Syllabus: Basic definitions and examples: subrepresentations and irreducible representations, tensor products of representations. Character theory. Representations as modules over the group ring, Artin-Wedderburn structure theorem for semisimple rings. Induced representations, Frobenius reciprocity, Mackey's irreducibility criterion.

Let's begin this course by recalling a famous result from group theory: Cayley's theorem. This result states that every finite group embeds in a symmetric group. It's easy to prove—one simply looks at the action of G on itself by left multiplication and one shows that this gives an embedding of G into S_n where $n = |G|$. This is not a difficult result, but it persists as part of undergraduate mathematics syllabi to this day because it allows one to understand finite groups as subgroups of symmetric groups, where one has all sorts of tools one can use.

In a similar vein, representation theory seeks to understand groups in terms of maps into general linear groups, $\mathrm{GL}_n(F)$, where F is a field. If one thinks about it, this is generally even more desirable than mapping into a symmetric group because now one has the full power of linear algebra at one's disposal. As it turns out, we will consider all homomorphisms—not just injective ones—from groups into general linear groups, and we will call such homomorphisms *representations* of our group. Let's begin by showing that every finite group embeds into $\mathrm{GL}_n(F)$ for some field F.

Exercise 1. *Let F be a field. Show that every finite group embeds into $\mathrm{GL}_n(F)$ for some $n \geq 1$.*

One might wonder whether this holds for infinite groups if one is allowed to chose the field F. As it turns out, this is not the case.

Exercise 2. *Let G denote the group consisting of bijective maps f from $\mathbb{Z}$ to itself such that f fixes all but finitely many integers. Show that there does not exist a field F and an $n \geq 1$ such that G embeds in $\mathrm{GL}_n(F)$.*

We'll do things a bit out of order from the syllabus. The reason for this is that I would like to lay the technical foundations for representation theory before doing the actual theory. We'll keep the analogy with Cayley's theorem in mind as we consider three different ways of considering representations of a group G. With Cayley's theorem, one can look at a finite group G as embedding in some S_n. But one can also look at G as *acting* on the set $X := \{1, \ldots, n\}$ via the induced action of S_n on X. Similarly, if we have a homomorphism $\phi : G \to \mathrm{GL}_n(F)$ then since $\mathrm{GL}_n(F)$ acts on the space V of $n \times 1$ column vectors via left multiplication, we see that V inherits a structure as a *G-module*. More precisely, a G-module V is an F-vector space endowed with a map $\cdot : G \times V \to V$ satisfying $g \cdot (v + \lambda w) = g \cdot v + \lambda g \cdot w$ and $(gh) \cdot v = g \cdot (h \cdot v)$ and $1 \cdot v = v$ for $g, h \in G$, $v, w \in V$, and $\lambda \in F$. One can understand this better in terms of a third point of view: the *group algebra*.

Group algebras

Let F be a field and let G be a group (not necessarily finite). We define the *group algebra* $F[G]$ to be the set of all elements of the form

$$\{\sum_{g \in G} \alpha_g g \colon \alpha_g \in F,\ \alpha_g = 0 \text{ for all but finitely many } g\}.$$

Then $F[G]$ has a natural addition, and multiplication is given by $(\alpha g) \cdot (\beta h) = (\alpha\beta)gh$ and then extending by linearity. This turns $F[G]$ into a (not necessarily commutative) ring with identity given by $1_F \cdot 1_G$.

For those who have not done any ring theory, we recall that if R is a ring, a *left R-module* M is just an abelian group endowed with a map $\cdot : R \times M \to M$ satisfying $(rs) \cdot m = r \cdot (s \cdot m)$, $r \cdot (m + n) = r \cdot m + r \cdot n$, $(r + s) \cdot m = r \cdot m + s \cdot m$, and $1 \cdot m = m$ for $r, s \in R$ and $m, n \in M$. Right modules are defined analogously. A submodule of a left R-module M is a non-empty subset N that is an subgroup of M and is closed under left multiplication by R. Left and right ideals of R are respectively left and right R-submodules of R.

Exercise 3. *Show that a group homomorphism $\phi : G \to \mathrm{GL}_n(F)$ extends to a ring homomorphism $\psi : F[G] \to \mathrm{M}_n(F)$ and that if V is the set of $n \times 1$ column vectors with entries in F then V is a left $F[G]$-module.*

In this way, we can view G-modules as being the same as left $F[G]$-modules. Thus a representation of a group can be understood in three ways: as a homomorphism from the group into a general linear group; as a homomorphism of the group algebra into a matrix ring; or in terms of having a finite-dimensional G-module.

The most important result about group algebras comes from work of Artin and Wedderburn. In short, this implies that if F is an algebraically closed field of characteristic zero and G is a finite group then $F[G]$ is isomorphic to a finite direct product of matrix rings over F. This isomorphism and these matrix rings will completely determine the representation theory of the group G.

The Artin-Wedderburn theorem

We recall that a ring R is *left artinian* if every descending chain of left ideals of R

$$I_0 \supseteq I_1 \supseteq I_2 \supseteq \cdots$$

terminates; i.e., there exists some n such that $I_n = I_{n+1} = \cdots$.

Exercise 4. *Show that being left artinian is equivalent to every non-empty set of left ideals having a minimal element (not necessarily unique) with respect to inclusion.*

We recall that, given a field F, a ring R is an F-algebra if it has a homomorphism $\phi : F \to R$ sending 1_F to 1_R and such that the image of ϕ is contained in the centre of R. Intuitively, it just means that R has a copy of F inside and R is an F-vector space in addition to being a ring.

Proposition 0.1. *Let R be a ring and suppose that R is a finite-dimensional F-algebra. Then R is artinian.*

As it turns out, work of Artin and Wedderburn completely gives the structure theory of Artinian rings. We recall that a two-sided ideal I of a ring R is *nil* if for every $x \in I$ there is some $n = n(x) > 0$ such that $x^n = 0$.

Theorem 0.1. *(Artin-Wedderburn) Let R be an artinian ring and suppose that R has no nonzero nil two-sided ideals. Then R is isomorphic to a direct product of matrix rings*

$$\prod_{i=1}^{s} M_{n_i}(D_i),$$

where the D_i are division rings.

In the case that R is commutative, notice that the matrix rings must be 1×1 matrix rings and the division rings are fields, so in this case the conclusion is stronger: R is a direct product of fields when R is a commutative artinian ring with no nonzero nil ideals.

In the case that R is a finite ring, another result of Wedderburn states that finite division rings are actually finite fields and so the conclusion becomes that R is a direct product of matrix rings over finite fields when R is a finite artinian ring with no nonzero nil ideals. Let's do Wedderburn's theorem for the sake of completeness.

Exercise 5. *Prove that a finite division ring is a field using the following steps. Let D be a finite division ring.*

1 *Show that the centre Z of D is a field and has size q for some prime power q. Show that D has size q^n for some $n \geq 1$.*
2 *Let $G = D^*$ be the multiplicative group of D. Then $|G| = q^n - 1$. Use the class equation to show that*

$$q^n - 1 = |Z| + \sum_g |C_g| = q - 1 + \sum_g (q^n - 1)/|C(g)|,$$

where the sums runs over a complete set of non-central conjugacy class representatives and C_g denotes the conjugacy class of g and $|C(g)|$ denotes the centralizer of g in G.
3 *Show that if $g \in D^*$ then the centralizer of g in D is a division ring E that properly contains Z. Conclude that $|C(g)| = q^m - 1$ for some m.*
4 *Show that $q^m - 1$ divides $q^n - 1$ if and only if m divides n. Conclude that $|C(g)| = q^d - 1$ for some d dividing n and $d > 1$.*
5 *Rewrite the class equation as*

$$q^n - 1 = (q - 1) + \sum_{j=1}^{r} (q^n - 1)/(q^{d_j} - 1),$$

where r is the number of non-central conjugacy class representatives $d_1, \ldots, d_r > 1$ are divisors of n.
6 *Remember! Our goal is to show that D is a field, so we want to show $D = Z$ and so $n = 1$. Let $P(x) = \prod(x - \zeta)$, where ζ runs over all primitive n-th roots of unity. You can use the following fact: $P(x)$ is a monic polynomial with integer coefficients. (We'll show this later on when we talk about characters, but if you know a bit of Galois theory, you can convince yourself that the coefficients of $P(x)$ are fixed by the Galois group of $\mathbb{Q}(\exp(2\pi i/n))$ over $\mathbb{Q}$ and so the coefficients are rational; also ζ is an algebraic integer since it satisfies $\zeta^n - 1 = 0$—since the algebraic integers form a ring we see the coefficients are rational algebraic integers and hence integers. If you don't understand this, don't worry about it.) Show that $(x^n - 1) = P(x)Q(x)$ where $Q(x)$ is a monic integer polynomial and $x^d - 1$ divides $Q(x)$ in $\mathbb{Z}[x]$ for every divisor d of n with $d < n$.*
7 *Now show from step 5 that $P(q)$ divides $q - 1$.*
8 *Now we're ready to finish. Show that if $n > 1$ then $|P(q)| > q - 1$ and conclude that $n = 1$ and $D = Z$.*

We'll prove the Artin-Wedderburn theorem later, but for now let's find some nice consequences.

Exercise 6. *Let R be a finite ring and suppose that for each $x \in R$ there is some $n = n(x) > 1$ such that $x^n = x$. Show using the Artin-Wedderburn theorem and Wedderburn's theorem on finite division rings that R is commutative.*

Notice that when $n(x) = 2$ for all n this is rather easy:

$$x + y = (x + y)^2 = x^2 + xy + yx + y^2 = x + xy + yx + y.$$

Hence $xy + yx = 0$. But $(-1)^2 = -1$ and so R has characteristic 2 and so $xy = yx$ and we see that R is commutative. In the case when R is a finite-dimensional F-algebra we can say even more.

Exercise 7. *Let F be an algebraically closed field and let R be a finite-dimensional F-algebra and suppose that R has no nonzero nil ideals. Show that R is isomorphic to a finite direct product of matrix rings over F.*

And let's prove one more nice fact. Given a ring R, we say that a proper two-sided ideal P is *prime* if whenever $a, b \in R$ are such that $aRb \subseteq P$ we must have $a \in P$ or $b \in P$. We say that R is a *prime ring* if (0) is a prime ideal of R. One can think of a prime ring as being a generalization of being an integral domain in the commutative case. So we have a few examples.

Example 0.2. Let R be a commutative ring. Then R is a prime ring if and only if it is an integral domain.

Recall that a ring is simple if and only if it has no nonzero proper two-sided ideals.

Example 0.3. Let R be a simple ring. Then R is a prime ring.

Example 0.4. If R is a non-trivial ring with identity then $R \times R$ is never prime.

Throughout these notes, we assume that all rings have 1 and $0 \neq 1$.

Exercise 8. *Let R be a ring. Show that if S is a subset of R that is closed under multiplication and does not contain zero then there is some prime ideal P of R with $P \cap S = \emptyset$. Conclude by taking $S = \{1\}$ that R always has at least one prime ideal.*

Exercise 9. *Let R be a ring (with identity). Show that the intersection of the prime ideals of R is a nil ideal. (Hint: use the preceding exercise.)*

We say that a ring R is (left) primitive if there is a left R-module M with the following two properties:

(i) M is simple; i.e., the only submodules of M are (0) and M;
(ii) M is faithful; i.e., if $r \cdot m = 0$ for all $m \in M$ then $r = 0$.

An ideal P of R is primitive if R/P is a primitive ring.

The faithful condition says that if we look at $\mathrm{Ann}(M) = \{r \in R \colon rM = (0)\}$ then $\mathrm{Ann}(M)$ is (0). Notice that $\mathrm{Ann}(M)$ is a two-sided ideal of R and it is called the *annihilator* of M. We shall see that we have the inclusions

$$\{\text{prime rings}\} \supset \{\text{primitive rings}\} \supset \{\text{simple rings}\}.$$

This exercise will show the inclusions are strict.

Exercise 10. *Let R be the $\mathbb{C}$-algebra with generators x and y and relations $xy = 2yx$. (You can think of R as being a polynomial ring in 2-variables x and y but where the variables skew-commute. Show that R is primitive and xR is a two-sided proper ideal of R. Conclude that R is prime, primitive, but not simple. Let $R = \mathbb{C}[x]$. Show that R is prime but not primitive. (Hint: primitivity is tricky. Here's how one can do it. First show that every nonzero ideal of R contains an element of the form $(xy)^a$ with $a \geq 1$. Now let $I = R(xy - 1)$. Show that I is a proper left ideal. Use Zorn's lemma to show that there is a maximal left ideal L containing I. Then $M := R/L$ is a simple left R-module. Show that it must be faithful.)*

To see that a primitive ring is prime, suppose that $aRb = 0$ and let M be a faithful simple R-module. If $b \neq 0$ then $bM \neq (0)$ since M is faithful. Thus $RbM = M$ since M is simple and RbM is a nonzero submodule. But then if $a \neq 0$ then $aRbM \neq (0)$ since M is faithful. It follows that if $aRb = (0)$ then $aRbM = (0)$ and so $a = 0$ or $b = 0$. To see that a primitive ring is simple, we recall that by Zorn's lemma R always has a maximal left ideal I. Then $M = R/I$ is a simple R-module. It suffices to show that M is faithful. But if it is not, then its annihilator is a nonzero two-sided ideal and hence all of R. But this is absurd since 1 cannot annihilate M.

We now come to the Jacobson density theorem, which shows that primitive rings embed densely in a ring of linear operators.

Theorem 0.2. *(Jacobson density theorem) Let R be a primitive ring and let M be a faithful simple left R-module. Then R embeds densely in $\mathrm{End}_\Delta(M)$ via the rule $r \mapsto \Phi_r$, where $\Phi_r(m) = rm$. Moreover, if $m_1, \ldots, m_n$ are left Δ-linearly independent elements of M and $w_1, \ldots, w_n$ are in M then there exists some $r \in R$ such that $r \cdot m_i = w_i$ for $i = 1, \ldots, n$.*

Proof. The fact that the map $r \mapsto \Phi_r$ is a homomorphism is routine. The fact that it is injective comes from looking at the kernel: if $\Phi_r = 0$ then $rm = 0$ for all $m \in M$ and so $r = 0$ since M is faithful.

So now we prove the density part by induction on n. When $n = 1$ we have $Rm_1 = M$ since m_1 is nonzero and M is simple, so there is some $r \in R$ such that $rm_1 = w_1$. Now suppose that the claim holds for sets of size less than n. Then we may assume without loss of generality that $rm_1 = \cdots = rm_{n-1} = 0$ implies $rm_n = 0$. This means (using the induction hypothesis) that we have a well-defined R-module homomorphism

$$\Psi : M^{n-1} \to M$$

given by $\Psi((rm_1, \ldots, rm_{n-1})) = rm_n$. Now by the induction hypothesis, there for $j = 1, \ldots, n-1$ there is some r_j such that $r_j m_i = \delta_{i,j} m_j$ for $i = 1, \ldots, n-1$. In particular, $\Psi((0, 0, \ldots, m_j, 0, \ldots, 0)) = \Psi((r_j m_1, \ldots, r_j m_{n-1})) = r_j m_n$. Now the map $f_j : M \to M$ given by $m \mapsto \Psi((0, 0, \ldots, m, 0, \ldots, 0))$, where m is in the j-th slot is an element of $\Delta = \mathrm{End}_R(M)$. So

we see that $f_j(m_j) = r_j m_n$. Now consider $(r_1 + \cdots + r_{n-1} - 1)m_i = r_i m_i - m_i = 0$. So since $(r_1 + \cdots + r_{n-1} - 1)$ kills $m_1, \ldots, m_{n-1}$, by our assumption, it must also kill m_n. This

$$\sum_{i=1}^{n-1} r_i m_n = m_n.$$

In other words,

$$\sum_{i=1}^{n-1} f_i(m_n) = m_n,$$

or

$$m_n - f_1 m_1 - \cdots - f_{n-1} m_{n-1} = 0,$$

contradicting independence over Δ. The result follows. □

Now if Δ is a division ring and M is a left Δ-module then it makes sense to call it a vector space. For example, M has a basis as a Δ-module, which one can prove by Zorn's lemma.

