Cartan subalgebra $\mathfrak{h}$ and fix a weight $\lambda \in \mathfrak{h}^*$. Let $\mathbb{C}\langle f_1, \dots, f_n\rangle$ be the free algebra generated by the noncommuting symbols $f_1, \dots, f_n$, and let $\tilde{M}_\lambda = \mathbb{C}\langle f_1, \dots, f_n\rangle v$ be the free module generated by v. There exists an action of $\mathfrak{g}$ on $\tilde{M}_\lambda$ such that

$$f_i \cdot \left(\prod_k f_{j_k} v\right) = \left(f_i \prod_k f_{jk}\right); \tag{13.6.4}$$

$$h_i \cdot \left(\prod_k f_{j_k} v\right) = \left(\lambda(h_i) - \sum_k a_{i,j_k}\right)\left(\prod_k f_{j_k} v\right); \tag{13.6.5}$$

$$e_i \cdot \left(\prod_{k=1}^{l} f_{j_k} v\right) = \sum_{k \mid j_k = i} f_{j_1} \cdots f_{j_{k-1}} h_i f_{j_{k+1}} \cdots f_{j_l} v. \tag{13.6.6}$$

Proof. The $\mathfrak{g}$-module defined here is simply the Verma module, the only difference is that we're not imposing any condition on the f_is in the monomials in $\tilde{M}_\lambda$, whereas in M_λ we impose the Serre relations. □

The **weight lattice** of $\mathfrak{g}$ is $P = \mathbb{Z}\Phi \subset E$, the lattice generated by the roots. For example, for A_1 the weight lattice is just $\mathbb{Z}$, for $\mathrm{A}_1 \oplus \mathrm{A}_1$ it's $\mathbb{Z}^2$, for A_2 it's a hexagonal lattice, and for B_2 it's again a square lattice, $\mathbb{Z}^2$ (but scaled differently to $\mathrm{A}_1 \oplus \mathrm{A}_1$).

Corollary 13.6.7 The Verma module, M_λ, has a weight decomposition. In this weight decomposition the weight lattice is $P = \lambda - \mathbb{Z}\Phi$, and the λ-weight eigenspace of M_λ is one-dimensional, further, all weight subspaces are finite dimensional.

We are now ready to give the result which links M_λ to the finite-dimensional representations.

Proposition 13.6.8 — Universal Property of Verma Modules Let $\mathfrak{g}$ be a semisimple Lie algebra and use notation as above. If V is a $\mathfrak{g}$-module and $v \in V$ is a highest weight vector ($h.v = \lambda(h)v$ for $h \in \mathfrak{h}$ and $e_i.v = 0$) then there exists a unique homomorphism $\varphi : M_\lambda \to V$ such that $\eta(v_\lambda) = v$ where $v_\lambda \in M_\lambda$ is the highest weight element of the Verma module M_λ. In particular, if such a nonzero v generates V, that is V is a highest weight representation with weight vector v, then V is a quotient of M_λ.

The above result says that M_λ is universal amongst highest weight representations of $\mathfrak{g}$. Any map into any highest weight representation, V, can be achieved by first mapping into M_λ, then mapping into V in a unique way (using φ).

Proposition 13.6.9 Every highest weight representation has a weight decomposition into finite-dimensional weight subspaces.

So, every highest weight module is a quotient of the Verma module. It turns out that only one of these quotients is irreducible.

Proposition 13.6.10 For every $\lambda \in \mathfrak{h}^*$ the Verma module, M_λ, has a unique simple quotient, L_λ. Further, L_λ arises as a quotient of any highest weight $\mathfrak{g}$-module with highest weight λ.

The idea of the above is that as long as we never include v_λ in any submodule of M_λ we never get all of M_λ, and so we can sum all proper submodules of M_λ, and we know that the result will still be a proper submodule. We can then quotient by this sum, and the result is L_λ, we've quotiented out all submodules which could appear, and thus L_λ is simple.

Corollary 13.6.11 Simple highest weight $\mathfrak{g}$-modules (for $\mathfrak{g}$ a semisimple Lie algebra over $\mathbb{C}$) are classified by their highest weight, $\lambda \in \mathfrak{h}^*$, by the bijection $\lambda \mapsto L_\lambda$.

Fourteen

Braids, Knots, and Hecke Algebras

14.1 The Pure Braid Group

We start with a technical definition, assuming the reader is familiar with the notion of a braid group, if not maybe skip the definition and look at the pictures.

> **Definition 14.1.1 — Pure Braid Group** Let $M_n = \{(z_1, \dots, z_n) \in \mathbb{C}^n \mid z_i \neq z_j \text{ for } i \neq j\}$, which is a topological space as a subspace of $\mathbb{C}^n$. The **pure braid group**, is the fundamental group, $\mathcal{PB}_n = \pi_1(M_n)$.

A pure braid is then a (homotopy class) continuous function $\beta\colon [0,1] \to M_n$ with $\beta(0) = \beta(1)$, given by $t \mapsto (\beta_1(t), \dots, \beta_n(t))$ where the β_i are continuous functions $[0,1] \to \mathbb{C} \setminus \{z_1, \dots, \widehat{z}_i, \dots, z_n\}$ such that at no $t \in [0,1]$ do we have $\beta_i(t) = \beta_j(t)$.

Let $\mathbb{C}_n$ be the n-punctured complex plane[1]. Then for a pure braid, β, fixing some $t \in [0,1]$ we can view $\beta(t)$ as a choice of n distinct points in $\mathbb{C}$. Further, as t varies these points move around continuously. We can draw the whole path by considering $t \in [0,1]$ as a third dimension, and considering the positions traced by these points as time goes from 0 to 1. By lining up the punctures we can then project this down onto two dimensions, but keeping track of when a path goes over or under another. This gives us the standard picture of a pure braid. For example, Figure 14.1 shows an element of $\mathcal{PB}_4$.

The group operation of $\pi_1(M_n)$ is path concatenation (with rescaling of time so that we still have $t \in [0,1]$). The corresponding operation for pure braids is given by taking $\beta\beta'$ to be given by concatenating the diagram for β below the diagram for β' (reading the braid from the top down we want to do β' first[2]).

[1] We're treating $\mathbb{C}$ as a topological space here, the position of the points doesn't matter, we won't use any algebraic properties of this copy of $\mathbb{C}$

[2] The alternative convention gives us a perfectly well defined group, but to match conventions with the symmetric group we want this order.

14.2 The Braid Group

So far we've restricted our pure braids so that if a strand ends at the same puncture it begins at. The braid group relaxes this condition.

Let M_n be as in Definition 14.1.1. There is an obvious action of S_n on M_n given by permuting elements within a tuple, and this defines an equivalence relation on M_n, in which two tuples are equivalent if they are related by permuting elements. Let M_n/S_n be the quotient of M_n by this equivalence relation.

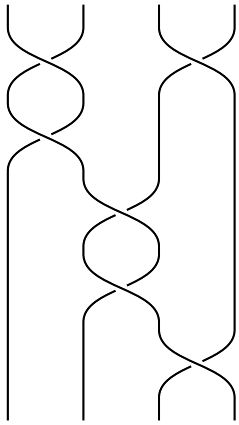


Figure 14.1: An element of the pure braid group, $\mathcal{PB}_4$.

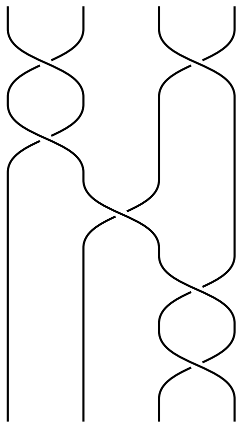


Figure 14.2: An element of the braid group, $\mathcal{B}_n$. Notice that the strands starting at 1, 2, 3 and 4 end at 1, 3, 4, and 2 respectively, defining a permutation (2 3 4).

Definition 14.2.1 — Braid Group The **braid group** is $\mathcal{B}_n = \pi_1(M_n/S_n)$.

In terms of the pictures of braids the only difference is that we no longer require that braids start and end at the same point. See Figure 14.2. The group operation is still concatenation. Notice that by tracking where each strand starts and ends we get a permutation, $w \in S_n$. It is always possible to write a braid, $b\colon [0,1] \to M_n/S_n$, as a composite, $b = \beta \circ p$, where β is a pure braid and $w \in S_n$ is a permutation such that $\beta(1) = w^{-1}(b(0))$.