Corollary 0.5. *If R is an artinian primitive ring then R is isomorphic to $M_n(D)$ for some division ring D and some $n \geq 1$.*

Proof. Let M be a faithful simple R-module and let $\Delta = \mathrm{End}_R(M)$. We claim that M is finite-dimensional as a left Δ-vector space. To see this, suppose that we have an infinite linearly independent set $m_1, m_2, \ldots$. Then by the Jacobson density theorem, for each $i > 0$ there is some $r_i \in R$ such that $r_i m_1 = \cdots r_i m_i = 0$ and $r_i m_{i+1} \neq 0$. Now let $I_i = \{r \in R \colon rm_1 = \cdots = rm_i = 0\}$. Then each I_i is a left ideal and notice that $I_i \supseteq I_{i+1}$ by definition. Finally, $r_i \in I_i$ but is not in I_{i+1} so we get an infinite descending chain of ideals

$$I_1 \supseteq I_2 \supseteq \cdots,$$

contradicting the left Artinian hypothesis. Thus M is finite dimensional—let's say the dimension is n. Now if Δ were a field, we'd feel pretty good: M would be a finite-dimensional vector space and we'd know that the endomorphism ring is matrices over Δ. In the division ring setting there is one subtlety. We have to introduce the *opposite ring*. Given a ring R the opposite ring R^{op} is just R as set with multiplication $r \star s = s \cdot r$; that is, we reverse the order of multiplication. If R is commutative then its opposite ring is itself. If R is a division ring then so is R^{op}. Now let $m_1, \ldots, m_n$ be a basis for M as a left Δ-vector space. We claim that

$$S := \mathrm{End}_\Delta(M) \cong M_n(\Delta)^{\mathrm{op}} \cong M_n(\Delta^{\mathrm{op}}).$$

To define the isomorphism, let $f \in S$. Then $f(m_i) = \sum_j a_{i,j} m_j$ for some $a_{i,j} \in \Delta$. We define a map $\Phi : S \to M_n(\Delta)^{\mathrm{op}}$ by $\Phi(f) = (a_{i,j})$. This is definitely well-defined and it's fine to see that $\Phi(f + g) = \Phi(f) + \Phi(g)$. Suppose that $\Phi(g) = (b_{i,j})$. Notice that

$$f \circ g(e_i) = f(\sum_k b_{i,k} e_k) = \sum b_{i,k} \sum_j a_{k,j} m_j.$$

Thus the (i,j) entry of $\Phi(f \circ g)$ is $\sum_k b_{i,k} a_{k,j}$, and so $\Phi(f \circ g)$ is just the product of $(b_{i,j})$ and $(a_{i,j})$ in $M_n(\Delta)$, which is the product of $(a_{i,j})$ and $(b_{i,j})$ in $M_n(\Delta)^{\mathrm{op}}$. So it is a homomorphism. To see that it is 1-to-1, we remark that if f is in the kernel then it must send each m_i to 0 and so it is the zero map. It remains to see why this map Φ is onto. But this just comes from the fact that M is a free Δ-module so we can send $m_1, \ldots, m_n$ wherever we'd like! Finally, we remark that $M_n(\Delta)^{\mathrm{op}} \cong M_n(\Delta^{\mathrm{op}})$, which is a matrix ring over a division ring. □

Remark 0.6. Observe that $M_n(D)$ is a simple ring.

Suppose we have some nonzero ideal I. Then there is some nonzero $A = (a_{i,j}) \in I$. Pick k, ℓ such that $a_{k,\ell} \neq 0$. Then $E_{i,k} A E_{\ell,j} = a_{k,\ell} E_{i,j}$. Since $a_{k,\ell}$ is in D and nonzero, $E_{i,j} \in I$. But now $1 = \sum E_{i,i} \in I$.

Corollary 0.7. *If R is a primitive left artinian ring then R is simple.*

In fact, prime, primitive, and simple are all equivalent for left artinian rings.

Theorem 0.3. *Let R be a prime left Artinian ring. Then $R \cong M_n(D)$ for a division ring D.*

Proof. We note that R is left Artinian means that every non-empty set of left ideals has a minimal element. Take S to be the set of nonzero left ideals and let L be a minimal element of R. Then L is a simple left R-module. Notice that if $a \in R$ annihilates L then we have $aL = (0)$. Now pick nonzero $b \in L$. Then we have $Rb \subseteq L$ and so $aRb = (0)$. But R is prime and so $a = 0$ and so L is faithful. Thus R is primitive and the result follows. □

Corollary 0.8. *Let P be a prime ideal of a left Artinian ring R. Then P is a maximal ideal.*

Proof. Since R is left Artinian, so is R/P by correspondence. Thus R/P is a prime Artinian ring and hence simple. So by correspondence P is a maximal two-sided ideal. □

Remark 0.9. If R is a ring and $P_1, \ldots, P_n$ are distinct maximal ideals then $P_1 \not\supseteq P_2 P_3 \cdots P_n$.

Proof. We prove this by induction on n. When $n = 2$ it is clear. Suppose that it is true up to $n-1$. Then $P_1 \not\supseteq I := P_2 \cdots P_{n-1}$. In particular, there is some $a \in I \setminus P_1$. Also there is some $b \in P_n \setminus I$. So if $P_1 \supseteq IP_n$ then we have $aRb \subseteq P_1$ with $a, b \notin P_1$, a contradiction since P_1 is prime. □

Proposition 0.10. *Let R be a left Artinian ring. Then R has only finitely many prime ideals.*

Proof. If $P_1, P_2, \ldots$ is an infinite set of distinct prime ideals, then since we know prime ideals are maximal in a left Artinian ring, we have $P_{n+1} \not\supseteq P_1 \cdots P_n$. But

$$P_1 \supseteq P_1P_2 \supseteq \cdots$$

is a descending chain, so we must have $P_1P_2 \cdots P_n = P_1P_2 \cdots P_nP_{n+1} \subseteq P_{n+1}$, contradiction. □

Corollary 0.11. *Let R be a left Artinian ring with no nonzero nil ideals. Then R is a product of matrix rings over division rings.*

Proof. Let $P_1, \ldots, P_n$ be the distinct prime ideals of R. By the assignment the intersection of the P_i is a nil ideal and hence zero. We need Chinese Remainder theorem: If $P_1, \ldots, P_n$ are distinct maximal ideals whose intersection is (0) then $R \cong \prod R/P_i$. Let's create a homomorphism $\Phi : R \to \prod R/P_i$ via $r \mapsto (r+P_1, \ldots, r+P_n)$. This is 1-to-1 since the intersection of the P_i is zero. To see that it is onto, by the remark we have that $P_i + \prod_{j \neq i} P_j = R$ since P_i is maximal. So there exists $b_i \in \cap_{j \neq i} P_j$ such that $b_i \in 1 + P_i$. Then $\Phi(b_i) = e_i$, where e_i is the i-th coordinate function. So $\Phi(\sum r_i b_i) = (r_1 + P_1, \ldots, r_n + P_n)$ and so Φ is onto. □

Applications to group algebras

As it turns out, when G is a finite group and F is a field, one generally has that $F[G]$ has no nonzero nil ideals—the only exceptions occur when the characteristic of F divides the order of G. This is a theorem of Maschke. Let's prove this now.

Theorem 0.4. *Let G be a finite group and let F be a field. Then $F[G]$ has no nonzero nil two-sided ideals unless F is a field of characteristic $p > 0$ and $p||G|$.*

Proof. Suppose that the characteristic of F does not divide the order of G. Then let I be a nil ideal of $F[G]$. If I is nonzero then there exists some $x = \sum_g a_g g \in I$. We pick some $h \in G$ such that $a_h \neq 0$. Then by multiplying on the left by h^{-1} we have $xh^{-1} = \sum_g b_g g \in I$ is nonzero with $b_1 \neq 0$; i.e., the coefficient of the identity is nonzero.

Now let's rethink things a bit. We can think of $V := F[G]$ as a $|G|$-dimensional F-vector space with basis given by $B := \{g \colon g \in G\}$. Then an element $y = \sum c_g g \in F[G]$ gives rise to a linear map $f_y : V \to V$ given by $v \mapsto yv$. Notice that if we work with the basis B then for $g \in G$ the element f_g permutes the basis and doesn't fix any elements of the basis if $g \neq 1$. Thus the trace of f_g is 0 if $g \neq 1$. On the other hand, the trace of $f_1 = |G|$. So the trace of f_y is just $c_1|G|$. In particular, the trace of $u := xh^{-1} = b_1|G| \neq 0$ since $b_1 \neq 0$ and the characteristic of F does not divide $|G|$. But by assumption u is nilpotent and so that means that $u^n = 0$ and so $f_u^n = 0$. But the trace of any nilpotent matrix is 0, so this is a contradiction. □

The other direction is done in this exercise.

Exercise 11. *Let F be a field of positive characteristic $p > 0$. Suppose that $p||G|$. Show that $x = \sum_{g \in G} g$ is in the centre of $F[G]$ and $x^2 = 0$. Conclude that $I := F[G]x$ is a two-sided ideal of $F[G]$ with $I^2 = (0)$.*

So now we've seen that $F[G]$ is artinian and Maschke's theorem tells us that if the characteristic of F does not divide the order of G then $F[G]$ has no nonzero nil ideals and so Artin-Wedderburn says that

$$F[G] \cong \prod_{i=1}^{m} M_{n_i}(D_i).$$

Recall that if F is algebraically closed then the D_i must be F when $F[G]$ is finite-dimensional and so we can say more: if F is an algebraically closed field of characteristic zero then

$$F[G] \cong \prod_{i=1}^{s} M_{n_i}(F)$$

for some $s \geq 1$ and some $n_1, \ldots, n_s$. Moreover, this isomorphism sends $\lambda \in F$ to $(\lambda I_{n_1}, \ldots, \lambda I_{n_s})$. So most of the time we will take $F = \mathbb{C}$. Then we can say quite a bit about this decomposition.

Theorem 0.5. *Let G be a finite group. Then*

$$\mathbb{C}[G] \cong \prod_{i=1}^{s} M_{n_i}(\mathbb{C})$$

and we have the following:

(i) $\sum n_i^2 = |G|$;
(ii) *s is equal to the number of conjugacy classes of G;*
(iii) $\#\{i \colon n_i = 1\} = |G/G'|$;

(iv) $n_i||G|$ *for all* i.

Exercise 12. *Let G be a finite group. Show that G has at most $(|G|+3k)/4$ conjugacy classes, where $k = |G/G'|$. For which dihedral groups does equality hold? For which symmetric groups does equality hold? (Feel free to use the fact that A_n is simple for $n \geq 5$.)*

Exercise 13. *Let G be a finite group and let $\rho : G \to \mathrm{GL}(V)$ be a d-dimensional representation. Show that if ρ is injective then the trace of $\rho(g)$ has modulus at most d, with equality only if g is central.*

We have already shown the decomposition. Some of these facts are a bit hard to prove, but we can prove the first one right away. Notice that $M_n(\mathbb{C})$ is n^2-dimensional as a $\mathbb{C}$-vector space. So $\prod_{i=1}^s M_{n_i}(\mathbb{C})$ is $(\sum n_i^2)$-dimensional. On the other hand, $\mathbb{C}[G]$ is $|G|$-dimensional. To do the second one, we have to look at the centre of the group algebra.

Given a conjugacy class $\mathcal{C}$ of G, we define $z_{\mathcal{C}} := \sum_{g \in \mathcal{C}} g$. Then we claim that $z_{\mathcal{C}}$ is central in $F[G]$. To see this, observe that it suffices to show that $hz_{\mathcal{C}} = z_{\mathcal{C}}h$ for all $h \in F[G]$. Now $h\mathcal{C}h^{-1} = \mathcal{C}$ by definition, so $hz_{\mathcal{C}}h^{-1} = z_{\mathcal{C}}$, giving us the result.

Proposition 0.12. *The elements $z_{\mathcal{C}}$ form a basis for the centre of $F[G]$ as $\mathcal{C}$ runs over the conjugacy classes.*

Proof. Let $z = \sum a_g g$ be in the centre of $F[G]$. Then for all $h \in G$ we have

$$\sum a_g g = z = hzh^{-1} = \sum a_g hgh^{-1} = \sum a_{h^{-1}gh} g.$$

It follows that $a_g = a_{h^{-1}gh}$ for all $g, h \in G$. In particular, the map $g \mapsto a_g$ is constant on conjugacy classes of G. It follows that we can write $z = \sum_{\mathcal{C}} a_{\mathcal{C}} z_{\mathcal{C}}$. Hence we see that the elements span. Since they have disjoint supports, we see that they are linearly independent and the result follows. □

Corollary 0.13. *The dimension of the centre of $F[G]$ is equal to the number of conjugacy classes of G.*

Now given a ring R, we let $Z(R)$ denote its centre. Then

$$Z(R_1 \times R_2 \times \cdots \times R_s) \cong Z(R_1) \times \cdots \times Z(R_s).$$

In particular, since the centre of $M_{n_i}(\mathbb{C})$ is just $\mathbb{C}I_{n_i}$, we see that the centre of $\prod_{i=1}^s M_{n_i}(\mathbb{C})$ is just $\mathbb{C}^s$ and so it is s dimensional. This proves the second claim from the theorem. Let's make a few remarks.

Remark 0.14. Let G be an abelian group of size n. Then $\mathbb{C}[G] \cong \mathbb{C}^n$.

Well, we can only get 1×1 matrices since $\mathbb{C}[G]$ is commutative, right?

Remark 0.15. $\mathbb{C}[S_3] \cong \mathbb{C}^2 \times M_2(\mathbb{C})$.

This is easy now! Note that S_3 has three conjugacy classes. How many ways can you write 6 as a sum of three squares? Let's prove that $d := \#\{i \colon n_i = 1\} = |G/G'|$. To see this is a bit trickier. Let's first remark that if N is a normal subgroup of G then we have a reduction (F-algebra) homomorphism $F[G] \to F[G/N]$, in which all elements of N are sent to 1 in $F[G/N]$. Let us fix a $\mathbb{C}$-algebra isomorphism

$$\Phi : \mathbb{C}[G] \to R := \mathbb{C}^d \times \prod_j M_{n_j}(\mathbb{C}),$$

where the $n_j > 1$. We have a surjective map $f : R \to \mathbb{C}^d$ obtained by quotienting by the ideal $I = (0) \oplus \prod_j M_{n_j}(\mathbb{C})$. Then $\Psi := f \circ \Phi$ gives a map from $\mathbb{C}[G] \to \mathbb{C}^d$. Notice that $\Psi(ghg^{-1}h^{-1}) = 1$ since the right side is commutative. Since all commutators are sent to 1, this map Ψ factors through $F[G/G']$. That is, we have a map $\Psi' : \mathbb{C}[G/G'] \to \mathbb{C}^d$ and so $|G/G'| \geq d$. Now suppose that $|G/G'| > d$. Then $\mathbb{C}[G/G'] \cong \mathbb{C}^e$ for some $e > d$. But $\mathbb{C}[G/G']$ is a homomorphic image of $\mathbb{C}[G]$ and so $\mathbb{C}^e$ must be a homomorphic image of $\mathbb{C}^d \times \prod_j M_{n_j}(\mathbb{C})$. Notice that if $s : \mathbb{C}^d \times \prod_j M_{n_j}(\mathbb{C}) \to \mathbb{C}^e$ then each $(0) \times M_{n_j}(\mathbb{C}) \times (0)$ must map to 0, because $M_{n_j}(\mathbb{C})$ is noncommutative and simple. Thus the kernel of s contains all of $(0) \times \prod_j M_{n_j}(\mathbb{C})$ and so s induces a surjective $\mathbb{C}$-algebra map $\mathbb{C}^d \to \mathbb{C}^e$. But this is impossible with $d < e$. The result now follows. The last piece, showing that each n_i divides the order of G is hardest. We'll have to develop some character theory to do this.

Exercise 14. *Let C_i be the cyclic group of order i. Show that $\mathbb{C}[C_4] \cong \mathbb{C}[C_2 \times C_2]$ but show that $\mathbb{Q}[C_4] \ncong \mathbb{Q}[C_2 \times C_2]$.*

Exercise 15. *Find two non-abelian groups G, H such that $G \ncong H$ but $\mathbb{C}[G] \cong \mathbb{C}[H]$.*

Characters and representations

We can genuinely start representation theory at this point. But we now have the advantage of knowing lots of structure theory about the group algebra in advance. Let G be a group. A *representation* of G is a group homomorphism $\phi : G \to \mathrm{GL}_n(F)$ for some $n \geq 1$ and some field F. There are two representations everyone (and I mean everyone!) should know. The first is the trivial representation. This is the map $\phi : G \to F^* = \mathrm{GL}_1(F)$ that sends every element $g \to 1$. The second is the *left regular representation*. For this one, let $V = F[G]$. Then $\mathrm{GL}(V)$, the collection of F-linear isomorphisms of V, is isomorphic to $\mathrm{GL}_n(F)$, where $n = |G|$. Then each $g \in G$ gives a linear map $\phi_g : F[G] \to F[G]$ via left multiplication. If we pick the basis $\{g : g \in G\}$ for V, then we can then get an explicit representation of G in terms of permutation matrices.

Example 0.16. Let Q be the quaternion group $\{\pm 1, \pm i, \pm j, \pm k\}$. Then we have a representation $Q \to \mathrm{GL}_2(\mathbb{C})$.

(This is given in class.) We will often find it easier to work with a vector space V and work with $\mathrm{GL}(V)$ rather than with $\mathrm{GL}_n(F)$. The advantage here is that V becomes a left $F[G]$-module via the rule $(\sum a_g g) \cdot v = \sum a_g (g \cdot v)$ and we do not have to worry about bases. If $\phi : G \to \mathrm{GL}(V)$ is a representation of G then $\mathrm{Tr}(\phi(1)) = \dim(V)$ is called the *degree* of the representation ϕ. We will sometimes call V a representation of G, too—that's just how life is. So we will often adopt two points of view. We make this as a remark right now.

Remark 0.17. We have a correspondence between F-representations $\phi : G \to V$ and left $F[G]$-modules, given by taking our representation V and making it an $F[G]$-module by taking the action of G on V induced by ϕ and extending via F-linearity. In the other direction, we take an $F[G]$-module, and we define $\phi(g)$ to be the endomorphism from V to V given by left multiplication by g. This is an isomorphism since g is invertible in $F[G]$.

Notice that if $W \subseteq V$ is a subspace that has the property that $\phi(g)(w) \in W$ for all $g \in G$ and all $w \in W$ then restriction of ϕ to W gives a representation $\psi : G \to \mathrm{GL}(W)$. In this case, we call W a *subrepresentation* of V. We can also form a quotient representation $\omega : G \to \mathrm{GL}(V/W)$ defined by $\omega(g)(v + W) = \phi(g)(v) + W$. This is well-defined since W is invariant under the action of G. In the other point of view, W is a subrepresentation if it is a nonzero $F[G]$-submodule of V and we can form the quotient module V/W.

Given two representations $\rho_i : G \to \mathrm{GL}(V_i)$ for $i = 1, 2$, we say that V_1 is isomorphic to V_2 (as representations) if there is a linear isomorphism $f : V_1 \to V_2$ such that $\rho_1(v_1) = \rho_2(f(v_1))$ for all $v_1 \in V_1$. Then $f^{-1} : V_2 \to V_1$ is the inverse of f.

Let's see an example. Notice that if $V = \mathbb{C}e_1 + \mathbb{C}e_2 + \mathbb{C}e_3 \cong \mathbb{C}^3$, then we have a map $\phi : S_3 \to \mathrm{GL}(V)$ by declaring that $\sigma(e_i) = e_{\sigma(i)}$. Then $\tau \circ \sigma(e_i) = \tau(e_{\sigma(i)}) = e_{\tau(\sigma(i)}$, so, yes, it works. Now let W be the span of $e_1 + e_2 + e_3$. Then W is a subrepresentation. Notice it is naturally isomorphic to the trivial representation. Also V/W is a 2-dimensional representation of S_3.

Exercise 16. *Let $\phi : G \to \mathrm{GL}_d(F)$ be a representation of G. Show that ϕ extends to an F-algebra homomorphism (not necessarily onto) from $F[G]$ to $M_d(F)$.*

Exercise 17. *Use the preceding exercise and the example of the 2-dimensional representation of S_3 above to find an explicit isomorphism from $\mathbb{C}[S_3]$ to $\mathbb{C}^2 \times M_2(\mathbb{C})$. (The maps to $\mathbb{C}^2$ come from the trivial representation and the 1-dimensional representation that sends g to $\mathrm{sgn}(g)$ for $g \in S_3$.)*

If $\rho_i : G \to \mathrm{GL}(V_i)$ are representations for $i = 1, 2$ with V_1, V_2 vector spaces over F, we can make a new representation $\rho = (\rho_1 \oplus \rho_2) : G \to V_1 \oplus V_2$ with $\rho(g)(v_1, v_2) = (\rho_1(g)(v_1), \rho_2(g)(v_2))$. This is easily checked to be a representation and is called the *direct sum* of the representations V_1 and V_2.

There is one other nice way of constructing representations we should get out of the way: the tensor product. If $\rho_i : G \to \mathrm{GL}(V_i)$ are representations for $i = 1, 2$ with V_1, V_2 vector spaces over F, we can make a new space $V = V_1 \otimes_F V_2$ as follows. Let $\{b_i\}_{i=1}^d$ and $\{c_j\}_{j=1}^e$ be bases for V_1 and V_2 respectively. Then V has dimension de and a basis given by the symbols $b_i \otimes c_j$. The representation is defined as follows. If $g \in G$ then $\rho_1(g)(b_i) = \sum \alpha_{g,i} b_i$ and $\rho_2(g)(c_j) = \sum \beta_{g,j} c_j$. We then define $\rho = \rho_1 \otimes \rho_2(g)(b_i \otimes c_j) = \sum_i \sum_j \alpha_{g,i} \beta_{g,j} b_i \otimes c_j$ and extend via linearity. This is called the tensor product of the representations V_1 and V_2.

Exercise 18. *Show that the tensor product representation $\rho_1 \otimes \rho_2$ of two representations $\rho_i : G \to \mathrm{GL}(V_i)$ over a field F is indeed a group homomorphism from G to $\mathrm{GL}(V_1 \otimes_F V_2)$.*

Definition 0.18. *A representation $\rho : G \to \mathrm{GL}(V)$ is* irreducible *if the only G-invariant subspaces $W \subseteq V$ are $W = (0)$ and $W = V$. This is equivalent to saying that V is a simple $F[G]$-module (i.e., it is nonzero and its only submodules are itself and (0)).*

Semisimplicity

A ring R is called *semisimple* if R, when regarded as a left R-module, is a direct sum of simple submodules. Notice that it is generally pretty hard to be semisimple. For example, $\mathbb{Z}$ doesn't even have *any* simple submodules, because every nonzero ideals properly contains another nonzero ideal. If D is a division ring, then D is already a simple D module, so D is semisimple (in fact simple) as a D-module. More generally, $M_n(D)$ is semisimple. To see this, recall that the column space $D^{n \times 1}$ is a simple module: if you give me a nonzero vector in $D^{n \times 1}$, we can find a transformation that sends it to any other vector you want. (Exercise for you!) Then $M_n(D) = M_1 \oplus \cdots M_n$, where M_i is the matrices whose only nonzero entries are in the i-th column. Then each $M_i \cong D^{n \times 1}$ is simple, so we get semisimplicity. More generally, if $R_1, \ldots, R_s$ are semisimple then $R_1 \times \cdots \times R_s$ is semisimple too. To do this, it suffices to prove it when $s = 2$ and then use induction. So if R_1, R_2 are semisimple, then we have $R_1 = M_1 \oplus \cdots \oplus M_d$ and $R_2 = N_1 \oplus \cdots \oplus N_e$, where the M_i and N_j are simple R_1- and R_2-modules. Notice that each M_i can be given a structure as an $R_1 \times R_2$ module, via the rule $(r_1, r_2) \cdot m = r_1 m$ and similarly, we have $(r_1, r_2) \cdot n = r_2 n$ gives each N_j a structure as an $R_1 \times R_2$-module. Moreover, the $R_1 \times R_2$-submodules of each M_i are just the R_1-submodules of M_i, so they are all simple. But now we have

$$R_1 \times R_2 = M_1 \oplus \cdots M_n \oplus N_1 \oplus \cdots \oplus N_2$$

and this is an isomorphism of $R_1 \times R_2$-modules. It follows that an Artinian ring with no nonzero nil ideals is a semisimple ring by the remarks above and the Artin-Wedderburn theorem. In fact, if

$$R = \prod_{i=1}^{s} M_{n_i}(D_i)$$

and we let M_i denote the simple R-module induced from the column vectors $D_i^{n_i \times 1}$ for $i = 1, 2, \ldots, s$ as above then we have just shown that we have

$$R \cong M_1^{n_1} \oplus M_2^{n_2} \oplus \cdots \oplus M_s^{n_s}$$

as an R-module.

In fact, we have the even stronger result for semisimple rings.

Theorem 0.6. *Let R be a semisimple ring. Then every left R-module is a direct sum of simple modules.*

To prove this, we first require a lemma.

Lemma 0.19. *If R is a ring and M is a left R-module that can be written as a sum of simple submodules then M is equal to a direct sum of simple submodules.*

So this statement is saying that if we can write things as a sum of simples, then we can refine things and get it as a direct sum decomposition.

Proof. Write $M = \sum_{i \in I} M_i$, where I is some index set and the M_i are simple sub modules. Now let S denote the collection of subsets J of I for which $\sum_{i \in J} M_i$ is direct. Notice S is non-empty, because each singleton of I is in S. We order S by inclusion and we observe that S is closed under unions of chains. This means by Zorn's lemma we can find a subset J of I that is maximal in S. We claim that $M = \bigoplus_{i \in J} M_i =: N$. To see this, suppose that it is not the case. Then there is some $i \in I$ such that $M_i \not\subseteq N$. Then $M_i \cap N = (0)$ since M_i is simple and so $M_i \oplus N$ is direct, contradicting maximality of J. $\square$

As a corollary, we see that if M is equal to a direct sum of simple submodules then so is every quotient of M.

Exercise 19. *Show that if M is a left R-module and is equal to a direct sum of simple submodules then so is every quotient of M.*

Now we see that if R is semisimple then every left R-module is semisimple (i.e., a direct sum of simple submodules). To see this, we let M be a left R-module and let $\{x_\alpha\}_{\alpha \in I}$ be a generating set for M. Then we have a surjection $R^I = \oplus_{i \in I} Re_i \to M$ which sends $e_i \to x_i$. By the first isomorphism theorem, $M \cong R^I/L$, where L is the kernel. So M is a quotient of a free module. By the exercise it suffices to prove that free R-modules are semisimple. But since R is semisimple, we have $R = M_1 \oplus \cdots M_s$. Now we just take $R^I \cong \oplus_{i=I}(M_1 \oplus \cdots \oplus M_s)$, which is a direct sum of simple modules.

Corollary 0.20. *Let F be a field of characteristic zero or of characteristic $p > 0$ prime to the order of G. Then every $F[G]$-module is semisimple. In particular, every representation V of G (over F) is isomorphic to a direct sum of irreducible representations (over F).*

Of course, one might worry about uniqueness. How do we know we cannot have different decompositions of a module as a direct sum of simple submodules. As it turns out, this cannot occur. We only prove this in the finite direct sum case, since that is all we are interested in.

Proposition 0.21. *Suppose that $M = M_1 \oplus \cdots M_s = N_1 \oplus \cdots \oplus N_t$ where the M_i and N_j are simple submodules of M. Then $s = t$ and after permuting indices we have $M_i \cong N_i$ as left R-modules for $i = 1, \ldots, s$.*

Proof. Suppose that we have two decompositions into simples $M = M_1 \oplus \cdots M_s = N_1 \oplus \cdots \oplus N_t$. We'll prove this by induction on the min of s and t. If the min is 1 then M is simple and there is nothing to prove. Let's suppose that the min is at least 2 and it holds for all smaller mins. Then for each i, we let $P_i = \sum_{j \neq i} M_j$. Notice that $M/P_i \cong M_i$ and so P_i is a maximal proper submodule of M by correspondence. Moreover, the intersection of the P_i is nonzero. In particular, there is some i such that $N_t \not\subseteq P_i$. Since N_t is simple we then have $N_t \cap P_i = (0)$ and so $N_t + P_i$ is direct. Since P_i is maximal, we see $M = P_i \oplus N_t$. Thus $N_t \cong M/P_i \cong M_i$. By relabelling, we may assume that $i = s$ and so $N_t \cong M_s$. But now $N_1 \oplus \cdots \oplus N_{t-1} \cong M/N_t = P_s = M_1 \oplus \cdots \oplus M_{s-1}$. The result follows by induction. $\square$

We can now get a handle on all simple modules for a product of matrix rings over division rings.

Theorem 0.7. *Let $R = \prod_{i=1}^{s} M_{n_i}(D_i)$. Then let M_i denote the simple R-module induced from the column vectors $D_i^{n_i \times 1}$ for $i = 1, 2, \ldots, s$. Then every simple module of R is isomorphic to some M_i and*

$$R \cong M_1^{n_1} \oplus M_2^{n_2} \oplus \cdots \oplus M_s^{n_s}.$$

We have already shown

$$R \cong M_1^{n_1} \oplus M_2^{n_2} \oplus \cdots \oplus M_s^{n_s}.$$

Now let N be a simple left R-module. One can show N is isomorphic to one of the M_i quickly and directly, but let's take a leisurely route with the advantage that we learn about Hom sets.

Given two R-modules M and N, we let $\mathrm{Hom}_R(M,N)$ denote the set of R-module homomorphisms from M to N. In the case that $M=N$ then $\mathrm{Hom}_R(M,N)$ is often denoted $\mathrm{End}_R(M)$, End standing for endomorphisms, and this becomes a ring with multiplication being given by composition, just as $\mathrm{End}_F(V)\cong M_d(F)$ when V is a d-dimensional F-vector space. When R is understood, we will often suppress it. Notice that $\mathrm{Hom}_R(M,N)$ is itself a left R-module, given by $r\cdot f(m)=f(rm)$ for $f\in\mathrm{Hom}(M,N)$ and $r\in R$. With this we have the following routine facts.

Exercise 20. *Show that if M, M_α, $\alpha\in I$, and N_1 and N are R-modules and $N_1\cong N$ then $\mathrm{Hom}_R(M,N)\cong\mathrm{Hom}_R(M,N_1)$ and similarly if the M and N_1, N are interchanged. Show that*

$$\mathrm{Hom}_R(\bigoplus_{\alpha\in I}M_\alpha,M)\cong\prod_{\alpha\in I}\mathrm{Hom}(M_\alpha,M).$$

Lemma 0.22. *Let M and N be two simple left R-modules. Then $\mathrm{Hom}_R(M,N)=(0)$ if $M\ncong N$. If M is simple $\mathrm{End}_R(M)$ is a division ring.*

Proof. Suppose we have a nonzero $f:M\to N$ then the image of f is a nonzero submodule of N and since N is simple, we see that f is onto. The kernel of f is a submodule of M and since f is nonzero it is proper and since M is simple, we see that the kernel is trivial so f is 1-to-1. Thus if $M\ncong N$ then the set of homomorphisms is trivial. Now if $M=N$ then we just showed that every nonzero homomorphism has an inverse and so $\mathrm{End}_R(M)$ is a division ring. □

Now we are ready to finish the proof. If N is a simple R-module, then $\mathrm{Hom}(R,N)$ is nonzero since $N\cong R/I$ for some maximal left ideal I so reduction mod I gives us a homomorphism. On the other hand,

$$\mathrm{Hom}(R,N)=\mathrm{Hom}(\oplus M_i^{n_i},N)\cong\bigoplus_i\mathrm{Hom}(M_i,N)^{n_i}.$$

By Schur we know that $M_i\cong N$ for some i since the Hom set is nonzero. (Recall that direct sums and direct products are really the same thing for finite index sets.)