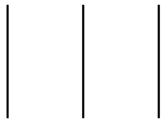
Theorem 14.2.2 — Artin. The braid group has the standard presentation

$$\mathcal{B}_n = \langle \sigma_1, \dots, \sigma_{n-1} \mid \sigma_i\sigma_{i+1}\sigma_i = \sigma_{i+1}\sigma_i\sigma_{i+1}, \sigma_i\sigma_j = \sigma_j\sigma_i \text{ for } |i-j| > 1 \rangle.$$

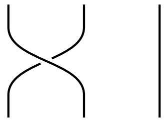
The relationship

$$\sigma_i\sigma_{i+1}\sigma_i = \sigma_{i+1}\sigma_i\sigma_{i+1} \tag{14.2.3}$$

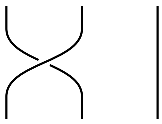
is called the **braid relation**. The identification between this presentation and $\mathcal{B}_n$ is pretty simple. For simplicity we'll just look at the $n = 3$ case, but for other values of n the pictures generalise in the obvious way. First, the identity, e, is simply leaving all strands fixed:

$e =$  . (14.2.4)

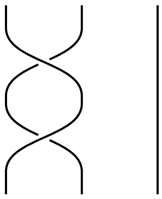
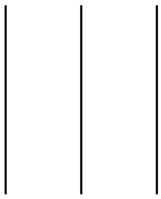
Then σ_1 is the braid

$\sigma_1 =$  , (14.2.5)

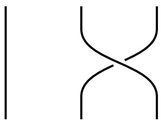
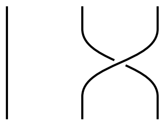
and σ_1^{-1} is given by crossing in the other direction:

$\sigma_1^{-1} =$  . (14.2.6)

This makes sense, since we then have

$\sigma_1\sigma_1^{-1} =$  $=$  $= e.$ (14.2.7)

Similarly, we have

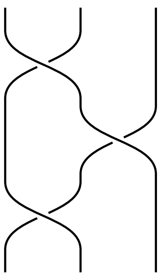
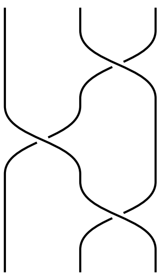
$\sigma_2 =$  and $\sigma_2^{-1} =$  . (14.2.8)

So, σ_i means passing the ith strand over strand $i+1$, and σ_i^{-1} means passing the ith strand *under* strand $i+1$.

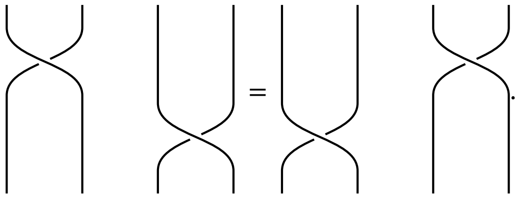
The braid relation in this case tells us that

$$\sigma_1\sigma_2\sigma_1 = \sigma_2\sigma_1\sigma_2, \tag{14.2.9}$$

which is just the following picture:

 $=$  . (14.2.10)

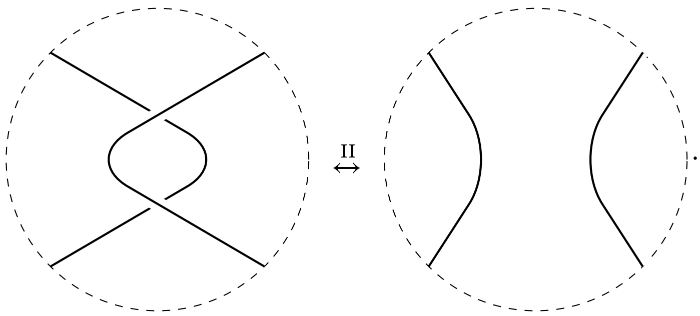
For $n = 3$ we never have $|i - j| > 1$, so let's look at $n = 4$, where this relation simply tells us that "sufficiently separated" swaps commute:



(14.2.11)

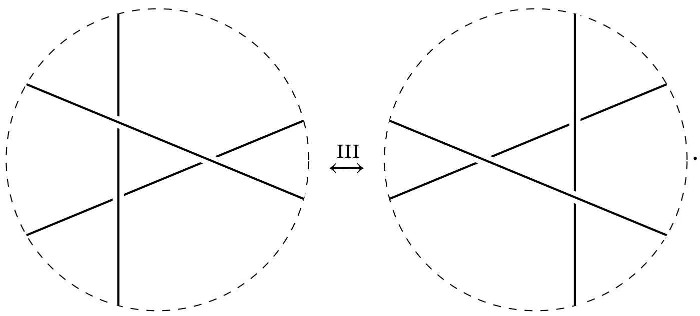
So far, we've just been looking at braids and deciding if they're equal if they intuitively give the same picture after rearranging strands without passing them through each other. This can be made rigorous as follows.

First, we define the **Reidemeister moves** of types II and III. These are "local" operations on braids, in that we can apply them to any portion of the diagram without changing the rest of the diagram. To represent this we use a dashed circle to "zoom in" on just a portion of the diagram: The Reidemeister move of type II is



(14.2.12)

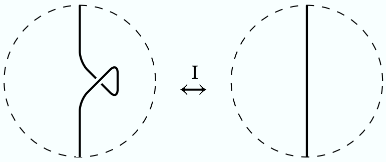
The Reidemeister move of type III is



(14.2.13)

These are simply capturing the fact that we want $\sigma_i^{-1}\sigma_i = e$ and the braid relation.

Remark 14.2.14 The Reidemeister moves first arose in knot theory, in which there is a third Reidemenster move, move number I, which is

 (14.2.15)

We don't consider this as strands in braids aren't allowed to loop back up.

Proposition 14.2.16 Two braids are the same if and only if they are related by an isotopy and a sequence of Reidemeister moves of type II and III.

Remark 14.2.17 In physics a braid describes the adiabatic exchange of indistinguishable quasi particles in two dimensions. This is important in, for example, the fractional quantum Hall effect. This idea has applications to quantum computing. Particles whose exchange is governed by the braid group are called *anyons* (cf. bosons and whose exchange is governed by the trivial and antisymmetric representations of the symmetric group).

Remark 14.2.18 From a geometric view point $\mathcal{B}_n$ is the "mapping class group" of the punctured disc with n points. We swap the plane to a disc just because it's nicer to work with compact things, and we're not allowing punctures at infinity anyway.
The mapping class group is defined as follows. Let S be a surface, and $Q \subset S$ a finite set of marked points. Denote by $\mathrm{Homeo}(S, Q)$ the group of homeomorhpisms of S which fix Q as a set and fix the boundary pointwise. That is, $\varphi \in \mathrm{Homeo}(S, Q)$ is such that for every marked point, p, there is some marked point q (not necessarily distinct) such that $\varphi(p) = q$, and for every boundary point, x we have $\varphi(x) = x$. The **mapping class group** of the marked surface (S, Q) is

$$\mathrm{Mod}(S, Q) = \mathrm{Homeo}^+(S, Q)/\,\mathrm{Homeo}_0(S, Q) \tag{14.2.19}$$

where $\mathrm{Homeo}^+(S, Q)$ denotes the collection of orientation preserving homeomorphisms in $\mathrm{Homeo}(S, Q)$, and $\mathrm{Homeo}_0(S, Q)$ denotes the connected component of $\mathrm{Homeo}(S, Q)$ containing the identity (in the compact-open topology).
This group is also sometimes called the modular group, hence the notation $\mathrm{Mod}(S, Q)$. This is because when we take the torus with no marked points the mapping class group ends up being isomorphic to the modular group, $\mathrm{SL}_2(\mathbb{Z})$.
When we say that $\mathcal{B}_n$ is the mapping class group of the disc with n punctured points we mean that if D is this punctured disc and Q is our set of punctures then there is an isomorphism $\mathcal{B}_n \to \mathrm{Mod}(D, Q)$ given by $\sigma_i \mapsto H_i$ where σ_i is a generator of the braid group in the standard presen-

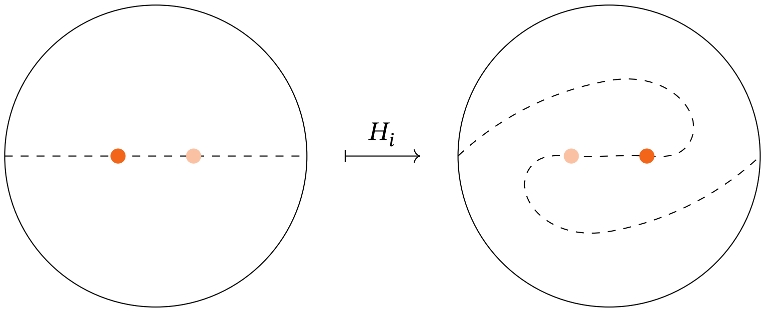


Figure 14.3: Half twist of the disc exchanging i and $i+1$. The dashed line shows some curve and its image under H_i.

tation and H_i is the homeomorphism of the n-punctured disc given by a half twist exchanging the points numbered i and $i+1$. See Figure 14.3.

14.3 Coxeter Groups

Definition 14.3.1 — Coxeter System Let S be a finite set. A **Coxeter matrix** for S is a symmetric matrix $M = (m_{st})_{s,t\in S}$ such that

1. $m_{ss} = 1$ for all $s \in S$;
2. $m_{st} \in \{2, 3, \dots\} \cup \{\infty\}$ for all distinct $s, t \in S$.

We call (S, M) a **Coxeter system**.

For example, take $S = \{1, \dots, 4\}$ and

$$M = \begin{pmatrix} 1 & 3 & 2 & 2 \\ 3 & 1 & 3 & 2 \\ 2 & 3 & 1 & 3 \\ 2 & 2 & 3 & 1 \end{pmatrix}. \tag{14.3.2}$$

Definition 14.3.3 — Coxeter Group Given a Coxeter system, (S, M), the corresponding **Coxeter group**, is the pair (W, S), where W is the group given by the presentation

$$W := \langle s \in S \mid (st)^{m_{st}} = 1 \forall s, t \in S \rangle. \tag{14.3.4}$$

The **rank** of W is defined to be $|S|$.

Notice that if we take $s = t$ then we have $m_{ss} = 1$, so $(ss)^{m_{ss}} = s^2$, and so in a Coxeter group we always have $s^2 = 1$, and hence $s^{-1} = s$ for all generators. If $m_{st} = 2$ then $(st)^2 = 1$, which means that $st = ts$, since $s^{-1} = s$ and $t^{-1} = t$. Thus, a 2 in the matrix means that the corresponding generators commute. More generally, we can always take the equation $(st)^{m_{st}} = stst \cdots st = 1$, which has $2m_{st}$ factors, and rearrange to get

$$\underbrace{sts\cdots}_{m_{st}\text{ terms}} = \underbrace{tst\cdots}_{m_{st}\text{ terms}}. \tag{14.3.5}$$

For example, when $m_{st} = 3$ this relation is

$$sts = tst. \tag{14.3.6}$$

We call this the **braid relation**.