Corollary 0.23. *$\mathbb{C}[G]$ has only s irreducible representations up to isomorphism, where $s=\#$ of conjugacy classes of G.*

We have $\mathbb{C}[G]\cong\prod_{i=1}^s M_{n_i}(\mathbb{C})$. We have seen that there are exactly s different simple modules up to isomorphism and since simple $\mathbb{C}[G]$-modules correspond to irreducible representations we have the result. Notice also that the representations themselves come from the isomorphism

$$\phi:\mathbb{C}[G]\cong\prod_{i=1}^s M_{n_i}(\mathbb{C})$$

and the projection maps $\pi_i:\prod_{j=1}^s M_{n_j}(\mathbb{C})\to M_{n_i}(\mathbb{C})$. Then $\pi_i\circ\phi$ restricted to G gives an irreducible representation of G into $\mathrm{GL}_{n_i}(\mathbb{C})$ of degree n_i. Thus the degrees of the irreducible representations are $n_1,n_2,\ldots,n_s$. So we have that

$$|G|=\sum n_i^2,$$

where $n_1,\ldots,n_s$ are the degrees of the irreducible representations. So to sum up so far, we have shown the following about complex representations of G:

(i) Every representation is a direct sum of irreducibles and this decomposition is unique;
(ii) There are only finitely many distinct irreducible representations (up to isomorphism) and this is the number of conjugacy classes;
(iii) There are $|G/G'|$ irreducible representations of degree 1;
(iv) The sum of the squares of the degrees of the irreducible representations is $|G|$;
(v) The left regular representation of G decomposes as a sum of n_i copies of the i-th irreducible representation.

The last statement follows from the fact that the left regular representation comes from looking at $\mathbb{C}[G]$ as a module over itself and we have shown how to decompose it as a sum of simples.

Characters

We've already developed an impressive amount of theory, but now we're going to continue and study characters. Then we'll have really laid the foundations of representation theory and we can start looking to apply it. For now, we will restrict our attention to complex representations; i.e., the case when our field $F=\mathbb{C}$. Given a representation $\rho:G\to\mathrm{GL}_n(\mathbb{C})$, we define the *character* $\chi_\rho(g)$ to be the Trace of $\rho(g)$. We call χ_ρ the *character* associated to the representation ρ. Since Trace is a similarity invariant, it does not depend upon our choice of basis and so we can make sense of the character when we just have a map $G\to\mathrm{GL}(V)$ by picking whichever bases we like and computing the trace. Notice that a character is a *class function*; that is, it takes the same value on a conjugacy class. The reason for this is that if $g,h\in G$ then $\rho(hgh^{-1})=\rho(h)\rho(g)\rho(h)^{-1}$ and so $\rho(hgh^{-1})$ is similar to $\rho(g)$ and so they have the same trace. Thus the character of a representation is completely determined by what it does on a set of conjugacy class representatives. Notice that if $\rho_i:G\to\mathrm{GL}(V_i)$ for $i=1,2$ then $\rho_1\oplus\rho_2$ maps to $\mathrm{GL}(V_1\oplus V_2)$ and $\chi_{\rho_1+\rho_2}(g)=\chi_{\rho_1}(g)+\chi_{\rho_2}(g)$. In other words, if V is a complex representation and we write $V=\bigoplus V_i^{n_i}$ as a direct sum of irreducible representations then the character associated to the character V is $\sum n_i\chi_i$, where χ_i is the character associated to V_i. Hence it suffices to understand the characters of irreducible representations. As it turns out there is a tremendous amount of structure here.

Orthogonality relations

Let ϕ, ψ be two complex-valued functions on a finite group G. Then we define an "inner product"

$$\langle \phi, \psi \rangle := \frac{1}{|G|} \sum_{g \in G} \phi(g)\psi(g^{-1}).$$

In the case when we have a compact Hausdorff group, we can still make sense of this inner product by using an integral with Haar measure. Notice this inner product is bilinear and symmetric since we can make a change of variables sending $g \mapsto g^{-1}$. Given a matrix M we let $M_{i,j}$ denote the (i,j)-entry of M.

Lemma 0.24. *Let $\rho_1 : G \to \mathrm{GL}(V)$ and $\rho_2 : G \to \mathrm{GL}(W)$ be two distinct irreducible representations of G. Then if we fix a basis and regard $\rho_i(g)$ as being matrices then we have $\langle (\rho_1)_{i,j}, (\rho_2)_{k,\ell} \rangle = 0$. If ρ is irreducible, then $\langle \rho_{i,j}, \rho_{k,\ell} \rangle = 0$ unless $i = \ell$ and $j = k$ in which case the inner product is $1/d$, where d is the degree of ρ.*

Proof. Let $T : V \to W$ be a linear map. Then we can form a new "averaging" linear map $S : V \to W$ as follows

$$S(v) = \frac{1}{|G|} \sum_{g \in G} \rho_2(g)(T(\rho_1(g^{-1})v)).$$

Notice that for $h \in G$,

$$S(h \cdot v) = S(\rho_1(h)v) = \frac{1}{|G|} \sum_{g \in G} \rho_2(g)(T\rho_1(g^{-1}\rho_1(h)v)).$$

Now we let $u = g^{-1}h$. Then this becomes

$$\frac{1}{|G|} \sum_{u \in G} \rho_2(hu^{-1})T\rho_1(h)v = \rho_2(h) \cdot S(v) = h \cdot S(v).$$

Thus S is actually a homomorphism of $\mathbb{C}[G]$-modules. But Schur's lemma tells us that S must be zero unless $V \cong W$. Now if we regard T as being a matrix in our bases, this tells us that

$$\sum_{g \in G} \rho_2(g)T\rho_1(g)$$

for all matrices $T : V \to W$. In particular, if we let $d = \dim(V)$ and $e = \dim(W)$ and fix bases $v_1, \ldots, v_d$ and $w_1, \ldots, w_e$ for V and W then if we take T to be the matrix $E_{i,j}$ (i.e., it sends v_j to w_i and sends all other v_k to zero) then if we take the (k,ℓ)-entry we see that

$$0 = \sum_g \rho_2(g)_{k,i}\rho_1(g^{-1})_{j,\ell}$$

This gives the first part. For the second part, we take $V = W$ and we'll let ρ denote our map. Recall that if $V = W$ then Schur's lemma tells us that $\mathrm{End}_{\mathbb{C}[G]}(V)$ is a division ring. Notice it must be $\mathbb{C}$ because it is finite-dimensional as a $\mathbb{C}$-vector space. Thus if we start with T, we see that S becomes a scalar multiple of the identity. By taking traces, we see that S and T have the same trace. Thus if we take $T = E_{i,j}$ we see that S becomes 0 if $i \neq j$ so

$$0 = \sum_g \rho(g)_{k,i}\rho(g^{-1})_{j,\ell}$$

if $i \neq j$. If $i = j$ then S becomes $1/d\mathrm{I}$ where d is the degree of the representation so we see

$$1/d\mathrm{I} = \sum_g \rho(g)E_{i,i}\rho(g^{-1})$$

and we get the remaining result by taking entries. □

Corollary 0.25. *(Orthogonality of characters) If χ_1, χ_2 are two characters of irreducible representations then $\langle \chi_1, \chi_2 \rangle = 0$ if the representations are distinct and is 1 otherwise.*

Proof. Let ρ_1, ρ_2 denote the correspond representations. Then $\chi_1 = \sum(\rho_1)_{i,i}$ and $\chi_2 = \sum(\rho_2)_{i,i}$. If the characters are distinct this follows from the lemma. If $\rho_1 = \rho_2 = \rho$, then since $\langle \rho_{i,i}, \rho_{j,j} \rangle = \delta_{i,j}1/d$, we see that $\langle \rho_1, \rho_2 \rangle = \sum_{i=1}^d 1/d = 1$. And we're done! □

As a corollary, we see that if V is a representation and we decompose it as a sum of distinct irreducibles:

$$V = \oplus_{i=1}^s V_i^{m_i}$$

then $\chi_V = \sum_{i=1}^s m_i\chi_{V_i}$. So

$$\langle \chi_V, \chi_V \rangle = \sum_{i=1}^s \sum_{j=1}^s m_i m_j \delta_{i,j} = m_1^2 + \cdots + m_s^2.$$

Hence V is irreducible if and only if $||\chi_V||^2 = \langle \chi_V, \chi_V \rangle = 1$; otherwise the norm is strictly greater than 1. We can also find the decomposition: if $\langle \chi, \chi_i \rangle = m_i$ then the representation contains m_i copies of V_i.

Corollary 0.26. *Let ρ be the left regular representation of G; that is, $\rho : G \to \mathrm{GL}(V)$, where $V = \mathbb{C}[G]$ and $\rho(g)$ is left multiplication by g. Then if χ is the character of ρ then $\chi(g) = |G|\delta_{g,1}$ and $\chi = \sum_i n_i\chi_i$, where χ_i ranges over the irreducible characters of G and n_i is the degree of χ_i.*

Proof. The fact that $\chi(g) = |G|\delta_{g,1}$ comes from looking at the matrix $\rho(g)$ in the basis given by the elements of G. If we use the group algebra representation we can see this too, but let's use the orthogonality relations. Notice that

$$\langle \chi, \chi_i \rangle = 1/|G| \cdot \sum_g \chi(g)\chi_i(g^{-1}) = \chi_i(1) = n_i,$$

and so χ contains n_i copies of χ_i. □

Character tables

Given a group, a lot of information can be obtained from its character table. This is just an array that lists the distinct irreducible characters along with their values on a set of conjugacy class representatives. In fact, it is often very easy to compute the character table for a small group just using basic facts we have already developed. Let's do one for S_3. Recall that S_3 has three inequivalent irreducible representations of degrees 1, 1, 2. We know that one degree 1 representation is the trivial one; the other is called the alternating representation and uses the sign function—let's call the corresponding characters χ_1 and χ_2. We'll call the degree two representation χ_3. Also, S_3 has three conjugacy classes—we pick representatives id, (12), and (123). So far we know

	id	(12)	(123)
χ_1	1	1	1
χ_2	1	-1	1
χ_3	a	b	c

for some a, b, c. Now $\chi(1)$ is just the degree of the representative since 1 gets sent to the identity and the trace of the identity is the degree. So $a = \chi_3(1) = 2$. Notice that the conjugacy class (12) has 3 elements and the class of (123) has 2 elements. So $0 = \langle \chi_3, \chi_1 \rangle = 1/6 \cdot (a + 3b + 2c)$. So we get $a + 3b + 2c = 0$ or $3b + 2c = -2$. Similarly, with χ_2 we get $a - 3b + 2c = 0$ so $-3b + 2c = -2$ so $b = 0$ and $c = -1$ and the character table is

	id	(12)	(123)
χ_1	1	1	1
χ_2	1	-1	1
χ_3	2	0	-1

Dare we try S_4? Yes—this is easier with a bit more structure theory but we can still handle it. So S_4 has five conjugacy classes [id], of size 1, [(12)] of size 6, [(123)] of size 8; [(12)(34)] has size 3; [(1234)] has size 6. So we have five irreducible characters. The trivial and the alternating are the degree one characters. Let a, b, c denote the degrees of χ_3, χ_4, χ_5. Then $1^2 + 1^2 + a^2 + b^2 + c^2 = 24$, so we see that $a = 2$, $b = c = 3$ is the only solution. So $\chi_3(1) = 2$, $\chi_4(1) = \chi_5(1) = 3$. So far we have

	id	(12)	(123)	(12)(34)	(1234)
χ_1	1	1	1	1	1
χ_2	1	-1	1	1	-1
χ_3	2	x	y	z	w
χ_4	3	a	b	c	d
χ_5	3	e	f	g	h

Now to help us out, let's make one exercise.

Exercise 21. *Let $\rho : G \to \mathrm{GL}(V)$ be an irreducible representation and $\phi : G \to \mathbb{C}^* = \mathrm{GL}(\mathbb{C})$ be a 1-dimensional representation. Show that $\phi \otimes \rho$ is an irreducible representation of G of the same degree as ρ and its character at g is just $\chi_\rho(g)\phi(g)$.*

Notice that by the exercise $\chi_3\chi_2$ is an irreducible character of degree 2 and so $\chi_3\chi_2 = \chi_3$. In particular, $x = w = 0$. Also $\chi_4\chi_2$ is irreducible of degree 3 so either $\chi_4\chi_2 = \chi_5$ or it is χ_4. We can't have the latter since that would give $a = d = e = h = 0$ and this would be impossible with the orthogonality relations. So $\chi_4\chi_2 = \chi_5$ which means $a = -e$, $b = f$, $c = g$, $d = -h$. So let's clean it up!

	id	(12)	(123)	(12)(34)	(1234)
χ_1	1	1	1	1	1
χ_2	1	-1	1	1	-1
χ_3	2	0	y	z	0
χ_4	3	a	b	c	d
χ_5	3	-a	b	c	-d

In particular, $x = w = 0$. Notice that the inner product of χ_3 and χ_1 being zero gives: $2 + 8y + 3z = 0$. The inner product of χ_3 with itself then gives

$$24 = 4 + 8y^2 + 3z^2 = 4 + 8y^2 + 3(8y + 2)^2/9.$$

This equation has roots $y = -1$ and $y = 7/11$. We'll see later that if the trace is rational then it has to be an integer, so let's use this fact and we get $y = -1$ and $z = 2$. Character tables are like sudokus—as you fill in more stuff, it becomes successively easier to complete. So now the orthogonality of χ_3 and χ_4 gives $0 = 6 - 8b + 6c$ and orthogonality of χ_4 with the sum of χ_1 and χ_2 gives $0 = 3 + 8b + 3c = 0$. So we solve and we get $c = -1$ and $b = 0$. So now we have

	id	(12)	(123)	(12)(34)	(1234)
χ_1	1	1	1	1	1
χ_2	1	-1	1	1	-1
χ_3	2	0	-1	2	0
χ_4	3	a	0	-1	d
χ_5	3	-a	0	-1	-d

The inner product of χ_4 with χ_2 gives $0 = 3 - 6a - 3 - 6d$, so $a = -d$. Taking the norm gives $24 = 9 + 6a^2 + 3 + 6a^2$ so $a = \pm 1$. Which one do we take? It doesn't matter—one gives χ_4 and one gives χ_5, so now we finally have it!

	id	(12)	(123)	(12)(34)	(1234)
χ_1	1	1	1	1	1
χ_2	1	-1	1	1	-1
χ_3	2	0	-1	2	0
χ_4	3	1	0	-1	-1
χ_5	3	-1	0	-1	1

Exercise 22. *Compute the character table for A_4.*

Exercise 23. *Compute the character table for the dihedral group of size* 14.

Algebraic integers and characters

One of the things we used in finding the character table of S_4 was that if a character takes a rational value, then it must in fact be an integer. We will prove this now, by developing the theory of algebraic integers. A number $\alpha \in \mathbb{C}$ is called an algebraic integer if it is a root of a monic polynomial with integer coefficients. Obviously, the integers are algebraic integers. After all, n is a root of $x - n$. Notice that $\sqrt{2}$ is an algebraic integer. And for us, it is important to note that $\exp(2\pi i/n)$ is an algebraic integer since it satisfies $x^n - 1 = 0$. Notice that if α is a rational number then it is an algebraic integer if and only if α is an integer. To see this, suppose that $P(\alpha) = 0$ with $P(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_0$. Write $\alpha = a/b$ with $\gcd(a, b) = 1$, $b > 0$. Then $P(a/b) = 0$ gives $a^n + a_{n-1}a^{n-1}b + \cdots + a_0 b^n$. In particular, b divides a^n and so $b = 1$. We will now show an important fact about algebraic integers: they form a ring! To do this, we need a lemma.

Lemma 0.27. *Let $\alpha \in \mathbb{C}$. Then α is an algebraic integer if and only if there is a nonzero finitely generated $\mathbb{Z}$-submodule M of $\mathbb{C}$ such that $\alpha M \subseteq M$.*

Proof. If α is an algebraic integer, then we have $\alpha^n \in \sum_{i<n} \mathbb{Z}\alpha^i =: M$ for some n and so $\alpha M \subseteq M$. For the converse, suppose that M is a finitely generated submodule of $\mathbb{C}$. Since it is torsion-free, we see it is a free module and has basis $c_1, \ldots, c_m$ for some m. Then $\alpha c_i = \sum_j m_{i,j} c_j$ for some integers $m_{i,j}$. In other words, $\alpha \mathbf{c} = T\mathbf{c}$ where $\mathbf{c}$ is the vector $[c_1, \ldots, c_m]^T$ and T is the matrix $(m_{i,j})$. Now by the Cayley-Hamilton theorem $P(T) = 0$, where P is the characteristic polynomial of T (monic, integer coefficients). Thus $0 = P(T)\mathbf{c} = P(\alpha)\mathbf{c}$. Since $\mathbb{C}$ is a domain, we then see $P(\alpha) = 0$ and the result follows. We have a map given by $x \mapsto T(x)$ as before. Then x satisfies its char poly by CH so we are done. □

Theorem 0.8. *The set of algebraic integers forms a subring of $\mathbb{C}$.*

Proof. If a, b are algebraic integers of degrees m and n, let $M = \sum \mathbb{Z}a^i b^j$ with $i < m$ and $j < n$. Then we see that $\alpha M \subseteq M$ and $\beta M \subseteq M$ and so $(\alpha + \beta)M \subseteq M$ and $\alpha\beta M \subseteq M$ and so $\alpha + \beta$ and $\alpha\beta$ are algebraic integers. □

Corollary 0.28. *Let χ be a character. Then $\chi(g)$ is an algebraic integer. In particular, if $\chi(g)$ is rational then it is an integer.*

Proof. We have $g^n = 1$ for some n. Thus $A = \rho(g)$ is a matrix satisfying $A^n = I$. Hence its eigenvalues are roots of unity and so algebraic integers. But the trace is the sum of the eigenvalues and so it is an algebraic integer too. □

Exercise 24. *Suppose that a, b are integers and $\tan(2\pi a/b)$ is rational (and defined). What values can it take?*

Recall that in $\mathbb{C}[G]$ we have the central elements $z_{\mathcal{C}}$ corresponding to the conjugacy class $\mathcal{C}$. This is just the sum of the elements in $\mathcal{C}$. We have a map $\rho_i := \pi_i \circ \phi : \mathbb{C}[G] \to M_{n_i}(\mathbb{C})$ coming from the Artin-Wedderburn theorem. Since $z_{\mathcal{C}}$ is central, we see that $\rho_i(z_{\mathcal{C}}) = \alpha \mathbf{I}_{n_i}$. In particular, the trace of $\rho(z_{\mathcal{C}})$ is just $n_i\alpha$. On the other hand, it is $|\mathcal{C}|\chi_i(g)$ where g is a representative of the conjugacy class, since all elements of the conjugacy class will have the same trace. Thus $\alpha = |\mathcal{C}|\chi_i(g)/n_i$. Now we claim that α is an algebraic integer. Thus once we prove this claim, we have proved the following result.

Theorem 0.9. *If $g \in \mathcal{C}$, $\mathcal{C}$ a conjugacy class. Then for an irreducible character χ, we have*

$$\chi(g)|\mathcal{C}|/d$$

is an algebraic integer, where d is the degree of χ.

Let's show that α is an algebraic integer. Notice that if we have s conjugacy classes and $z_1, \dots, z_s$ are the central elements of the group algebra obtained by adding together all group elements in a conjugacy class, then we showed that the z_i form an $\mathbb{C}$-basis for the centre of $\mathbb{C}[G]$. Now $z_i z_j = \sum c_{i,j,k} z_k$. Notice that if we multiply $z_i z_j$ then all the coefficients of group elements have nonnegative integer coefficients. Since the supports of the z_k are disjoint, we then see that the $c_{i,j,k}$ must be integers. That is, if we let $M = \mathbb{Z}z_1 + \cdots + \mathbb{Z}z_s \subseteq \mathbb{C}[G]$, then we have $z_i M \subseteq M$. Now let $M' = \rho_i(M)$. Then since the z_j are central, we see that M' is a $\mathbb{Z}$-submodule of $M_{n_i}(\mathbb{C})$ consisting of scalar matrices. We thus identify M' with a finitely generated $\mathbb{Z}$-submodule of $\mathbb{C}$ (just identifying scalar matrices with the scalars). Also, since ρ_i is a homomorphism, we have $\rho_i(z_j)M' \subseteq M'$ and since M' is a finitely generated $\mathbb{Z}$-submodule of $\mathbb{C}$ that contains each $\rho_i(z_j)$, we see that the $\rho_i(z_j)$ are of the form $\alpha \mathbf{I}_{n_i}$ with α an algebraic integer, as claimed.

Now we're ready for something amazing. Notice that if χ_i is an irreducible character, then we have $\langle \chi_i, \chi_i \rangle = 1$. In other words,

$$\sum_{g \in G} \chi_i(g)\chi_i(g^{-1}) = |G|.$$

Now χ_i is constant on conjugacy classes. Let $\mathcal{C}_1, \dots, \mathcal{C}_s$ denote the conjugacy classes of G. Then we can rewrite this as

$$\sum_{j=1}^{s} \sum_{g \in \mathcal{C}_j} \chi_i(g)\chi_i(g^{-1}) = |G|.$$

If g and h are conjugate then so are g^{-1} and h^{-1} and so since χ_i is constant on conjugacy classes, we see that we can write this as

$$\sum_{j=1}^{s} |\mathcal{C}_j| \chi_i(g_j)\chi_i(g_j^{-1}) = |G|,$$

where we pick conjugacy class representatives $g_1, \dots, g_s$ for $\mathcal{C}_1, \dots, \mathcal{C}_s$. Now let's divide both sides by n_i, the degree of χ_i. Then we get

$$\sum_{j=1}^{s} \left(|\mathcal{C}_j| \chi_i(g_j)/n_i\right) \chi(g_j^{-1}).$$

Now by our theorem above we have that

$$(|\mathcal{C}_j| \chi_i(g_j)/n_i)$$

are algebraic integers for all j and so are the $\chi(g_j^{-1})$. Thus the left-hand side is an algebraic integer. But the right-hand side is $|G|/n_i$, which is rational. Since a rational algebraic integer is an integer, we see that $n_i || G|$. Let's record this fact.

Theorem 0.10. *Let G be a finite group. Then the degree of every irreducible representation of G divides the order of G.*

Let's get some nice corollaries of this fact.

Theorem 0.11. *Let p be a prime and let G be a group of order p^2. Then G is abelian.*

Proof. We write $\mathbb{C}[G] = \prod_{i=1}^{s} M_{n_i}(\mathbb{C})$. Then the $n_i | p^2$ and so they are either $1, p, p^2$. Since $\sum_i n_i^2 = |G| = p^2$, we see in fact that they are either all equal to 1 or $s = 1$ and $n_1 = p$. But the latter is impossible, since we always have at least one 1-dimensional representation (since $|G/G'| \geq 1$). □

Theorem 0.12. *Let G be a group of order p^n. Then G is solvable.*

Proof. We prove this by induction on n. The result is true for $n = 1, 2$ by the above. We again again write $\mathbb{C}[G] = \prod_{i=1}^{s} M_{n_i}(\mathbb{C})$. Then the $n_i | p^n$ and so they are all powers of p. Thus the number of j such that $n_j = 1$ must be a multiple of p using the fact that $p^n = |G| = \sum_j n_j^2$. Hence $|G/G'| \geq p$ and so $|G'| \leq p^{n-1}$. It follows by induction that G' is solvable and so G is solvable. □

In fact, there is a theorem of Burnside that shows that any group of order $p^n q^m$ is solvable, where p and q are primes. As far as I know, the only proofs of this use representation theory. Notice that it even gets tough for $p^2 q^2$. If one knows Sylow theory then one can prove this case without representation theory.

Exercise 25. *Let G be a group of order pq with p and q primes with $p < q$. Give the possible isomorphism types of $\mathbb{C}[G]$. Your answer will depend upon whether or not p divides $q - 1$.*

Burnside's Theorem

Burnside's theorem shows that if G is a group of size $p^a q^b$ with p, q primes then G is solvable. We have proven this when $b = 0$. The key lemma in Burnside's theorem is the following:

Lemma 0.29. *Let G be a group and suppose there is a prime p such that G has a conjugacy class of size p^n with $n \geq 1$. Then G is not simple.*

Exercise 26. *Let E be a finite Galois extension of $\mathbb{Q}$ and suppose that $\alpha \in E$ is an algebraic integer, all of whose Galois conjugates have modulus < 1. Then show that $\alpha = 0$.*

Proof. Suppose that G is simple. Let $x \in \mathcal{C}$ where $\mathcal{C}$ is a conjugacy class of size p^n. Then we have shown that for each irreducible character χ_i, we have

$$(|\mathcal{C}|\chi_i(x)/n_i)$$

is an algebraic integer.

Observe that if $p \not| n_i$ then $\gcd(n_i, p^n) = 1$ and so we have $an_i + bp^n = 1$ for some integers a and b. Thus $\chi_i(x)/n_i = (an_i + bp^n)\chi_i(x)/n_i = a\chi_i(x) + b|\mathcal{C}|\chi_i(x)/n_i$. Notice that both terms on the right-hand side are algebraic integers and so $\chi_i(x)/n_i$ must be an algebraic integer if $p \not| n_i$. Now $\chi_i(x)$ is the sum of the eigenvalues of $\rho_i(x)$. There are n_i eigenvalues and they are all n-th roots of unity. So $\chi_i(x)/n_i = \left(\sum_{j=1}^{n_i} \omega_j\right)/n_i$. So notice that $\chi_i(x)/n_i \in \mathbb{Q}(\exp(2\pi i/n))$ and all of its Galois conjugates are ≤ 1 in absolute value. Thus by the exercise, we have that either $\chi_i(x)/n_i$ has some conjugate of modulus 1, in which case all the roots of unity are the same and we see that $\rho_i(x) = \omega \mathbf{I}_{n_i}$ is central; or $\chi_i(x)/n_i = 0$. If the former case, we have that the kernel of ρ_i contains all elements of the form $xgx^{-1}g^{-1}$. Since G is simple we see that either ρ_i is trivial or x commutes with all elements of G. But in the latter case we see that $x \in Z(G)$ and since $x \neq 1$ we get that G is not simple. Hence we have that if $p \not| n_i$ then either ρ_i is the trivial representation or $\chi_i(x) = 0$.

Now we're done. Let's look at the left regular representation ρ and let χ be its character. Then if χ_1 is the trivial character, we have

$$0 = \chi(x) = \sum_{i=1}^{s} n_i\chi_i(x) = \chi_1(x) + \sum_{i>1:p\not| n_i} n_i\chi_i(x) + \sum_{i>1:p|n_i} n_i\chi_i(x).$$

We have just shown that $\chi_i(x) = 0$ for $p \not| n_i$ with $i > 1$ so we get $0 = 1 + p\beta$ with $\beta = \sum_{i>1:p|n_i}(n_i/p)\chi_i(x)$, an algebraic integer. But $\beta = -1/p$, which is rational and not an integer, a contradiction. The result follows and we see that G is not simple. □

Exercise 27. *Let G be a group of order $p^a m$ with m relatively prime to p. Then a subgroup of order p^a is called a* Sylow *p-subgroup of G. Lots of you have seen a proof that such subgroups exist, so let's do a weird proof you've probably never seen before with the following steps.*

(i) *Show that $\mathrm{GL}_n(\mathbb{F}_p)$ has size $(p^n - 1)(p^n - p)\cdots(p^n - p^{n-1})$. (Hint: invertible matrices have linearly independent rows. Pick your rows one by one.)*

(ii) *Show that the set of upper triangular matrices in $\mathrm{GL}_n(\mathbb{F}_p)$ with ones along the diagonal forms a Sylow p-subgroup of $\mathrm{GL}_n(\mathbb{F}_p)$.*

(iii) *Show that every finite group is a subgroup of $\mathrm{GL}_n(\mathbb{F}_p)$ for some n.*

(iv) *Show that if G has a Sylow p-subgroup and H is a subgroup of G then H has a Sylow p-subgroup. Then use the preceding steps to get the result. (Hint for the last step: Let P be a Sylow p-subgroup of G. Put an equivalence relation (double cosets) on G via $x \sim y$ if $x = hyp$ for some $h \in H$ and some $p \in P$. Show that this is an equivalence relation. Show that the equivalence class containing x has size $|H| \cdot |P|/|H \cap xPx^1|$. Now show if p^a is the largest power of p that divides $|H|$, then p^a always divides $|H| \cdot |P|/|H \cap xPx^{-1}|$ and show that there must be at least one x for which p^a exactly divides $|H| \cdot |P|/|H \cap xPx^{-1}|$. Now show that for such an x, $xPx^{-1} \cap H$ is a Sylow p-subgroup of H.*

We recall that a finite p-group always has a non-trivial centre. This follows immediately from the class equation, since all conjugacy classes of size > 1 must have size equal to a prime power, so we see that p divides the order of the centre.

Theorem 0.13. *(Burnside's Theorem) Let p and q be primes and let G be a group of size $p^a q^b$. Then G is solvable.*

Proof. We have shown the result when $a = 0$ or $b = 0$, so we may assume that $ab \neq 0$. We suppose that the conclusion does not hold. Then there exists some counter-example with $a + b$ minimal. By the exercise (or the Sylow theorems!) we know that G has a subgroup Q of size q^b and that Q has a non-trivial central element z. Let $\mathcal{C}$ be the conjugacy class of G that contains z. Then $|\mathcal{C}| = |G|/|C(z)|$, where $C(z)$ is the centralizer of z in G. Then $C(z)$ contains Q since z is central in Q. Thus $|Q||C(z)|$ and so $|C(z)| = q^b p^c$ for some c. Now if $c = a$ then $C(z) = G$ and so G has a non-trivial centre. If $Z(G) = G$ then G is abelian and we are done. Thus we may assume that $Z(G)$ is non-trivial and proper and so G has a non-trivial proper normal subgroup. Then by the induction hypothesis both $Z(G)$ and $G/Z(G)$ are solvable and so G is solvable. On the other hand if $c < a$ then $|\mathcal{C}| = |G|/|C(z)| = p^a q^b/q^b p^c = p^{a-c}$ and so G has a conjugacy class of size equal to a prime power. Then the lemma shows us that G is not simple, so it has a proper non-trivial normal subgroup N. By the induction hypothesis both N and G/N are solvable. Done! □

Exercise 28. *Given a finite group G, let $a(G)$ denote the average of the degrees of its irreducible representations and let $f(G) = a(G)/\sqrt{|G|}$. Finally for a natural number n, let $h(n)$ be the supremum of $f(G)$ as G ranges over groups of size n. Show that $h(n) \to 0$ as $n \to \infty$.*

Exercise 29. *Show that if $n \geq 5$ then S_n has no irreducible complex representations of degree 2 or 3.*

Exercise 30. *Suppose that G is a group of odd order. Show that the trivial character is the only irreducible character that is real-valued.*

Induced Representations

We now consider induced representations. Informally, this shows how, given a representation of a subgroup H of G, one can "lift" it to a representation of G. We'll look at this in two different ways. There is a straightforward definition that is hard to remember but does not require any background and one can also use the machinery of tensor products, which gives a more intuitive sense of what an induced representation is. So to start, let H be a subgroup of a finite group G. Then $\mathbb{C}[H]$ is a subalgebra of $\mathbb{C}[G]$. In this way, we may view $\mathbb{C}[G]$ as a left (and right!) $\mathbb{C}[H]$-module: notice that if $c \in \mathbb{C}[H]$ and $r \in \mathbb{C}[G]$ then we define $c\dot{r} = cr$, where the multiplication is performed in $\mathbb{C}[G]$.

Lemma 0.30. *As a left $\mathbb{C}[H]$-module, $\mathbb{C}[G]$ is free (that is, it has a basis) with a basis given by $\{g_1, \ldots, g_m\}$ where $g_1, \ldots, g_m$ is a set of right coset representatives for G/H.*

Proof. We have $G = Hg_1 \cup \cdots \cup Hg_m$. Recall that a representation V of H is really just a finite-dimensional $\mathbb{C}[H]$-module. Notice then that if $\sum_{g\in G} \alpha_g g$ is in $\mathbb{C}[G]$ then we may rewrite it as

$$\sum_{i=1}^{r} \sum_{h\in H} \alpha_{hg_i} hg_i = \sum_{i=1}^{r} \left(\sum_{h\in H} \alpha_{hg_i} h \right) g_i.$$

Thus $\mathbb{C}[G] = \sum_{i=1}^{r} \mathbb{C}[H]g_i$. We notice that this sum is direct by the disjointness of support. □

A similar argument shows that $\mathbb{C}[G]$ is in fact a free right $\mathbb{C}[H]$-module with basis given by a set of left coset representatives.

Tensor products

Now a general fact about rings is that given a subring R of a ring S, if one has a left R-module M then one can "extend scalars" and create a left S-module, which we denote by $M \otimes_R S$. We'll spend the next little while doing this construction rigorously, but then we'll give a concrete interpretation for group algebras.