Consider the Coxeter system $(\{s_1, s_2, s_3, s_4\}, M)$ with M given by Equation (14.3.2). The corresponding Coxeter group is

$$W = \langle s_1, s_2, s_3, s_4 \mid s_i^2 = 1, s_i s_{i+1} s_i = s_{i+1} s_i s_{i+1}, s_i s_j = s_j s_i \text{ for } |i-j| > 1\rangle. \tag{14.3.7}$$

We can recognise this as a presentation of S_5, where $s_i = (i\, i{+}1)$. This presentation generalises fully to $S = \{s_1, \dots, s_{n-1}\}$ to give the Coxeter presentation of S_n.

Definition 14.3.8 Let (W, S) be a Coxeter group. Given $w \in W$ we can write w as $s_{i_1} \cdots s_{i_k}$, with $s_i \in S$. We say that this is a **reduced expression** for w if k is minimal, and we define $\ell(w) = k$ to be the **length** of w.

Theorem 14.3.9 — Matsumoto. Let W be a group generated by $S = \{s_1, \dots, s_n\}$ subject to some relations, such that $s_i^2 = 1$. Then the following are equivalent:

- (W, S) is a Coxeter group.
- Any two reduced expressions for $w \in W$ can be transformed into each other by a series of braid relations.

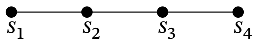
Note that two Coxeter groups may be isomorphic as groups, but not as Coxeter groups, since if they have different generating sets and/or relations then we consider them to be different as Coxeter groups, but not as groups. The problem of producing a general algorithm to decide if two Coxeter systems produce isomorphic groups is unsolved. A related open problem is deciding, given W, what subset, S, and relations can we take to make (W, S) a Coxeter group.

14.3.1 Classification of Coxeter Groups

Definition 14.3.10 — Coxeter Diagram Let (S, M) be a Coxeter system. The corresponding **Coxeter diagram** is constructed as follows:

- The vertices are elements of S;
- If $m_{st} < 3$ there are no edges between s and t;
- If $m_{st} = 3$ there is an unlabelled edge between s and t;
- If $m_{st} \geq 4$ there is an edge between s and t labelled with m_{st}.

For example, if we take M as in Equation (14.3.2) then the corresponding Coxeter graph has as vertices $\{s_1, s_2, s_3, s_4\}$, and there are edges connecting s_1 to s_2, s_2 to s_3, and s_3 to s_4, so the graph is

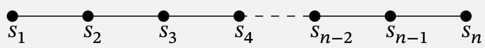


(14.3.11)

Remark 14.3.12 Notice that this is exactly the Dynkin diagram A_4. Dynkin diagrams and Coxeter diagrams are very closely related. There are some distinctions though: Dynkin diagrams can be directed, Coxeter graphs can be labelled, Dynkin diagrams can have multiple edges between a pair of vertices.
Both Dynkin and Coxeter diagrams encode a matrix, Cartan and Coxeter respectively, which encode some properties of an algebraic structure, a Lie algebra and group respectively.

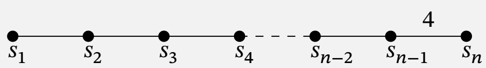
Theorem 14.3.13 — Classification of Coxeter Groups. Every Coxeter system has a Coxeter diagram falling into one of the following infinite families,

A_n



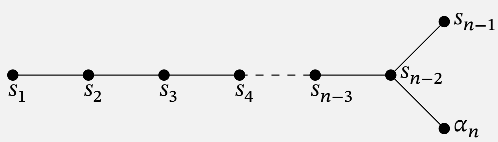
(14.3.14)

$B_n = C_n$



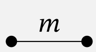
(14.3.15)

D_n



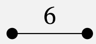
(14.3.16)

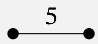
$I_2(m)$

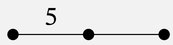


(14.3.17)

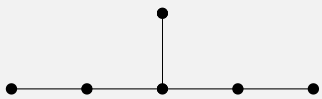
or is one of the following exceptional cases,

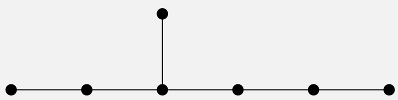
G_2  (14.3.18)

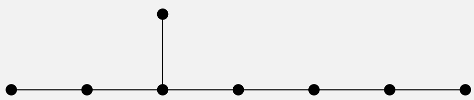
H_2  (14.3.19)

H_3  (14.3.20)

F_4  (14.3.21)

E_6  (14.3.22)

E_7  (14.3.23)

E_8  (14.3.24)

Note that $G_2 = I_2(6)$ and $H_2 = I_2(5)$. It is also common to see $I_2(m)$ written as I_m, but this is bad notation as $I_2(m)$ is rank 2, not m. We have both B_n and C_n because in generalisations (such as affine Coxeter groups) these become different. Sometimes we write BC_n for these finite-type diagrams which are the same.

For many of these the corresponding Coxeter group has quite a nice interpretation:

- A_n is the symmetric group S_{n+1}, it can also be thought of as the symmetries of the n-simplex.
- B_n is the symmetries of the n-cube, and can be thought of as signed permutations of $\{-n, \dots, n\}$, this group is $S_n \wr \mathbb{Z}/2\mathbb{Z} = S_n \ltimes (\mathbb{Z}/2\mathbb{Z})^n$.
- D_n is a subgroup of the symmetries of the n-cube consisting only of signed permutations which fix the number of minus signs.
- $I_2(m)$ is the symmetries of the regular m-gon (that is, the dihedral group of order $2m$).
- A_3, B_3, and H_3 are symmetries of the platonic solids (including reflections), being the tetrahedron, cube/octahedron, and dodecahedron/icosahedron symmetry groups respectively.

14.3.2 Reflection Groups

A **reflection** of a Euclidean space, V, with inner product $\langle -, - \rangle$, is a linear transformation, $s : V \to V$ such that there exists some $\alpha \in V$ with $s(\alpha) = -\alpha$ and s fixes the hyperplane perpendicular to α, $H_\alpha = (\mathbb{R}\alpha)^\perp$, pointwise. A **reflection group** is a group which is (isomorphic to) a subgroup of $O(V)$ generated only by reflections. It turns out that every Coxeter group can be viewed as a reflection group.

Let V be a finite-dimensional Euclidean space, and $\mathcal{H}$ a finite collection of hyperplanes. Removing these hyperplanes from V we get $V \setminus \bigcup_{H\in\mathcal{H}} H$. We call the connected components of this space **alcoves**. For example, if we take $V = \mathbb{R}^2$, the plane, then a hyperplane is just a line. If we take two non-parallel lines in the plane then they split the plane into four segments, which are our alcoves.

Fix some alcove, A, in V, then define

$$S_A = \{s_H \mid s_H \text{ is a reflection in } H \text{ with } H \in \mathcal{H} \text{ bounding } A\}. \tag{14.3.25}$$

Then if we take W to be the group generated by such reflections then (W, S_A) is a Coxeter group.

Conversely, if we're given a Coxeter system, (W, S), then we can define a Euclidean space,

$$V = \mathbb{R}S \cong \bigoplus_{s\in S} \mathbb{R}\boldsymbol{s} \tag{14.3.26}$$

where we're defining a basis vector, $\boldsymbol{s}$, for each $s \in S$. More abstractly, V is the free vector space on S. The inner product on this space is given on this basis by

$$\langle \boldsymbol{s}, \boldsymbol{t} \rangle = \cos\left(\frac{\pi}{m_{st}}\right). \tag{14.3.27}$$

This is always positive because $\pi/m_{st} \in [0, \pi]$ (defining $\pi/\infty = 0$), and it's symmetric because M is a symmetric matrix. We can then define a reflection in the hyperplane orthogonal to $\boldsymbol{s}$ by

$$\sigma_{\boldsymbol{s}}(\boldsymbol{v}) = \boldsymbol{v} - 2\langle \boldsymbol{v}, \boldsymbol{s} \rangle \boldsymbol{s} \tag{14.3.28}$$

for all $\boldsymbol{v} \in V$.

Proposition 14.3.29 — Tits Representation With the notation above the map $W \to \mathrm{GL}(V)$ defined by $s \mapsto \sigma_{\boldsymbol{s}}$ is a faithful representation, and $\langle -, - \rangle$ is W-invariant.

14.3.3 Generalisations

If we remove the requirement that $s_i^2 = 1$ (equivalently allowing the diagonals to not be 1), then the group that we get is called an **Artin group**. Matsumoto's theorem then states that two words in a Coxeter group are related only by the relations of the corresponding Artin group. Examples of Artin groups include all Coxeter groups, as well as the braid groups. Some other examples are $\langle S \mid st = ts \forall s, t \in S \rangle$, which is the free abelian group on S, and $\langle S \rangle$, the free group on S. This suggests the following generalisation of the braid group.

Definition 14.3.30 — Artin Braid Group The **Artin braid group** of a Coxeter group, (W, S), is

$$\mathcal{B}_W = \langle \{b_s \mid s \in S\} \mid \underbrace{b_s b_t b_s \cdots}_{m_{st}} = \underbrace{b_t b_s b_t \cdots}_{m_{st}} \rangle. \tag{14.3.31}$$

Note that the braid group $\mathcal{B}_n$ as defined in Definition 14.2.1 is simply $\mathcal{B}_{S_n}$, as can be seen by the standard presentation (Theorem 14.2.2) which mirrors the relations of the S_n presentation (Equation (14.3.7)) after removing the condition that $s_i^2 = 1$.