First, if you have seen tensor products in the commutative setting or in the setting of vector spaces, you are in good shape—still there are some subtleties that arise in the noncommutative setting. We start by letting R be a ring that is not necessarily commutative. Given a left R-module M and a right R-module N we can form a left R-module $M \otimes_R N$, which is called the tensor product of M and N, as follows.

First, we recall that in this setting, if A is an abelian group then a map $f : M \times N \to A$ is said to be *bilinear* if $f(m + m', n) = f(m, n) + f(m', n)$ for all $m, m' \in M$ and $n \in N$; $f(m, n + n') = f(m, n) + f(m, n')$ for all $m \in M$, $n, n' \in N$ and for $m \in M, n \in N$ and $r \in R$ we have $f(mr, n) = f(m, rn)$.

Important Remark! Notice that we really need the right/left pairing here to make this work in general. If M and N were both left R-modules then if we tried to impose $f(rm, n) = f(m, rn)$ then we'd have for $r, s \in R$

$$f(m, (rs)n) = f((rs)m, n) = f(sm, rn) = f(m, srn) = f(m, (sr)n),$$

so we'd have $f(m, (rs - sr)n) = 0$ for all $(m, n) \in M \times N$ and every $r, s \in R$. Now in certain rings, one can have $1 = rs - sr$ for some $r, s \in R$. For example, if one takes the ring of all linear operators on $\mathbb{C}[x]$ then the differentiation operator and multiplication by x operator have this relation. So in this situation one would have $f(m, n) = 0$ for all m, n. But the way we have defined it, we now have $f(m, (rs)n) = f(m(rs), n) = f(mr, sn) = f(m, rsn)$ and there is no problem now.

Second, we let T denote the free $\mathbb{Z}$-module on all symbols $e_{m,n}$ where $(m, n) \in M \times N$. That is, T is all finite integer linear combinations of elements of the form $e_{m,n}$. Then we let U denote the subgroup of T generated by the relations

$$e_{m+m',n} - e_{m,n} - e_{m',n} = 0$$

$$e_{m,n+n'} - e_{m,n} - e_{m,n'} = 0$$

$$e_{mr,n} = e_{m,rn}$$

for all $m, m' \in M$, $n, n' \in N$, and $r \in R$. What we are in fact doing is choosing our relations so that the function $M \times N \to T/U$ given by $(m, n) \mapsto e_{m,n} + U$ is now R-*bilinear*. We define $M \otimes_R N$ to be the abelian group T/U. We then use the symbol $m \otimes n$ (read "m tensor n") to denote the image of $e_{m,n}$ in T/U. Now in general, there is no additional structure, but in the case where M is both a left S-module and right R-module (we call this a S-R-bimodule), we can actually give $M \otimes_R N$ the structure of a left S-module as follows: we define $s \cdot (m \otimes n) = (sm) \otimes n$. Notice that if we did not have the bimodule structure on M we'd be in trouble. One might hope that we could still at least put a left R-module structure on the tensor product using the fact that N is a left R-module and define $r \cdot (m \otimes n) = m \otimes (rn)$, but this is problematic: by definition of our relations $m \otimes (rn) = (mr) \otimes n$ and so we'd have to have $(rs)m \otimes n = m \otimes (rs)n = m(rs) \otimes n$ and $(rs)m \otimes n = r \cdot (s \cdot m \otimes n) = r \cdot m \otimes sn = r \cdot (ms \otimes n) = ms \otimes rn = m(sr) \otimes n$. Notice this really only works if $m(rs - sr) = 0$ for all $r, s \in R$. But if we have that M is a bimodule then there is a whole other untapped side which we can use to endow the tensor product with a left module structure.

We remark that in general not every element of $M \otimes_R N$ is of the form $m \otimes n$—we must take sums of elements of this form. An element of the form $m \otimes n$ is called a *pure tensor*. People who have worked with tensor products before know

that it is actually hard to prove even basic properties about them. The way one does this is by using what is known as the *Universal property* of tensor products.

Universal property

For any abelian group A and any bilinear map $f : M \times N \to A$, there exists a unique homomorphism of abelian groups $\widehat{f} : M \otimes_R N \to A$ such that $\widehat{f} \circ i = f$, where $i : M \times N \to M \otimes_R N$ is the $\mathbb{Z}$-module homomorphism induced by $(m, n) \mapsto m \otimes n$.

Let's see that $M \otimes N$ has this universal property. We define $\widehat{f}(m \otimes n) = f(m, n)$ and extend via linearity. We must check that this is well-defined. Notice that f induces a group homomorphism $f' : T \to A$ via $f'(\sum c_i e_{m_i,n_i})$. Saying that f is bilinear is exactly the same as saying that f is zero on U. Thus we can define $\widehat{f} : T/U \to A$ via $\widehat{f}(t + U) = f'(t)$ and this is well-defined. Moreover, this is the only way to define this map!

The universal property actually means that the tensor product is the unique $\mathbb{Z}$-module (up to isomorphism) with this property given a right R-module M and a left R-module N. To see this, suppose that we have two abelian groups A and B that have the universal property for (M, N). Then we have maps $i_1 : M \times N \to A$ and $i_2 : M \times N \to B$ such that if $f : M \times N \to C$ is a $\mathbb{Z}$-module homomorphism then there exist f_1 and f_2 from A and B to C respectively such that $f_1 \circ i_1 = f = f_2 \circ i_2$. Now take $C = A$ and let $f = i_1$. Then there is a map $f_2 : B \to A$ such that $f_2 \circ i_2 = i_1$. Similarly, there exists a unique f_1 such that $f_1 \circ i_1 = i_2$. Then

$$(f_1 \circ f_2) \circ i_2 = f_1 \circ (f_2 \circ i_2) = f_1 \circ i_1 = i_2.$$

Notice also that $i_2 : M \times N \to B$ is onto since we can use the universal property to get that there is some $g : M \times N \to B$ such that $g \circ i_2 = i_2$. Since $g = \mathrm{id}$ works, by *uniqueness* of the universal property we see that g is the identity. But we see that $g = f_1 \circ f_2$ works too, so we see that $f_1 \circ f_2$ is the identity. By symmetry we get that $f_2 \circ f_1$ is the identity and so A and B are isomorphic.

The universal property is how one proves *anything* about tensor products in practice. Let's do a few examples.

Let $R = M_2(\mathbb{C})$ and let M be the right R-module $\mathbb{C}^{1\times 2}$, the 1×2 row vectors; let $N = \mathbb{C}^{2\times 1}$ be the column vectors. Then what is $M \otimes_R N$? Notice that $M \otimes_{\mathbb{C}} N \cong \mathbb{C}^4$, but when we tensor over R we get a different result. First, this will be an abelian group, so we just need to find out what it is as an abelian group. As before we let T be the free $\mathbb{Z}$-module on the generators $e_{v,w}$ where v is a row vector and w is a column vector. Now let $v_1 = [1, 0]$ and let $w_1 = [0, 1]$ be a nonzero column vector. Then for $v \in M$ and $w \in N$ there is some $r \in R$ such that $v = v_1 r$. So $e_{v,w} = e_{v_1 r,w} = e_{v_1,rw}$. Now if $w' = rw$ then we have that the tensor product is generated by the images in T/U of things of the form $e_{v_1,w'}$ with w' a column vector. If s is an element of R whose first row is $(1, 0)$ then $v_1 s = v_1$ and so in T/U we have $e_{v_1,w'} = e_{v_1 s,w'} = e_{v_1,sw'}$. Notice that by picking s appropriately, we may arrange things so that $sw' = \lambda w_1$ for some scalar λ. So now we see we are spanned by things of the form $e_{v_1,\lambda w_1}$. Now by the other bilinear relations, we see that $e_{v_1,\lambda_1 w_1} + e_{v_2,\lambda_2 w_1} = e_{v_1,(\lambda_1+\lambda_2)w_1}$ and so we see that we have a map from the abelian group $\mathbb{C}$ to $T/U = M \otimes_R N$ via the rule $\lambda \mapsto e_{v_1,\lambda w_1} + U$ and we have shown that this map is onto. Now it would be pretty difficult to show that this is 1-to-1 in general, but we can use the universal property to do this. Notice that we have a bilinear map $f : M \times N \to \mathbb{C}$ via the rule $(v, w) \mapsto v \cdot w$. Then under this map $(v_1, \lambda w_1)$ maps to λ. So if λ is in the kernel it must be zero. Thus the tensor product is just the complex numbers in this case.

Exercise 31. *Let R be a ring, let M be a right R-module, and let N be a left R-module. Suppose we regard R as a left R-module. Show that $M \otimes_R R \cong M$. Show that if we regard R as a right R-module then $R \otimes_R N \cong N$. (Hint: use the bilinear map $M \times R \to R$ given by $(m, r) \mapsto mr$.)*

Exercise 32. *Let R be the ring of upper-triangular 2×2 complex matrices and let $M = \mathbb{C}^{1\times 2}$ be the 1×2 row vectors; let $N = \mathbb{C}^{2\times 1}$ be the column vectors. What is $M \otimes_R N$?*

Exercise 33. *Use the universal property to show that if M_1 and M_2 are two right R-modules and N is a left R-module then $(M_1 \oplus M_2) \otimes_R N \cong (M_1 \otimes_R N) \oplus (M_2 \otimes_R N)$. Show that if M_1 and M_2 are also left S-modules then this is an isomorphism of left S-modules, too.*

Extension of scalars

One of the nicest features of tensor products is that they can be used to extend scalars. For this set-up, let R be a subring of a ring S and let M be a left R-module. Then we can create a left S-module from M via tensor products as follows. Notice that S is both a left S-module and a right R-module and thus we can form $S \otimes_R M$. If we only used the right R-module structure of S, this would just be an abelian group, but since S is a left S-module, we can give it a left S-module structure. We can give $S \otimes_R M$ a left S-module structure via the rule $s \cdot (s' \otimes m) = (ss') \otimes m$. Since the left S-module structure does not interfere with the right R-module structure this does not create any problems.

Exercise 34. *Let R be a subring of a ring S and suppose that S is a free right R-module with basis $s_1, \ldots, s_d$. Show that if N is a left R-module then as an abelian group $S \otimes_R N \cong N^d$ with isomorphism induced from*

$$(s_1 r_1 + \cdots + s_d r_d) \otimes n = \sum_i s_i \otimes (r_i n) \mapsto (r_1 n, r_2 n, \ldots r_d n).$$

In our setting, we have a subgroup H of a group G. Then $\mathbb{C}[H]$ is a subring of $\mathbb{C}[G]$ and a representation of H is a finite-dimensional left $\mathbb{C}[H]$-module V. Then we can form a $\mathbb{C}[G]$-module from V via

$$V\uparrow_H^G := \mathbb{C}[G]\otimes_{\mathbb{C}[H]} V.$$

This is called the *representation of G induced* from V. Personally, I find this a nicer way of thinking of induced representations, but it is less concrete than the way it is often introduced. When I learned representation theory, induced representations were taught in the concrete way. That meant that I memorized the way of computing the character, but I had no intuition for where it came from. This way is admittedly less concrete, but we can easily derive the concrete formulation from this set-up. So let's try to see how to think about this more concretely. Recall that $\mathbb{C}[G]$ is a free right $\mathbb{C}[H]$ module with basis $g_1,\dots,g_d$ where $G = g_1H\cup\cdots g_dH$ and the union is disjoint (i.e., the g_i are a set of left coset representatives). Then as an abelian group we have that $W := V\uparrow_H^G \cong V^d$ with the isomorphism coming via the rule

$$\left(\sum g_i f_i\right)\otimes v \mapsto (f_1v,\dots,f_dv),$$

where the f_i are in $\mathbb{C}[H]$.

In particular, $g_i\otimes v_j$ form a $\mathbb{C}$-basis for W.

So now let us suppose that V has dimension m and that $v_1,\dots,v_m$ is a basis for V. Then since V is a representation of H, we have a homomorphism $\rho: H\to \mathrm{GL}(V)$ and we have $\rho(h)(v_j) = \sum a_{i,j}(h)v_i$ for some scalars $a_{i,j}(h)$. Now W has a basis given by $g_i\otimes v_j$ with $i=1,\dots,d$ and $j=1,\dots,m$. So the question is: how does $g\in G$ act on this basis? To figure this out, notice that $g\cdot g_i = g_k h$ for some $h\in H$ and some $k\in\{1,\dots d\}$ and these are uniquely determined. Then by the definition of the $\mathbb{C}[G]$ module structure on the tensor product, we have

$$g\cdot(g_i\otimes v_j) = (gg_i)\otimes v_j = (g_kh\otimes v_j) = g_k\otimes h\cdot v_j = g_k\otimes(\sum a_{i,j}(h)v_i) = \sum a_{i,j}(h)g_k\otimes v_i.$$

So what $\rho(g)$ looks like is a $d\times d$ block matrix with copies of $\rho(h)$ or the zero matrix in the blocks.

The character of the induced representation

More important, is the character of the induced representation. This has an even more concrete formulation. Notice that if we use $g_i\otimes v_j$ as our basis then the trace of $\rho(g)$ is just the sum over the contribution of $g_i\otimes v_j$ in $g\cdot(g_i\otimes v_j)$. Now if $gg_i = g_kh$ and $k\neq i$ then the contribution is zero! So we really only need to consider the i for which $gg_i = g_ih$ for some $h\in H$. That is, we are interested in the set of i for which $g\in g_iHg_i^{-1}$. In that case, the coefficient of $g_i\otimes v_j$ in $g\cdot(g_i\otimes v_j)$ is just $a_{j,j}(h)$, where $h = g_i^{-1}gg_i$. Thus when we take the trace we sum over all h and pick up a contribution of $\chi_V(h)$ for this particular block. Thus we see that we get the formula:

$$\chi_{V\uparrow_H^G}(g) = \sum_{\{i: g_i^{-1}gg_i\in H\}} \chi_V(g_igg_i^{-1}).$$

It is often convenient to use the related formula, which we leave as an exercise:

$$\chi_{V\uparrow_H^G}(g) = \frac{1}{|H|}\cdot\sum_{\{x\in G:\ xgx^{-1}\in H\}}\chi_V(xgx^{-1}).$$

A lot of books simply defined the induced representation by this formula and so if you simply want to memorize this formula, that is fine too. But notice that we derived it in a very natural way. Also, using the power of tensor products we can get some nice results that would have to be painstakingly proven using the above formula otherwise.

Exercise 35. *Let $G = S_n$ and let H be the subgroup consisting of permutations that fix n; that is, $H\cong S_{n-1}$. Let V be the trivial representation of H; i.e., $V=\mathbb{C}$ and $h\cdot v = v$ for all $h\in H$, $v\in V$. Show that the character of $V\uparrow_H^G$ sends σ to the number of fixed points.*

Exercise 36. *Show that the average number of fixed points of a permutation in S_n is 1. Conclude that the character from the preceding exercise has one copy of the trivial representation.*

Let's do an example. Let $H = S_3$ and let V be the left regular representation. We regard H as the subgroup of $G = S_4$ consisting of permutations that fix 4. Find the character of the induced representation $V\uparrow_H^G$.

To do this, we note that a set of coset representatives is given by $(1,4),(2,4),(3,4),(4,4)$. Let $g_i = (i,4)$. Notice that if $\sigma\in S_4$ then $(i,4)\sigma(i,4)\in H$ if and only if $\sigma(i) = i$. So

$$\chi_{V\uparrow_H^G}(\sigma) = \sum_{\{i:\sigma(i)=i\}}\chi_V(g_i\sigma g_i^{-1}).$$

Now we know that the character of the left regular representation of H is 0 on all non-identity elements and is $|H|$ otherwise. If σ is not the identity then $g_i\sigma g_i^{-1}$ is never the identity so $\chi_V(g_i\sigma g_i^{-1})$ is zero and we see that $\chi_{V\uparrow_H^G}(\sigma) = 0$. On the other hand, if σ is the identity then we see that every point is fixed and we get $4|H| = |G|$ and so $\chi_{V\uparrow_H^G} = \chi_{reg,G}$ where $\chi_{reg,G}$ is the left regular representation of G. Thus $V\uparrow_H^G$ is the left regular representation of G.