Proposition 14.3.32 — Burau Let $\Lambda = \mathbb{Z}[t, t^{-1}]$. There is a group homomorphism

$$\rho_n : \mathcal{B}_n \to \mathrm{GL}_n(\Lambda) \tag{14.3.33}$$

given by

$$\rho_n(b_i) = \begin{pmatrix} 1 & & & & & & & \\ & \ddots & & & & & & \\ & & 1 & & & & & \\ & & & 1-t & t & & & \\ & & & 1 & 0 & & & \\ & & & & & 1 & & \\ & & & & & & \ddots & \\ & & & & & & & 1 \end{pmatrix} \tag{14.3.34}$$

where $1 - t$ appears in position (i, i). This representation of $\mathcal{B}_n$ is known as the **Burau representation**.

For example, if $n = 3$ then we have

$$\rho_3(b_1) = B_1 = \begin{pmatrix} 1-t & t & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \text{and} \quad \rho_3(b_2) = B_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1-t & t \\ 0 & 1 & 0 \end{pmatrix}. \tag{14.3.35}$$

One can then check that these two matrices satisfy $B_1 B_2 B_1 = B_2 B_1 B_2$.

Remark 14.3.36 For $n \leq 3$ the Burau representation is faithful (as the above shows in the $n = 3$ case, the $n = 2$ case is trivially faithful). For $n \geq 5$ the Burau representation is not faithful. For $n = 4$ faithfulness is unknown, but it is known that ρ_4 is a faithful representation if the Jones polynomial detects the unknot, an open question. See Section 14.4.

An interpretation of the Burau representation, due to Jones, is as follows. Picture a braid, $b \in \mathcal{B}_n$ as a bowling alley where each strand becomes a lane. Throw a bowling ball down one lane. Upon crossing over a lane below have the ball fall down to the lower lane with probability t (so it remains in the upper lane with probability $1-t$). Then the entry in position (i, j) of $\rho_n(b)$ is the probability that a ball thrown initially down lane i ends up in lane j.

For $n > 2$ the Burau representation is not irreducible, it admits a one-dimensional invariant subspace. Taking the quotient of Λ^n by this invariant subspace we we get the reduced Burau representation, which is irreducible.

Proposition 14.3.37 Let $\Lambda = \mathbb{Z}[t, t^{-1}]$. For n and integer greater than 2 there is a group homomorphism

$$\tilde{\rho}_n : \mathcal{B}_n \to \mathrm{GL}_{n-1}(\Lambda) \tag{14.3.38}$$

given by

$$\tilde{\rho}_n(b_1) = \begin{pmatrix} -t & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & I_{n-3} \end{pmatrix}, \tag{14.3.39}$$

$$\tilde{\rho}_n(b_i) = \begin{pmatrix} I_{i-2} & 0 & 0 & 0 & 0 \\ 0 & 1 & t & 0 & 0 \\ 0 & 0 & -t & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & I_{n-i-2} \end{pmatrix} \tag{14.3.40}$$

$$\tilde{\rho}_n(b_{n-1}) = \begin{pmatrix} I_{n-3} & 0 & 0 \\ 0 & 1 & t \\ 0 & 0 & -t \end{pmatrix} \tag{14.3.41}$$

where $2 \leq i \leq n-2$. This representation of $\mathcal{B}_n$ is called the **reduced Burau representation**.

Lemma 14.3.42 Let C be the $n \times n$ matrix

$$C = \begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 0 & 1 & 1 & \dots & 1 \\ 0 & 0 & 1 & \dots & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \end{pmatrix}. \tag{14.3.43}$$

Then

$$C^{-1}\rho_n(b_i)C = \begin{pmatrix} \tilde{\rho}_n(b_i) & 0 \\ X_i & 1 \end{pmatrix} \tag{14.3.44}$$

where X_i is the row of length $n-1$ which is $(0, \dots, 0)$ if $i \neq n-1$ and $(0, \dots, 0, 1)$ for $i = n-1$.

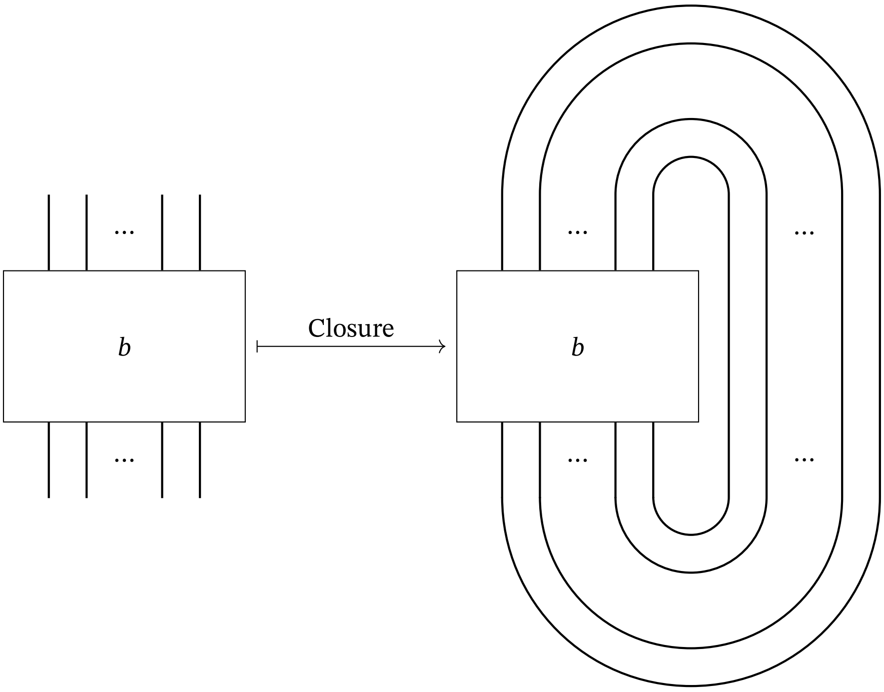


Figure 14.4: The closure of a braid, b, to produce a link.

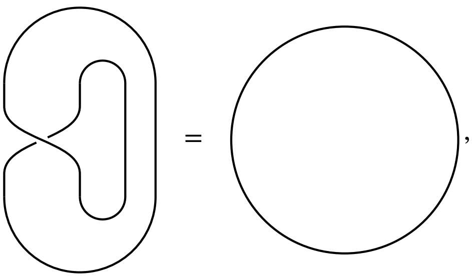
14.4 Knots

Definition 14.4.1 — Knot A **knot** is an embedding of S^1 in $\mathbb{R}^3$, considered up to ambient isotopy. A **link** is an embedding of $\bigsqcup_{k=1}^{n} S^1$, up to ambient isotopy.

We draw knots as their projection onto the plane. The simplest case is the unknot which *can*[3] be drawn with no crossings, in which case it just looks like what we would normally think of as a circle. A knot or link is oriented if we assign an orientation to each copy of S^1. Similarly, any braid can be oriented by declaring that all strands are oriented downwards.

Given a braid, b, its closure is given by joining the strand starting at position i to the strand ending at position i without introducing any new crossings. The result is a link. This is shown in Figure 14.4.

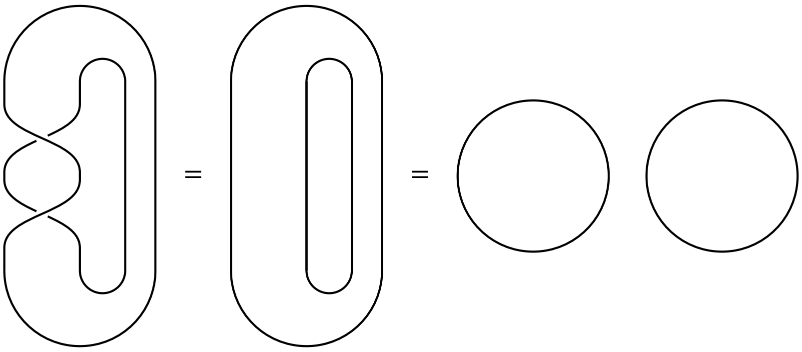
For example, in $\mathcal{B}_2$ the closure of σ_1 is the unknot,



(14.4.2)

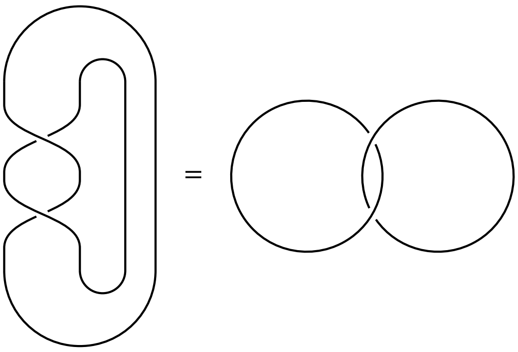
[3] It's possible also to draw it with crossings.

the closure of $\sigma_1\sigma_1^{-1} = e$ is the two component unlink,



(14.4.3)

and the closure of σ_1^2 is the **Hopf link**,



(14.4.4)

Theorem 14.4.5 — Alexander. Any oriented link can be obtained as the closure of an oriented braid.

Note that there will be many braids which close to give the same link. Thus, there generally multiple ways to take a link and cut it resulting in different braids with the same closure.

In order to consider all links at once it is not sufficient to consider $\mathcal{B}_n$ for some fixed number of strands, n. Instead we can define the **infinite braid group** to be the direct limit

$$\mathcal{B}_\infty := \varinjlim_n \mathcal{B}_n, \tag{14.4.6}$$

of the directed system given by the obvious inclusions $\iota\colon \mathcal{B}_n \to \mathcal{B}_{n+1}$, adding a strand on the right which doesn't cross any of the $n-1$ original strands. The resulting group can then be thought of as the braid group on an infinite number of strands. The way to reason about this is to only consider braids in which a finite number of strands are actually crossing, in which case we can just consider $\mathcal{B}_n$ with n chosen to be sufficiently large[4].

[4]cf. any element of the ring of symmetric functions can, despite having infinitely many variables, be considered as a polynomial in "sufficiently many" variables.

Definition 14.4.7 — Markov Moves The **Markov moves** are certain replacement rules in $\mathcal{B}_\infty$ which come in two types. Consider a and b to be elements of $\mathcal{B}_\infty$, and in particular, suppose that b has no crossings after string $n-1$, so can be viewed as an element of $\mathcal{B}_n$. Type I, conjugation, allows us to replace ab with ba. Type II comes in two subtypes, stabilisation, we can replace b with $\iota(b)b_n$ and destabilisation, we can replace b with $\iota(b)b_n^{-1}$ where the subscript n denotes that we start the braid b_n at strand n.

The ultimate goal of much of knot theory is to compute knot invariants which allow us to distinguish different knots. By "invariants" we mean objects (booleans, numbers, polynomials, vector spaces, etc.) which don't change with an ambient isotopy of the knot, so are the same however we draw the knot. By "distinguish" we mean that these values differ between different knots.

Some knot invariants successfully distinguish all knots, for example the complement of a knot (viewed as a topological space) is a "complete invariant". Unfortunately, these complete invariants tend to be hard to compute. Many more knot invariants fail to distinguish between many distinct knots. For example, whether a knot is 3-colourable[5] is an invariant, but only separates knots into one of two classes. This is still useful though, it distinguishes the unknot (not 3-colourable) from the trefoil (3-colourable).

[5] A knot is 3-colourable if each strand of the knot diagram (considering a strand passing under another to be a break in the strand) such that at least two colours are used and at each crossing all three strands are either the same colour or different colours

Perhaps the most celebrated knots invariants are polynomial invariants. The first of these discovered is the Alexander polynomial. Conway later showed that a slightly different form of this polynomial admits a description in terms of skein relations. Later more knot polynomials, such as the Jones and HOMFLYPT polynomials were found. These all have deep connections to representation theory. For example, the Jones polynomial can be understood through representations of the quantum group $U_q(\mathfrak{sl}_2)$.

Definition 14.4.8 — Alexander–Conway Polynomial Let L be an oriented link, and let $b \in \mathcal{B}_n$ be such that the closure of b is L. Let $t = q^2$ Then the **Alexander–Conway polynomial** of L is the element of $\mathbb{Z}[t^{\pm 1/2}] = \mathbb{Z}[q^{\pm}]$ given by

$$\nabla(L) := (-1)^{n+1} q^{\deg b} \frac{q - q^{-1}}{q^n - q^{-n}} \det(\tilde{\rho}_n(b) - I_{n-1}). \tag{14.4.9}$$

Note that

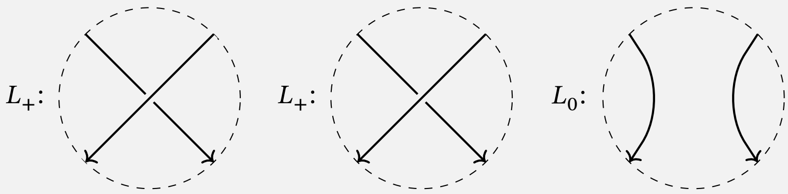
$$[n]_q = \frac{q^n - q^{-n}}{q - q^{-1}} \tag{14.4.10}$$

is the ***q*-analogue** n, so there's already some relation to quantum groups occurring here.

Theorem 14.4.11. The Alexander–Conway polynomial is uniquely determined on oriented links, L, by the following:

1. $\nabla(\text{unknot}) = 1$ for either choice of orientation;

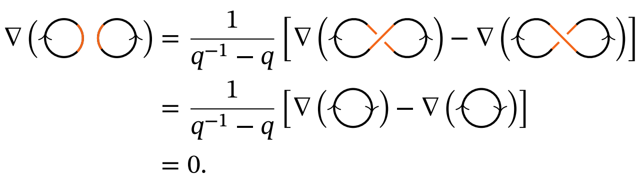
2. $\nabla(L_+) - \nabla(L_-) = (q^{-1} - q)\nabla(L_0)$, known as the **skein relation**. Here L_+, L_-, and L_0 are a Conway triple of links which differ only locally by the following:



(14.4.12)

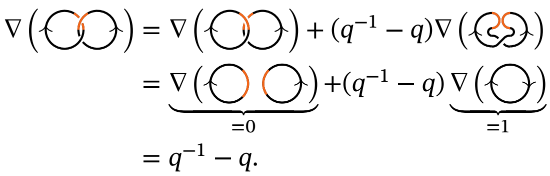
Outside of the circle these links will be the same.

The Skein relation gives a way to compute the Alexander–Conway polynomial. It will always be possible to use the skein relation to manipulate the polynomial into a combination of polynomials of links for which the Alexander–Conway polynomial is known. For example, starting with the two component unlink we can we can perform the following calculation:



(14.4.13)
(14.4.14)
(14.4.15)

We can then use this result to compute the Alexander–Conway polynomial of the Hopf link:



(14.4.16)
(14.4.17)
(14.4.18)

14.5 Iwahori–Hecke Algebra

In this section we define the Iwahori–Hecke algebra (often just called the Hecke algebra). Recall that a given root system, Φ, has a corresponding Euclidean space, $E = \mathrm{span}_{\mathbb{R}}\, \Phi$. The **Weyl group** of this root system, W, is the subgroup of $\mathrm{O}(E)$ generated by reflections in the hyperplanes orthogonal to the roots, that is, W is generated by the s_α. Weyl groups are Coxeter groups, but not every Coxeter group is a Weyl group. In fact, because root systems are classified by Dynkin diagrams so are Weyl groups, so there are Weyl groups of types A_n, B_n, C_n, D_n, E_6, E_7, E_8, G_2, and F_4.

It is possible to define the Iwahori–Hecke algebra for a general Weyl group, but we will only define it for the type A_n case.

Definition 14.5.1 The type A_n **Iwahori–Hecke algebra**, H_n, is the unital associative $\mathbb{Q}(q)$-algebra with generators $T_1, \dots, T_{n-1}$ and relations

1. $T_i T_{i+1} T_i = T_{i+1} T_i T_{i+1}$;
2. $T_i T_j = T_j T_i$ for $|i - j| \geq 2$;
3. $(T_i - q)(T_i + q^{-1}) = 0$.

Conventions vary quite a lot here, for example, it's common to swap q and q^{-1}, and other authors use $q^{1/2}$ instead of q.

Remark 14.5.2 We have chosen to make the definition here over the field $\mathbb{Q}(q)$. This is the field of rational functions[a] in q with coefficients in $\mathbb{Q}$. It is entirely possible to make the same definition with $\mathbb{C}(q)$ instead. It is also common to make this definition over $\mathbb{Q}$ (or $\mathbb{C}$) and just think of q as being a chosen value of $\mathbb{Q}$ (or $\mathbb{C}$). This has the advantage of being slightly less notation, but there are some subtleties that creep in, mostly about when certain denominators vanish. We can recover this version by simply picking a value of q at which to evaluate our rational functions.

[a] that is, $f \in \mathbb{Q}(q)$ is a ratio $f(q) = g(q)/h(q)$ with polynomials $g(q), h(q) \in \mathbb{Q}[q]$ and $h(q)$ not identically zero.

Notice that the first and second relations are shared by the braid group, so we can define H_n as the quotient

$$H_n = \mathbb{Q}(q)\mathcal{B}_n / \langle (T_i - q)(T_i + q^{-1}) \rangle \tag{14.5.3}$$

where $\mathbb{Q}(q)\mathcal{B}_n$ is the group algebra of $\mathcal{B}_n$ over the field, $\mathbb{Q}(q)$, of rational functions in q with coefficients in $\mathbb{Q}$.

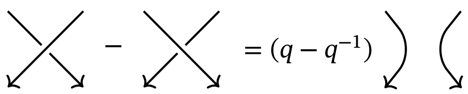
Rearranging the third relation we get

$$1 = T_i(T_i - q + q^{-1}). \tag{14.5.4}$$

This means that T_i^{-1} exists, and is equal to $T_i + q - q^{-1}$. Note that whenever we write q we're really thinking of it as $q1$ where 1 is the unit of H_n. We then have the relation

$$T_i - T_i^{-1} = (q - q^{-1})1. \tag{14.5.5}$$

This can be understood as being the skein relation



(14.5.6)

using the obvious notation for the generators T_i and identity inherited from $\mathcal{B}_n$ in the quotient.

Let $H_n^\times$ be the group of units of H_n. Note that $H_n^\times$ is more complicated than $H_n \setminus 0$, since we also have to account for inverses of sums of generators, as well as products of generators (any product of generators having an inverse as each generator has a multiplicative inverse). The map $\mathcal{B}_n \to H_n^\times$, sending generators to generators, $\sigma_i \mapsto T_i$, is injective exactly when the Burau representation is, so it is for $n = 3$, not for $n \geq 5$, and it's an open problem for $n = 4$.