Exercise 37. *Show that if H is a subgroup of G then if we induce the left regular representation of H up to G we obtain the left regular representation of G.*

Exercise 38. *Let H be the Klein-four subgroup of $G = S_4$. Then we have a 1-dimensional representation of H that sends $(12)(34)$ to -1 and $(13)(24)$ to -1 and all other elements to 1. Find the character obtained by inducing this representation up to S_4.*

Transitivity of induced representations

One of the tricky things about induced representations is figuring out how many copies of the various irreducible representations they have. One way of doing this is using Frobenius reciprocity, which allows one to compute inner products involving induced representations in terms of inner products of the underlying representations in the smaller subgroup. We first remark that given a class function $f : H \to \mathbb{C}$, if H is a subgroup of G, then we can define a new class function $f \uparrow_H^G : G \to \mathbb{C}$ via the rule

$$f \uparrow_H^G (g) = \frac{1}{|H|} \sum_{\{x \in G : xgx^{-1} \in H\}} f(xgx^{-1}).$$

When f is a character this is precisely the formula for the character of induced representation. To see that this is a class function for G, observe that if we replace g by ugu^{-1} in the above formula, then a simple change of variables gives us that the two values are the same.

Before we do Frobenius reciprocity, we'll give some of the basic properties of induced representations. The first important property is that if $H \leq G \leq K$ are groups and f is a class function on H then we have

$$(f \uparrow_H^G) \uparrow_G^K = f \uparrow_H^K .$$

We can do this via a routine calculation. In the exercises, we'll do a more "advanced" computation-free way of seeing this.

Proposition 0.31. *(Transitivity of induction)* $(f \uparrow_H^G) \uparrow_G^K = f \uparrow_H^K$.

Proof. We have

$$\begin{aligned}
(f \uparrow_H^G) \uparrow_G^K (k) &= \frac{1}{|G|} \sum_{\{x \in K :\, xkx^{-1} \in G\}} (f \uparrow_H^G)(xkx^{-1}) \\
&= \frac{1}{|G|} \sum_{\{x \in K :\, xkx^{-1} \in G\}} \frac{1}{|H|} \sum_{\{y \in G :\, yxkx^{-1}y^{-1} \in H\}} f(yxkx^{-1}y^{-1}) \\
&= \frac{1}{|G||H|} \sum_{\{x \in K :\, xkx^{-1} \in G\}} \sum_{\{y \in G :\, yxkx^{-1}y^{-1} \in H\}} f(yxkx^{-1}y^{-1}) \\
&= \frac{1}{|G||H|} \sum_{\{x \in K, y \in G :\, yxkx^{-1}y^{-1} \in H\}} f(yxkx^{-1}y^{-1}) \text{ (since } yxkx^{-1}y^{-1} \in H \implies xkx^{-1} \in G) \\
&= \frac{1}{|H|} \sum_{\{z \in K :\, zkz^{-1} \in H\}} f(zkz^{-1}) \text{ (let } z = yx) \\
&= f \uparrow_H^K .
\end{aligned}$$

□

We can see that via the following exercises too.

Exercise 39. *Let $R \subseteq S \subseteq T$ be rings. Let M be a left R-module. Show that as left T-modules we have $T \otimes_R M \cong T \otimes_S (S \otimes_R M)$. (Hint: you should use the Universal property. Google associativity of tensor products if you get stuck: notice that $T \otimes_S (S \otimes_R M) \cong (T \otimes_S S) \otimes_R M \cong T \otimes_R M$ if you have this associativity.*

Exercise 40. *Explain why the preceding exercise shows $(f \uparrow_H^G) \uparrow_G^K = f \uparrow_H^K$ when $H \leq G \leq K$ and f is a class function on H.*

Frobenius reciprocity

Given a class function $f : G \to \mathbb{C}$ on G, it is easy to create a class function for H via *restriction*; that is, we define

$$f \downarrow_H^G (h) = f(h)$$

for $h \in H$. Since conjugate elements of H are conjugate in G, we see that this is a class function for H. Frobenius reciprocity is the following beautiful formula.

Theorem 0.14. *Let H be a subgroup of G. Let f be a class function for G and let g be a class function for H. Then*

$$\langle f, g \uparrow_H^G \rangle_G = \langle f \downarrow_H^G, g \rangle.$$

Proof. This proof is a straightforward computation, but we'll do it anyway. We'll switch a sum, by setting $yx^{-1}y^{-1} = h$, where $h \in H$ is a new variable

$$\begin{aligned}
\langle f, g\uparrow_H^G\rangle_G = \frac{1}{|G|}\sum_{x\in G} f(x)g\uparrow_H^G(x^{-1}) & \\
&= \frac{1}{|G|}\left(\sum_{x\in G} f(x)\cdot\frac{1}{|H|}\sum_{\{y\in G:\, yx^{-1}y^{-1}\in H\}} g(yx^{-1}y^{-1})\right)\\
&= \frac{1}{|G||H|}\sum_{h\in H}\sum_{y\in G} g(h)f(yh^{-1}y^{-1})\\
&= \frac{1}{|G||H|}\sum_{h\in H} g(h)f(h^{-1})|G|\\
&= \frac{1}{|H|}\sum_{h\in H} g(h)f(h^{-1})\\
&= \langle f\downarrow_H^G, g\rangle_H.
\end{aligned}$$

□

This is fairly simple to prove, but it has a lot of interesting consequences. Let's look at a few ways we can use this. Let $\mathrm{fix}_n : S_n \to \mathbb{Z}$ be the map that sends σ to the number of fixed points. Let $F(m,n) = \frac{1}{n!}\sum_{\sigma\in S_n}\mathrm{fix}_n(\sigma)^m$. I computed a few values of $F(m,n)$. Here they are:

	$m=1$	$m=2$	$m=3$	$m=4$	$m=5$	$m=6$
$n=1$	1	1	1	1	1	1
$n=2$	1	2	4	8	16	32
$n=3$	1	2	5	14	41	122
$n=4$	1	2	5	15	51	187
$n=5$	1	2	5	15	52	202
$n=6$	1	2	5	15	52	203

Notice each column is eventually constant (or at least empirically this is the case). The limiting values, being $1, 2, 5, 15, 52, 203$. When I put this in OEIS, I got sequence $A000110$, which told me that this sequence is the *Bell numbers* (no relation ... well, except in the sense that I'm related to all living things on earth). The m-th Bell number, B_m, is the number of partitions of an m-set. So our guess is that

$$\frac{1}{n!}\sum_{\sigma\in S_n}\mathrm{fix}_n(\sigma)^m = B_m$$

whenever $n \geq m$.

Now let's prove this! Let $H = S_{n-1} \leq G = S_n$, where we regard H as the subgroup of permutations that fix n. Let f be the trivial character of H. Then we showed in the exercises that $f\uparrow_H^G(\sigma) =: \mathrm{fix}_n$ is the number of fixed points of $\sigma \in S_n$. Let's notice that f is an irreducible representation of H but $f\uparrow_H^G$ is not irreducible. To see this, notice that

$$\langle f\uparrow_H^G, f\uparrow_H^G\rangle_G = \langle \mathrm{fix}_n\uparrow_H^G, \mathrm{fix}_n\rangle_G = \langle \chi_{triv}, \mathrm{fix}_n\downarrow_H^G\rangle_H.$$

Notice that fix_n on H is $\mathrm{fix}_{n-1} + \chi_{triv}$ and so we see that our inner product is just $1 + \langle\chi_{triv}, \mathrm{fix}_{n-1}\rangle_H = 1 + 1$, where the last step follows from the assignment.

Here's something that's not easy to show in general.

Proposition 0.32. *Let $d \leq n$ be natural numbers. Then*

$$\frac{1}{n!}\sum_{\sigma\in S_n}\mathrm{fix}_n(\sigma)^d = B_d,$$

where B_d is the number of set partitions of a set of d elements.

Proof. We first note that we have $B_d = \sum_{j=0}^{d-1}\binom{d-1}{j}B_j$. To see this, consider a partition of the set $\{1, 2, \ldots, d\}$. Then the set that contains d has size $j+1$ for some $j \in \{0, \ldots, d-1\}$. Then we have $\binom{d-1}{j}$ ways of picking the j non-d elements of the set containing d. After we do this, we are then left with a set partition of a set of size $d-j-1$ and so we see

$$B_d = \sum_{j=0}^{d-1}\binom{d-1}{j}B_{d-1-j} = \sum_{j=0}^{d-1}\binom{d-1}{j}B_j.$$

We now prove this by induction on n. The case when $n = 1$ is trivial. Assume it holds up to $n-1$. We'll now find the same recurrence is satisfied for the left side of our formula. Notice that

$$\frac{1}{n!}\sum_{\sigma\in S_n}\mathrm{fix}_n(\sigma)^d = \langle \mathrm{fix}_n^{d-1}, \mathrm{fix}_n\rangle_{S_n}.$$

Now

$$\begin{aligned}\langle \mathrm{fix}_n^{d-1}, \mathrm{fix}_n\rangle_{S_n} &= \langle \mathrm{fix}_n^{d-1}, \chi_{triv}\uparrow_{S_{n-1}}^{S_n}\rangle_{S_n}\\ &= \langle \mathrm{fix}_n^{d-1}\downarrow_{S_{n-1}}^{S_n}, \chi_{triv}\rangle_{S_{n-1}}\end{aligned}$$

Now fix_n on S_{n-1} is just $1+\mathrm{fix}_{n-1}$ since we get the extra fixed point since n is always fixed in S_{n-1}. Thus we have

$$\begin{aligned}\langle \mathrm{fix}_n^{d-1}, \mathrm{fix}_n\rangle_{S_n} &= \langle \mathrm{fix}_n^{d-1}\downarrow_{S_{n-1}}^{S_n}, \chi_{triv}\rangle_{S_{n-1}}\\ &= \langle (\mathrm{fix}_{n-1}+1)^{d-1}, \chi_{triv}\rangle_{S_{n-1}}\\ &= \sum_{j=0}^{d-1}\binom{d-1}{j}\langle \mathrm{fix}_{n-1}^{j}, \chi_{triv}\rangle_{S_{n-1}}.\end{aligned}$$

Notice that since $j \le d-1 \le n-1$, we have by induction hypothesis

$$\langle \mathrm{fix}_{n-1}^{j}, \chi_{triv}\rangle_{S_{n-1}} = \frac{1}{(n-1)!}\sum_{\sigma\in S_{n-1}}\mathrm{fix}_{n-1}(\sigma)^j = B_j,$$

Thus

$$\langle \mathrm{fix}_n^{d-1}, \mathrm{fix}_n\rangle_{S_n} = \sum_{j=0}^{d-1}\binom{d-1}{j}B_j.$$

But we saw that this is B_d and so the result follows by induction. □

In fact, Frobenius reciprocity is great for computing permutation statistics, since it often yields recurrences of this type. Let's see another example.

Example 0.33. Let $H = D_4 = \langle (1234), (13)\rangle$ be a subgroup of size 8 of $G = S_4$. If one induces the trivial character on H up to G, how does it decompose into irreducible representations?

We can use Frobenius reciprocity. Recall that S_4 has five irreducible characters, which we've computed already. Let's call then $\chi_1, \ldots, \chi_5$. If we restrict the induced character of the trivial character, what do we get? Notice that for $h \in H$ we have

$$T(h) := \chi_{triv}\uparrow_H^G(h) = \frac{1}{|H|}\sum_{\{x\in G:\ xhx^{-1}\in H\}}\chi_{triv}(xhx^{-1}) = \#\{x\in G:\ h\in xHx^{-1}\}.$$

Notice that D_4 comprises the identity, two four-cycles, two two-cycles, and the three non-identity elements of the Klein four-group. The Klein-four group, so $T(h) = 3$ for all h in the Klein-four group; The two two-cycles are (13) and (24). We are looking at the set of $x \in G$ such that $x(13)x^{-1} = (13)$ or (24). This is just H, so $T(h) = 1$ for these elements. Similarly, $T(h) = 1$ for the four-cycles, too. Thus $\langle \chi_{triv}\uparrow_H^G, \chi_i\rangle_G = \langle T, \chi_i\downarrow_H^G\rangle_H$. Now we computed the character table of S_4 before and we had

	id	(12)	(123)	(12)(34)	(1234)
χ_1	1	1	1	1	1
χ_2	1	-1	1	1	-1
χ_3	2	0	-1	2	0
χ_4	3	1	0	-1	-1
χ_5	3	-1	0	-1	1

So now we just look at the characters on these values with the weights indicated and use Frobenius reciprocity and see what happens.

Exercise 41. *Let $H = S_3$ and let $G = S_4$ where we regard S_3 as the permutations that fix 4. Let χ_1, χ_2, χ_3 be the irreducible characters of H. Find how they decompose into irreducible characters when they are induced up to G.*

Mackey's criterion

We have seen that in general, the induced representation of an irreducible representation need not be irreducible. Mackey's criterion tells you when an induced representation is irreducible. Let H, K be two subgroups of a group G. Then we'll let χ be a character of H. One thing, we'd like to understand is the decomposition of $\left(\chi\uparrow_H^G\right)\downarrow_K^G$. To see this, we first need to look at the *double coset* decomposition. Given subgroups H and K of a group G, we can put an equivalence relation $\sim$ on G via the rule $x \sim y$ if there exist some $k \in K$ and $h \in H$ such that $kxh = y$. Check that this is an equivalence relation. The equivalence classes are things of the form KxH with $x \in G$. We note that unlike with left or right cosets,

double cosets need not have the same size. For example, if $G = S_3$ and $H = \{id, (12)\}$ and $K = \{id, (12)\}$ then $KH = H$; $K(13)H = \{(13), (132), (123), (23)\}$. So what is the size of KxH? To figure this out, observe that

$$|KxH| = |KxHx^{-1}| = |K \cdot (xHx^{-1})|.$$

Now xHx^{-1} is a group. So this size is just $|K||xHx^{-1}|/|K \cap xHx^{-1}|$ because each product in $K(xHx^{-1})$ is counted $|H \cap xHx^{-1}|$ times. We pick a set of double coset representatives S for G with the groups H and K. Then given a character χ of H, for each $s \in S$ we can create a character χ_s of $H_s := K \cap sHs^{-1}$ via $\chi_s(shs^{-1}) = \chi(h)$ when $shs^{-1} \in K$. Since H_s is a subgroup of K, we can induce χ_s up to K to obtain a character for K, $\chi_s \uparrow_{H_s}^K$. These give a decomposition of the character $\left(\chi \uparrow_H^G\right) \downarrow_K^G$, as we shall show.