The idea motivating the definition of H_n is that if we "set $q = 1$" we recover[6] S_n. More formally, we have

[6] Here S_n is playing the role of the Weyl group of type A_n.

$$H_n/\langle q-1\rangle \cong \mathbb{Q}S_n \tag{14.5.7}$$

with the isomorphism simply being $T_i \mapsto s_i$. Thus, we can interpret H_n as a ***q*-deformation** of S_n (or rather its group algebra).

Much of the representation theory of S_n lifts to H_n. The Hecke algebra also has the advantage that certain results in representation theory actually end up being easier to view in H_n, and then remain true after setting $q = 1$.

This idea is very similar to that of the quantum group, $U_q(\mathfrak{sl}_n)$, and indeed the Schur–Weyl duality of the commuting actions $\mathfrak{sl}_n \curvearrowright V^{\otimes r} \curvearrowleft S_r$ lifts to a quantum version, $U_q(\mathfrak{sl}_n) \curvearrowright V^{\otimes r} \curvearrowleft H_r$.

Proposition 14.5.8 Let $w = s_{i_1} \dots s_{i_r}$ be a reduced expression for $w \in S_n$, and let $T_w = T_{i_1} \cdots T_{i_r}$. Then $\{T_w \mid w \in S_n\}$ is a basis of H_n.

We won't prove this result, but the key step is to show that for any two reduced expressions of w the resulting T_w is the same. The proof of this part uses the skein relation to turn one expression into the other.

14.5.1 Representations of Iwahori–Hecke Algebras in Type A

14.5.1.1 Irreducible Representations

For λ a partition of n let $S_\lambda = S_{\lambda_1} \times \cdots \times S_{\lambda_k}$ be the corresponding row Young subgroup. Inspired by the representation theory of the symmetric group we define an analogue of the symmetriser,

$$a_\lambda = \sum_{w \in S_n} q^{\ell(w)} T_w. \tag{14.5.9}$$

Then we can define a left H_n-module, $M_\lambda \coloneqq H_n a_\lambda$. For the case of the S_n we had the decomposition

$$M_\lambda = V_\lambda \oplus \bigoplus_{\mu > \lambda} K_{\mu\lambda} V_\mu \tag{14.5.10}$$

where the V_μ are the irreducible Specht modules and $K_{\mu\lambda}$ are the Kostka numbers. We take this as inspiration to define the Specht modules for the Hecke algebra. Let I_λ be the two-sided ideal generated by a_μ for $\mu > \lambda$. Then the Hecke-algebra Specht module is

$$V_\lambda \coloneqq M_\lambda/(M_\lambda \cap I_\lambda). \tag{14.5.11}$$

The construction is such that the quotient by I_λ kills the $\bigoplus K_{\mu\lambda} V_\mu$ part of the sum in the $q = 1$ case leaving us with the symmetric-group Specht modules. Note that if $\lambda \neq \mu$ then $V_\lambda = V_\mu$ only if these module are trivial.

Theorem 14.5.12. Every simple H_n-module is of the form

$$\tilde{V}_\lambda = V_\lambda / \operatorname{Rad} V_\lambda \tag{14.5.13}$$

for some λ a partition of n.

The above theorem says that every irreducible representation of the Hecke algebra is the unique simple quotient of some Specht module. We can characterise the nontrivial simple H_n-modules in the following by placing a condition on how fast the rows of the Young diagram are allowed to decrease in length. This characterisation only works after we specialise, setting q to be some complex number. For example, if we take $q = i$ then $t = q^2 = -1$ and $1 + (-1) = 0$, so in the following theorem $e = 2$.

Theorem 14.5.14. Let e be the smallest positive integer such that $1 + t + \cdots + t^{e-1} = 0$, with $t = q^2$, and $e = \infty$ if no such integer exists. Then $\tilde{V}_\lambda$ is nontrivial if and only if $\lambda_i - \lambda_{i+1} < e$ for all $i \geq 1$.

14.5.1.2 Seminormal Representations

Recall that for the symmetric group, S_n, we had Young's seminormal form. To construct this fix a partition, λ, of n. We take a space spanned by v_T where T is a standard λ-tableau. For the generators, $s_i \in S_n$, we define

$$v_{s_i.T} = \begin{cases} v_{s_i.T} & \text{if } s_i\,.\,T \text{ is a standard } \lambda\text{-tableau,} \\ 0 & \text{otherwise,} \end{cases} \tag{14.5.15}$$

where on the right $s_i\,.\,T$ means s_i acting on the boxes of T according to their labels. We also define the content of a box to be $C_T(k) = j - i$ when $T(i, j) = k$, that is the box in position (i, j) is the one labelled k. We can then define the action of S_n on this space by

$$s_i\,.\,v_T = \frac{1}{C_T(i+1) - C_T(i)} v_T + \left(1 + \frac{1}{C_T(i+1) - C_T(i)}\right) v_{s_i.T}. \tag{14.5.16}$$

The benefit of this definition is that the Jucys–Murphy elements, $L_j \coloneqq \sum_{1 \leq i < j} (i\,j)$, act as a scalar:

$$L_j\,.\,v_T = c_T(j) v_T. \tag{14.5.17}$$

In order to perform a similar construction for the Hecke algebra we need to complexify, performing a basis change we get the complex Hecke algebra,

$$H_n^{\mathbb{C}} \coloneqq H_n \otimes_{\mathbb{Q}} \mathbb{C}. \tag{14.5.18}$$

Alternatively, we can define this just as we defined the rational Hecke algebra, but replacing $\mathbb{Q}(q)$ with $\mathbb{C}(q)$. The analogue of the Jucys–Murphy elements is then

$$L_j = \sum_{1 \leq i < j} T_{(i\,j)}. \tag{14.5.19}$$

Lemma 14.5.20 Let $M_j = T_{j-1} \cdots T_2 T_1^2 T_2 \cdots T_{j-1}$. Then

$$L_j = \frac{M_j - 1}{q - q^{-1}}. \tag{14.5.21}$$

The corresponding representation space is

$$V_\lambda = \mathrm{span}_{\mathbb{C}(q)}\{v_T \mid T \text{ is a standard } \lambda\text{-tableau}\}. \tag{14.5.22}$$

Theorem 14.5.23. Under the action

$$T_i.v_T = \frac{q - q^{-1}}{1 - q^{C_T(i) - C_T(i+1)}} v_T + \left(q^{-1} + \frac{q - q^{-1}}{1 - q^{C_T(i) - C_T(i+1)}}\right) v_{s_i.T}. \tag{14.5.24}$$

V_λ is an $H_n^{\mathbb{C}}$-module.

Lemma 14.5.25 The Jucys-Murphy elements act diagonally on V_λ, specifically,

$$M_j \,.\, v_T = q^{2C_T(j)} v_T \tag{14.5.26}$$

so

$$L_j \,.\, v_T = \frac{q^{2C_T(j)} - 1}{q - q^{-1}} v_T. \tag{14.5.27}$$

Theorem 14.5.28. The modules V_λ with λ a partition of n define a complete set of pairwise non-isomorphic simple $H_n^{\mathbb{C}}$-modules.

Remark 14.5.29 Write $H_n(q)$ for $H_n^{\mathbb{C}}$ with q specialised to some value in $\mathbb{C}$.

- $H_n(1) \cong \mathbb{C}S_n$
- If q is not a root of unity then $H_n(q)$ is semisimple, and is isomorphic to $\mathbb{C}S_n$ as vector spaces. The simple modules, V_λ^q, are therefore labelled by partitions of n, with a Gelfand–Tsetlin basis with a q-deformed version of the S_n action. We also get the decomposition

$$H_n(q) \cong \bigoplus_\lambda \mathrm{End}\, V_\lambda^q. \tag{14.5.30}$$

- If $q^{2d} = 1$ then the representation theory of $H_n(q)$ is similar to the representation theory of S_n over a field of characteristic d. In particular, if $n < d$ then $H_n(q)$ is still semisimple.

Fifteen

Quantum Groups

We will now give a *very* brief introduction to quantum groups. We assume familiarity with the notion of a Hopf algebra. We also use the language of monoidal categories, but these aren't essential to understanding what's going on, they're just motivating.

If you're not comfortable with monoidal categories just consider $\Bbbk\text{-}\mathsf{Vect}_{\mathbb{C}}$. This is equipped with a tensor product, which is such that $V \otimes W$ is a vector space for all vector spaces V and W (over $\mathbb{C}$). This has the property that it is associative up to isomorphism, $\alpha_{U,V,W} \colon U \otimes (V \otimes W) \xrightarrow{\sim} (U \otimes V) \otimes W$. There is also a unit (up to isomorphism) of the tensor product, which is just $\mathbb{C}$ as a vector space over itself, so we have isomorphisms $\lambda_V \colon \mathbb{C} \otimes V \xrightarrow{\sim} V$ and $\rho_V \colon V \otimes \mathbb{C} \xrightarrow{\sim} V$. For each vector space, V, we also have its dual, $V^* = \mathrm{Hom}(V, \mathbb{C})$, which is again a vector space, this property is called rigidity. This tensor product is braided (in fact, it's symmetric), meaning we have an explicit isomorphism $\sigma_{V,W} \colon V \otimes W \xrightarrow{\sim} W \otimes V$.

There are some compatibility conditions on all of these, namely $- \otimes -$ and $(-)^*$ are functorial, and $\alpha_{U,V,W}$, λ_V, ρ_V, and $\sigma_{V,W}$ are all component of some natural transformations,

- $\alpha \colon - \otimes (- \otimes -) \Rightarrow (- \otimes -) \otimes -$;
- $\lambda \colon \mathbb{C} \otimes - \Rightarrow \mathrm{id}$;
- $\rho \colon - \otimes \mathbb{C} \Rightarrow \mathrm{id}$;
- $\sigma \colon - \otimes - \Rightarrow - \otimes -$.

There are some diagrams formed from these maps which must commute.

To get the definition of a monoidal category we just replace vector spaces with the appropriate category.

15.1 Quasitriangularity

For a general algebra, A, the category A-Mod, of A-modules and A-module homomorphisms is "just" a category.

If instead we have a bialgebra, B, then the category B-Mod, of B-modules is a monoidal category. The tensor product M and N in B-Mod is defined to be the tensor product of the underlying vector spaces, $M \otimes N$, with the action defined on simple tensors by

$$b \,.\, (m \otimes n) = \Delta(b)(m \otimes n). \tag{15.1.1}$$

If in Sweedler notation $\Delta(b) = \sum b_{(1)} \otimes b_{(2)}$ then this action is given by

$$b \,.\, (m \otimes n) = \sum (b_{(1)} \,.\, m) \otimes (b_{(2)} \,.\, n) \tag{15.1.2}$$

where on the right we just have the actions of B on M and N respectively.

If we further add an antipode, χ, to get a Hopf algebra, H, then the category H-Mod, of H-modules is a **rigid monoidal category**. That is, every object has a dual, in the case of H-Mod (over $\mathbb{C}$) the dual of M is the dual module, $M^* = \mathrm{Hom}(M, \mathbb{C})$, with the action defined by

$$(h \,.\, f)(m) = f(h \,.\, m), \tag{15.1.3}$$

where on the right we have the action of H on M.

We see that adding structure, going from an algebra to a bialgebra to a Hopf algebra, adds structure to the category of modules. A natural question then is what structure do we need to add to a bialgebra (Hopf algebra) to get a *braided* (rigid) monoidal category? We'll assume a Hopf algebra, since that's the most useful case, but most of what we're about to do works with a bialgebra. To get a braiding we're looking for an element which acts on the tensor product in the way a braiding would.

Definition 15.1.4 — Quasitriangular Let H be a Hopf algebra. We say that H is **quasitriangular** if there exists some invertible element[a] $\mathcal{R} \in H \hat{\otimes} H$, called the **universal *R*-matrix**, such that

- $\mathcal{R}\Delta(x)\mathcal{R}^{-1} = \Delta^{\mathrm{op}}(x)$ for all $x \in H$ (note $\Delta^{\mathrm{op}} = P \circ \Delta$ where $P(u \otimes v) = v \otimes u$);
- $(\Delta \otimes \mathrm{id})(\mathcal{R}) = \mathcal{R}_{13}\mathcal{R}_{23}$;
- $(\mathrm{id} \otimes \Delta)(\mathcal{R}) = \mathcal{R}_{13}\mathcal{R}_{12}$

where subscripts on $\mathcal{R}$ denote which factors of a tensor product it acts on. For example, $\mathcal{R}_{13}$ is the image of $\mathcal{R}$ under the map $H^{\otimes 2} \to H^{\otimes 3}$ given by $a \otimes b \mapsto a \otimes 1 \otimes b$.

[a] $H \hat{\otimes} H$ is some appropriate completion of $H \otimes H$ to include infinite sums. Outside of this definition we'll often drop this notation, and either have an implicit completion or it won't actually be needed.

This is a slightly complicated definition, but the key idea is that we're only imposing on $\mathcal{R}$ the requirements such that when given a tensor product of H-modules the obvious action of $\mathcal{R}$ on this tensor product is a braiding. Specifically, the braiding is

$$\sigma_{U,V}(u \otimes v) = P(\mathcal{R} \,.\, (u \otimes v)) = \sum P(\mathcal{R}_1 \,.\, u \otimes \mathcal{R}_2 v) = \sum \mathcal{R}_2 \,.\, v \otimes \mathcal{R}_1 \,.\, u \tag{15.1.5}$$

where in the penultimate equality we're choosing some expansion of $\mathcal{R} \in H \otimes H$ of the form $\mathcal{R} = \sum \mathcal{R}_1 \otimes \mathcal{R}_2$ in Sweedler notation.

Example 15.1.6 Consider a Lie algebra, $\mathfrak{g}$. Then the universal enveloping algebra, $U(\mathfrak{g})$, is a Hopf algebra with

$$\Delta(x) = x \otimes 1 + 1 \otimes x. \tag{15.1.7}$$

We have $\Delta^{\mathrm{op}} = \Delta$ in this case, and this is trivially quasitriangular taking $\mathcal{R} = 1 \otimes 1$.

One can show that a consequence of the coassociativity of H is that the universal R-matrix always satisfies the **Yang–Baxter equation**:

$$\mathcal{R}_{12}\mathcal{R}_{13}\mathcal{R}_{23} = \mathcal{R}_{23}\mathcal{R}_{13}\mathcal{R}_{12}, \tag{15.1.8}$$

which is an equation in $H \otimes H \otimes H$. This comes from requiring that Δ^{op} also makes H into a Hopf algebra, and then using the definition of $\mathcal{R}$ to replace Δ^{op} with Δ conjugated by $\mathcal{R}$.

It is useful to introduce the flip operator, $P\colon H \otimes H \to H \otimes H$, $P(a \otimes b) = b \otimes a$, and then define $\check{\mathcal{R}} = P \circ \mathcal{R}$. The flip operator also satisfies the Yang–Baxter equation. Starting with the Yang–Baxter equation for $\mathcal{R}$ and acting with P_{ij}, which is just the flip operation acting on the ith and jth components it is possible to manipulate the Yang–Baxter equation for $\mathcal{R}$ into the form of the braid equation for $\check{\mathcal{R}}$:

$$\check{\mathcal{R}}_{23}\check{\mathcal{R}}_{12}\check{\mathcal{R}}_{23} = \check{\mathcal{R}}_{12}\check{\mathcal{R}}_{23}\check{\mathcal{R}}_{12}. \tag{15.1.9}$$

Note that $\check{\mathcal{R}}$ doesn't (necessarily) satisfy the Yang–Baxter equation[1]. It is $\check{\mathcal{R}}$ which acts as the braiding in H-Mod.

[1] Often people don't distinguish very well between $\mathcal{R}$ and $\check{\mathcal{R}}$ or between the Yang–Baxter and braid equations.

Definition 15.1.10 — Quantum Group A **quantum group** is a (not necessarily commutative or cocommutative) quasitriangular Hopf algebra.

15.2 Quantum Schur–Weyl Duality

Recall that "classical" Schur–Weyl duality is a statement on the compatibility of the actions of S_n and GL_m on $(\mathbb{C}^m)^{\otimes n}$. Specifically, it says that if $M = \mathbb{C}^m$ then we have the decomposition

$$M^{\otimes n} \cong_\lambda V_\lambda \otimes L_\lambda \tag{15.2.1}$$

where V_λ and L_λ are simple S_n-modules and simple $U(\mathfrak{sl}_m)$-modules respectively. Further, these modules are related by

$$V_\lambda \cong \mathrm{Hom}_{U(\mathfrak{gl}_m)}(L_\lambda, M^{\otimes n}), \quad \text{and} \quad L_\lambda \cong \mathrm{Hom}_{\mathbb{C}S_n}(V_\lambda, M^{\otimes n}). \tag{15.2.2}$$

Quantum Schur–Weyl duality is the corresponding statement that we get if we replace $\mathbb{C}S_n$ with its deformation, the Hecke algebra, H_n. The question then is what is the "correct" replacement for $U(\mathfrak{sl}_m)$. The answer is the quantum group $U_q(\mathfrak{sl}_m)$. The full definition of this is a fairly complicated algebra with generators and relations. We won't give these relations here. The generators are e_i, f_i, and k_i for $i = 1, \dots, \operatorname{rank} \mathfrak{g}$. This can also be generalised to other Dynkin types, replacing k_i with k_λ where λ is an element of the weight lattice.

15.3 $U_q(\mathfrak{sl}_2)$

Definition 15.3.1
Warning: Conventions differ in the exact definitions, usually differing by scaling some elements by some power of q.
The quantum group, $U_q(\mathfrak{sl}_2)$, is the unital associative algebra generated by e, f, k and k^{-1} subject to the relations

- $kk^{-1} = 1 = k^{-1}k$;
- $ke = q^2 ek$;
- $kf = q^{-2} fk$;
- $[e, f] = \frac{k-k^{-1}}{q-q^{-1}}$.

The coproduct of this algebra is defined by

$$\Delta(k^{\pm 1}) = k^{\pm 1} \otimes k^{\pm 1}, \quad \Delta(e) = e \otimes k^{-1} + 1 \otimes e, \quad \text{and} \quad \Delta(f) = f \otimes 1 + k \otimes f,$$

the counit is defined by

$$\varepsilon(k^{\pm 1}) = 1, \quad \text{and} \quad \varepsilon(e) = \varepsilon(f) = 0, \tag{15.3.2}$$

and the antipode is defined by

$$\chi(k^{\pm 1}) = k^{\mp 1}, \quad \chi(e) = -ek, \quad \text{and} \quad \chi(f) = -k^{-1} f. \tag{15.3.3}$$

Strictly, $U_q(\mathfrak{sl}_2)$ is just a Hopf algebra, it isn't actually quasitriangular, however, if we instead work in a completion of $U_q(\mathfrak{sl}_2)$ then we are allowed the element

$$\mathcal{R} = \left(\sum_{n=0}^{\infty} q^{n(n+1)/2} \frac{(1-q^2)^n}{[n]_q!} e^n \otimes f^n \right) q^{-h \otimes h/2} \tag{15.3.4}$$

which does act as a universal R-matrix for this completion. Here $[n]_q!$ is the **quantum factorial**, defined in terms of the **quantum integer**

$$[n]_q = \frac{q^n - q^{-n}}{q - q^{-1}}, \tag{15.3.5}$$

such that $[n]_q! = [n]_q [n-1]_q \cdots [1]_q$.

The idea behind these definitions is that when we set $k = q^h = e^{\hbar h}$ and take $\hbar \to 0$ we recover (at least formally) $U(\mathfrak{sl}_2)$.

The fundamental representation of $U_q(\mathfrak{sl}_2)$, also known as the vector representation, is $L_1 \cong \mathbb{C}^2 = \mathbb{C}v_0 \oplus \mathbb{C}v_1$, with the action

$$ev_0 = fv_1 = 0, \quad ev_1 = v_0, \quad fv_0 = v_1, \quad kv_0 = qv_0, \quad \text{and} \quad kv_1 = q^{-1} v_1. \tag{15.3.6}$$

Schur-Weyl duality is then the statement that

$$L_1^{\otimes r} = \sum_{m=0}^{r} V_m \otimes L_m \tag{15.3.7}$$

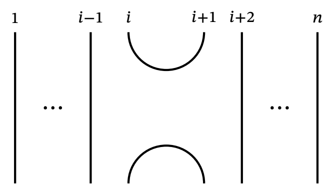
where L_m are more simple $U_q(\mathfrak{sl}_2)$-modules and the V_m are simple TL_n-modules, where TL_n is the **Temperley–Lieb** algebra, defined as a quotient of the Hecke algebra,

$$\mathrm{TL}_n = H_n / \langle 1 + T_i + T_{i+1} + T_i T_{i+1} + T_{i+1} T_i + T_i T_{i+1} T_i \rangle. \tag{15.3.8}$$

The Temperley–Lieb algebra can also be given as the unital algebra generated by e_i for $i = 1, \ldots, n-1$ subject to the relations

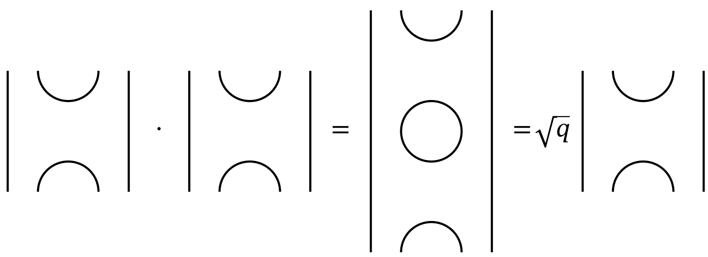
- $e_i^2 = \delta e_i$ where δ is some fixed complex number;
- $e_i e_{i+1} e_i = e_i$;
- $e_i e_{i-1} e_i = e_i$;
- $e_i e_j = e_j e_i$ for $|i - j| > 2$.

Graphically, we can represent the generator e_i as



(15.3.9)

The product of two such elements is their vertical concatenation, which we interpret with the rule that any circle is just the scalar δ. For example, with $n = 3$ the relation $e_2^2 = \delta e_2$ becomes



(15.3.10)

The representation theory of $U_q(\mathfrak{sl}_2)$ is not that different from the representation theory of $U(\mathfrak{sl}_2)$. In particular, there is a notion of a Verma module, where for the quantum case we replace the highest weight, $n \in \mathbb{Z}$, with the quantum integer $[n]_q$.

We can often understand a topological object, such as the plane, by looking at certain algebras of functions on the space. For the plane, we may look at the algebra of polynomial functions on the plane, which is just $\mathbb{C}[x, y]$. There is a natural action of $\mathfrak{sl}_2$ on this, in which it acts by difference operations, in particular, if we define

$$e = x\partial_y, \quad f = y\partial_x, \quad \text{and} \quad h = x\partial_x - y\partial_y \tag{15.3.11}$$

then we can check that these satisfy the $\mathfrak{sl}_2$ commutation relations.

There is a general philosophy of quantisation that to get the equivalent "quantum" result we should replace commuting variables with non-commuting variables. The **quantum plane**[2] is $\mathbb{C}_q[x, y] \coloneqq \mathbb{C}\langle x, y\rangle / \langle yx - qxy\rangle$. That is, polyno-

[2] In analogy to the normal plane the quantum plane is really the space for which this is the algebra of (polynomial) functions, but that's not really a space that exists, so we just call this the quantum plane.

mials in x and y, which no longer commute, but instead $yx = qxy$. Then there is an action of $U_q(\mathfrak{sl}_2)$ on this in which we replace the derivatives above with the corresponding **quantum derivatives**, defined by

$$\delta_x f(x) = \frac{f(qx) - f(q^{-1}x)}{qx - q^{-1}x}. \tag{15.3.12}$$

Then defining

$$e = x\delta_y, \quad \text{and} \quad f = y\delta_x \tag{15.3.13}$$

and k acts by

$$k \,.\, x = qx, \quad \text{and} \quad k \,.\, y = q^{-1}y. \tag{15.3.14}$$

Working with the universal R-matrix of $U_q(\mathfrak{sl}_2)$ (or rather its completion) is somewhat tricky. It's usually better to just compute the R-matrix, R, for the given representation. To do so we appeal to the relations defining $U_q(\mathfrak{sl}_2)$. We'll do this here for the fundamental representation, for which a useful basis is $\{v_{-1}, v_1\}$ with the action defined by

$$k^{\pm 1} \,.\, v_i = q^{\pm i} v_i, \quad e \,.\, v_{-1} = v_1, \quad e \,.\, v_1 = 0, \quad f \,.\, v_{-1} = 0, \quad \text{and} \quad f \,.\, v_1 = v_{-1}.$$

We see from this that v_1 is a highest weight vector (it's an eigenvector of k and annihilated by e) and v_{-1} is a lowest weight vector (it's an eigenvector of k and annihilated by f). The tensor product of highest (lowest) weight vectors is again a highest (lowest) weight vector, and so we have $\Delta(e)(v_1 \otimes v_1) = \Delta(f)(v_{-1} \otimes v_{-1}) = 0$.

We can act with R on these, and we should still get zero. Note that R is just the image of $\mathcal{R}$ induced by the representation map $U_q(\mathfrak{sl}_2) \to \operatorname{End}\mathbb{C}^2$ and the coproduct, so $R \in \operatorname{End}(\mathbb{C}^2 \otimes \mathbb{C}^2) \cong \operatorname{End}(\mathbb{C}^4)$, and thus we can write R as

$$R = \begin{pmatrix} a & 0 & 0 & 0 \\ 0 & b & c & 0 \\ 0 & c' & b' & 0 \\ 0 & 0 & 0 & a' \end{pmatrix} \tag{15.3.15}$$

for some $a, b, c, a', b', c' \in \mathbb{C}$ where we're using the ordered basis $\{v_1 \otimes v_1, v_1 \otimes v_{-1}, v_{-1} \otimes v_1, v_{-1} \otimes v_{-1}\}$.

We can then do more calculations, such as

$$\begin{aligned} R\Delta(e)v_{-1} \otimes v_{-1} &= R(e \otimes k^{-1} + 1 \otimes e)v_{-1} \otimes v_{-1} && (15.3.16) \\ &= R(ev_{-1} \otimes k^{-1}v_{-1} + 1v_{-1} \otimes ev_{-1}) && (15.3.17) \\ &= R(v_1 \otimes qv_{-1} + v_{-1} \otimes v_1) && (15.3.18) \\ &= qR(v_1 \otimes v_{-1}) + R(v_{-1} \otimes v_1) && (15.3.19) \\ &= qcv_1 \otimes v_{-1} + qb'v_{-1} \otimes v_1 + bv_1 \otimes v_{-1} + c'v_{-1} \otimes v_1 \end{aligned}$$

and

$$\begin{aligned} \Delta^{\mathrm{op}}(e)v_{-1} \otimes v_{-1} &= (k^{-1} \otimes e + e \otimes 1)v_{-1} \otimes v_{-1} && (15.3.20) \\ &= k^{-1}v_{-1} \otimes ev_{-1} + ev_{-1} \otimes 1v_{-1} && (15.3.21) \\ &= qv_{-1} \otimes v_1 + v_1 \otimes v_{-1}. && (15.3.22) \end{aligned}$$

The axioms of a universal R-matrix mean that these two results should be equal, and so if we equate coefficients we find that

$$qc + b = 1, \quad \text{and} \quad qb' + c' = q. \tag{15.3.23}$$

Continuing like this we eventually can eliminate all unknowns, and we get

$$\check{R} = \begin{pmatrix} q & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & q - q^{-1} & 0 \\ 0 & 0 & 0 & q \end{pmatrix}. \tag{15.3.24}$$

Lemma 15.3.25 Let $V = \mathbb{C}^2$. The action of $\mathcal{B}_n$ on $V^{\otimes n}$ acting by permutations factors through $H_n(q)$ to $\mathrm{TL}_n(q)$ where $\mathrm{TL}_n(q)$ is the Temperley–Lieb algebra with $\delta = q + q^{-1}$.

These two results are useful to get knot invariants, first look at the braid given by slicing the knot, then pass through the quotient to the Temperley–Lieb algebra, and we'll get a quantity that should be an invariant. For example, the Jones polynomial arises in this way, although it was originally discovered using operator algebras, since the underlying Temperley–Lieb algebra structure was not known at the time of discovery.