Theorem 0.15. *We have*

$$\left(\chi \uparrow_H^G\right) \downarrow_K^G = \sum_{s \in S} \chi_s \uparrow_{H_s}^K .$$

Proof. We have $G = \cup_{s \in S} KsH = \cup_{s \in S} \cup_{xH \in R_s sH}$, where both unions are disjoint and R_s is a set of left coset representatives for $K/K \cap sHs^{-1}$. We already know the first union holds, but let's look at the second union. Notice that $ksH = k's'H$ if and only if $s = s'$ by our choice of S. Moreover $ksH = k'sH$ if and only if $(k')^{-1}k \in sHs^{-1} \cap K$. That is they generate the same left coset. Now let $T = \cup_{s \in S} R_s \cdot s$. Then G is the disjoint union of tH with $t \in T$. Thus T is a set of left coset representatives for H as a subgroup of G. It follows that the induced representation of χ up to G at $k \in K$ is given by

$$\sum_{\{t \in T:\ t^{-1}kt \in H\}} \chi(t^{-1}kt) = \sum_{s \in S} \sum_{r \in R_s} \chi(s^{-1}r^{-1}krs) = \sum_{s \in S} \sum_{\{r \in R_s:\ r^{-1}kr \in sHs^{-1}\}} \chi_s(r^{-1}hr).$$

But since R_s is a transversal for $K/K \cap sHs^{-1}$, we see this is just

$$\sum_{s \in S} \sum_{\{r \in R_s:\ r^{-1}kr \in H_s\}} \chi_s(r^{-1}kr) = \sum_{s \in S} \chi_s \uparrow_{H_s}^K .$$

□

Now we can give Mackey's criterion for when an induced character is irreducible. Let χ be a character of a subgroup H of a group G. We recall that a character f of a group G is irreducible if and only if $||f||^2 = \langle f, f \rangle = 1$. The reason for this is that if we write $f = \sum a_i \chi_i$ with the $a_i \geq 0$ integers and the χ_i the irreducible characters, then the square of the norm of f is just $\sum a_i^2$. If this is 1 then $f = \chi_i$ for some i and otherwise if is at least two. So now we have $\chi \uparrow_H^G$ is irreducible if and only if

$$\langle \chi \uparrow_H^G, \chi \uparrow_H^G \rangle_G = 1.$$

By Frobenius reciprocity, this holds if and only if

$$\langle \chi, \chi \uparrow_H^G \downarrow_H^G \rangle_H = 1.$$

Now we just saw that we can write

$$\chi \uparrow_H^G \downarrow_H^G = \sum_{s \in S} \chi_s \uparrow_{H_s}^H,$$

where s runs over a set of double coset representatives with $G = \cup_{s \in S} HsH$. So now we have that

$$\langle \chi, \chi \uparrow_H^G \downarrow_H^G \rangle_H = \sum_{s \in S} \langle \chi, \chi_s \uparrow_{H_s}^H \rangle_H.$$

But notice that by Frobenius reciprocity, we have

$$\langle \chi, \chi_s \uparrow_{H_s}^H \rangle_H = \langle \chi \downarrow_{H_s}^H, \chi_s \rangle_{H_s} = \frac{1}{|H_s|} \sum_{g \in H_s} \chi(g)\chi_s(g^{-1}) = \frac{1}{|H_s|} \sum_{g \in H_s} \chi(g)\chi(s^{-1}g^{-1}s).$$

Now we can assume that $s = 1$ is in our set of double coset representatives. Then for $s = 1$, we have $H_s = H$ and

$$\frac{1}{|H_s|} \sum_{g \in H_s} \chi(g)\chi(s^{-1}g^{-1}s) = \langle \chi, \chi \rangle_H.$$

Thus

$$\langle \chi \uparrow_H^G, \chi \uparrow_H^G \rangle_G = \langle \chi, \chi \rangle_H + \sum_{s \in S \setminus \{1\}} \langle \chi \downarrow_{H_s}^H, \chi_s \rangle_{H_s}.$$

So notice that the right-side is equal to $\langle \chi, \chi \rangle_H \geq 1$ plus nonnegative integers. The left-side is one if and only if we have:

(i) $\langle \chi, \chi \rangle_H = 1$; and
(ii) $\langle \chi \downarrow_{H_s}^H, \chi_s \rangle_{H_s} = 0$ for $s \notin H$.

We record this more formally.

Theorem 0.16. *(Mackey's criterion) Let χ be a character of a subgroup H of a group G. Then the induced character is irreducible if and only if χ is irreducible and $\langle \chi \downarrow_{H_s}^H, \chi_s \rangle_{H_s} = 0$ for $s \in G \setminus H$.*

Linear groups

We've reached the end of the syllabus, but not the end of the lecture period. I thought it would be interesting to do a few results about subgroups of $\mathrm{GL}_n(\mathbb{C})$, since representation theory, at its very heart, is about looking at images of groups inside GL_n. Some of these are more useful in the context of representations of infinite groups, but Jordan's theorem, which we will soon give, is a very useful for representations of finite groups. We note that in a lot of these theorems we'll give explicit bounds—this is more done for the sake of proving that we can find explicit bounds rather than actually trying to find the *best* bound possible. Where best bounds are known, we state those, although they are generally much harder. My point of view is that the explicit bounds are not really the most interesting part, but rather it is the fact that they even exist that is interesting.

Jordan's theorem

Jordan's theorem is a result about finite subgroups of $\mathrm{GL}_n(\mathbb{C})$.

Theorem 0.17. *Let n be a natural number. Then there exists a constant C_n, depending only upon n, such that if G is a finite subgroup of $\mathrm{GL}_n(\mathbb{C})$ then G has an abelian subgroup A of index at most C_n.*

Notice that if G is small, this really says nothing. The trivial group will have index at most C_n if $|G| \leq C_n$. On the other hand, this is very useful when G is large. It says that G is "almost abelian" in this case. We will use this to prove results about representations of S_n as an application. This proof really requires a bit of analysis, so we will give the outline and then do the steps in details.

Steps in the proof of Jordan's theorem

We recall that if V is a finite-dimensional complex inner product space, then a linear operator $T : V \to V$ is called *unitary* if T^* and T are inverses, where $T^* : V \to V$ is the linear map, defined by the rule $\langle T^*v, w\rangle = \langle v, Tw\rangle$ for all $v, w \in V$. It is a straightforward exercise in linear algebra that T^* exists, is unique, and is linear. We also have that $(TS)^* = S^*T^*$ (exercise!) and so we can show that the unitary matrices form a subgroup of GL_n.

(i) The proof is by induction on n, with the $n = 1$ case being trivial.
(ii) First show that we may assume that G is a subgroup of the unitary group with respect to some inner product on $\mathbb{C}^n$. (We have more or less done the necessary work to show this.)
(iii) We have an operator norm on matrices where we take $||A||$ to be the supremum of $||Av||$ where v ranges over all vectors of norm 1. We show that if matrices $A, B \in G$ are sufficiently close to the identity (i.e., $||A-I||$ and $||B-I||$ are small) then $[A, B] = ABA^{-1}B^{-1}$ is even closer to the identity.
(iv) This means that there is some $\epsilon > 0$ such that $||A - I||, ||B - I|| < \epsilon \implies ||AB - BA|| < \epsilon$.
(v) Now we look at the subgroup of G generated by elements that are within ϵ of the identity. We show that this index is bounded purely in terms of n and that this is an abelian subgroup under general circumstances; if it is not, we can decompose G into smaller groups and use an induction hypothesis.

So what is the inner product that turns G into a subgroup of the unitary group? We have already seen it. We take the ordinary Euclidean inner product and we average over the elements of G. Let $\langle \cdot, \cdot \rangle$ be the ordinary Euclidean inner product on $\mathbb{C}^n$. We then define

$$\langle v, w\rangle_G = \frac{1}{|G|} \sum_{g \in G} \langle gv, gw\rangle.$$

Exercise 42. *Let G be a finite subgroup of $\mathrm{GL}_n(\mathbb{C})$ Show that the bilinear map $\langle \cdot, \cdot \rangle_G$ defined in class is an inner product on $\mathbb{C}^n$ and that the map $T_g : \mathbb{C}^n \to \mathbb{C}^n$ given by $T_g(v) = gv$ is unitary with respect to this inner product; i.e., $T_g T_g^* = T_g^* T_g = I$.*

So we have accomplished Step 1 in our plan. We have now shown that the maps T_g (which we identify with g) are unitary with respect to the inner product $\langle v, w\rangle_G$.

So now we get to the interesting step. We'd like to show that if U is the unitary group in $\mathrm{GL}_n(\mathbb{C})$ then there is some $\epsilon > 0$, depending only upon n, such that if $||A - I||, ||B - I|| < \epsilon$ then $||[A, B] - I|| < \min(||A - I||, ||B - I||)$. This is not too hard. So first observe that

$$||[A, B] - I|| = ||ABA^{-1}B^{-1} - I|| = ||AB - BA||$$

where the last step follows from the fact that unitary matrices have norm 1 and take the unit ball in $\mathbb{C}^n$ to itself. Now we write

$$||AB - BA|| = ||(A - I)B - B(A - I)|| = ||(A - I)(B - I) - (B - I)(A - I)|| \leq ||(A - I)(B - I)|| + ||(B - I)(A - I)||,$$

where the last step follows from the triangle inequality. Finally, we recall that $||CD|| \leq ||C|| \cdot ||D||$, and so $||(A - I)(B - I)|| + ||(B - I)(A - I)|| \leq 2||A - I|| \cdot ||B - I||$. So as long as $\epsilon < 1/2$ we see that we get the desired result. Now we're going to take $\epsilon = 1/(10n)$ as it turns out. If we take ϵ too big (even if less than $1/2$), we'll run into problems with getting our group to be abelian.

Lemma 0.34. *Let $G \leq U$ be a finite subgroup of the unitary subgroup of $\mathrm{GL}_n(\mathbb{C})$. Then there exists a constant κ_n, depending only on n, such that the subgroup H of G generated by elements $A \in G$ with $||A - I|| < 1/(10n)$ has index at most κ_n in G.*

Proof. To prove this, we first make the remark that $\mathrm{GL}_n(\mathbb{C})$ has a topology induced by the operator norm (we saw that the operator norm makes $M_n(\mathbb{C})$ into a metric space) and that if$A \in U$ then $||A|| = 1$ and so U is contained in the closed unit ball of GL_n. Now let $\epsilon = 1/(10n)$. We cover the closed unit ball in $\mathrm{GL}_n(\mathbb{C})$ by a family of ϵ-balls $B(A, \epsilon) = \{B \colon ||B - A|| < \epsilon$. Since the closed ball is compact, we only need finitely many balls.

Exercise 43. *Let $A = (a_{i,j})$ be an $n \times n$ matrix. Show that if all $a_{i,j}$ satisfy $a_{i,j} < \epsilon/n$ then A is the open unit ball of $M_n(\mathbb{C})$ of radius ϵ. Conclude that if $N = 4/\epsilon$ then we need at most N^{n^2} balls of radius ϵ to cover the closed unit ball. (In our proof of Jordan's theorem, we take $\epsilon = 1/(10n)$, so we can take the constant $\kappa_n = (40n)^{n^2}$ in the proof of Lemma 0.34)*

We call this number κ_n: it only depends upon ϵ and n and hence only on n. Now we claim that $[G : H] \leq \kappa_n$. To see this, suppose that we could find distinct right coset representatives $B_1, \ldots, B_{\kappa_n+1} \in G$. Since $G \subseteq U$, by the pigeon-hold principle, there are $i \neq j$ such that B_i and B_j are in the same ϵ-ball. But this means that $||B_i - B_j|| < \epsilon \implies ||B_i B_j^{-1} - I|| < \epsilon \implies B_i B_j^{-1} \in H$, so they are in the same right coset, contradiction. □

Lemma 0.35. *Let $G \leq U$ be a finite subgroup of the unitary subgroup of $\mathrm{GL}_n(\mathbb{C})$. Then the subgroup H of G generated by elements $A \in G$ with $||A - I|| < 1/(10n)$ either consists of scalar matrices or H has a central element that is not a scalar multiple of the identity.*

Proof. If H consists of scalar matrices, then we are done, so we may assume that H has non-scalars in it. Pick $A \in H$ not equal to a scalar multiple of the identity that is at least as close to the identity as any other non-scalar matrix in H. Since G is finite, this is not a problem. Then by definition of H, we see that $||A - I|| < 1/(10n)$, since H is generated by such elements and if they were all scalar matrices then H would be made up of scalar matrices. If B is one of our generators of H (i.e., $||B - I|| < 1/(10n)$) then we see that $||[A, B]] - I|| < ||A - I||$ and so by minimality of $||A - I||$ we see that $[A, B] = \lambda I$ for some λ. We claim that $\lambda = 1$ for all generators B. This means that A is central in H and also A is non-scalar by construction. To see this, notice that $[A, B] = ABA^{-1}B^{-1}$ has determinant 1 and so $\lambda^n = 1$ and so λ is an n-th root of unity. But $||[A, B] - I|| = ||(\lambda - 1)I|| = |\lambda - 1|$. By assumption, $|\lambda - 1| \leq 1/(10n)$. But $\lambda = \exp(2\pi ij/n)$ for some $j \in \{0, 1, \ldots, n - 1\}$. If $j \neq 0$ then we have $|\lambda - 1| \geq |\sin(2\pi ij/n)|$. Now $\sin(x) \geq 2x/\pi$ on $[0, \pi/2]$ and so we see that if $j \in \{1, \ldots, n - 1\}$ then $|\sin(2\pi ij/n)| \geq 4j/n \geq 4/n > \epsilon$. It follows that $\lambda = 1$ and so $[A, B] = 1$ and so A commutes with all generators. □

OK, so now let's put the proof of Jordan's theorem together.

Theorem 0.18. *Let G be a subgroup of $\mathrm{GL}_n(\mathbb{C})$. Then there exists a constant C_n, depending only upon n, such that G has an abelian subgroup of index at most C_n.*

Proof. We prove this by induction on n. When $n = 1$, G is abelian and so we can take $C_1 = 1$. Now suppose that we have computed $C_1, \ldots, C_{n-1}$ and that the claim holds up to $n - 1$. We note that we can take $C_i < C_{i+1}$ for all i since a subgroup of GL_i is a subgroup of GL_{i+1}.

By our lemma G has a subgroup H of index at most $\kappa_n \leq (40n)^{n^2}$ such that H either consists of scalar matrices or H has a central element z that is not a scalar multiple of the identity; in the former case we see that H is abelian and we have an abelian subgroup of index at most κ_n. In the latter case, since z has finite order, we know that after conjugating we can assume that z is a block diagonal matrix, with each block a scalar multiple of the identity and where the scalars in different blocks are distinct. Then since H commutes with z, we know the centralizer of z is isomorphic to $T := \mathrm{GL}_{m_1}(\mathbb{C}) \times \cdots \times \mathrm{GL}_{m_r}(\mathbb{C})$ where $m_1 + \cdots + m_r = n$ and the $m_i < n$ are the sizes of the blocks. Now $H \leq T$. Let $\pi_i(H)$ be the image of H in the projection to $\mathrm{GL}_{m_i}(\mathbb{C})$. Then by the induction hypothesis, $\pi_i(H)$ has an abelian subgroup A_i of index at most C_{m_i}.

Exercise 44. *Let $H \leq G_1 \times G_2 \times \cdots \times G_r$ where the G_i are groups and let π_i bet he projection onto G_i. Show that if $\pi_i(H)$ has an abelian subgroup of index at most N_i for $i = 1, \ldots, r$, then H has an abelian subgroup of index at most $N_1 \cdots N_r$.*

By the exercise, H has an abelian subgroup of index at most $C_{m_1} \cdots C_{m_r} \leq C_{n-1}^r \leq C_{n-1}^n$. But since $[G : H] \leq \kappa_n \leq (40n)^{n^2}$, we see that G has an abelian subgroup of index at most $C_{n-1}^n (40n)^{n^2}$. So as long as we chose $C_n \geq C_{n-1}^n (40n)^{n^2}$, we see that we get the result with C_n for n. □

Working by induction, we see that the constants $C_n = (40n)^{n^{2n}}$ will certainly work, since

$$(40n)^{n^{2n}} \geq \left((40(n-1))^{(n-1)^{2(n-1)}}\right)^n (40n)^{n^2},$$

since

$$(n-1)^{2(n-1)} \cdot n + n^2 \leq n^{2n-1} + n^2 \leq n^{2n-1} + n^{2n-1} \leq 2n^{2n-1} \leq n^{2n}$$

for all $n \geq 2$. So we get a pretty bad bound:

Theorem 0.19. *A finite subgroup of $\mathrm{GL}_n(\mathbb{C})$ has an abelian subgroup of index at most $C_n := (40n)^{n^{2n}}$.*

As a corollary, we have the following:

Theorem 0.20. *A finite subgroup of $\mathrm{GL}_n(\mathbb{C})$ has a normal abelian subgroup of index at most $C_n! := (40n)^{n^{2n}}!$.*

Proof. Pick at abelian subgroup A of index $m \leq C_n$. Then G acts on the left cosets of A via left multiplication. This gives a homomorphism from $G \to S_m$. The kernel, N, is a normal subgroup and it must be contained in A since A is the collection of elements that fix the coset A and N is the elements that fix all cosets. Clearly the index of N is at most $m!$. This gives the result. □

What's the best possible result? For a long time, the best bounds were quite astronomical (like the one in the theorem). In 2005, Collins showed that if $G \leq \mathrm{GL}_n(\mathbb{C})$ and $n \geq 71$ then G has an abelian subgroup of index at most $(n+1)!$. This is the best possible for $n \geq 71$, too! The reason for this is that an embedding of S_{n+1} into GL_n (think of the induced representation of the trivial subgroup on the subgroup isomorphic to S_n and how it decomposes into irreducibles) and any normal abelian subgroup of S_{n+1} is just the identity for $n \geq 4$.

Some subgroups of GL_2

In fact, one can classify all subgroups of GL_2, but this is a lot of work and appears in rather long papers. Let's see what can be said about subgroups of SL_2. Even here it is quite a bit of work (and involves some non-trivial geometry) to classify the subgroups. However, one can show that sufficiently large subgroups of SL_2 fall into a few nice categories. Thus one has certain infinite families of finite groups along with a few "sporadic" examples.

Exercise 45. *Let $A \leq \mathrm{SL}_2(\mathbb{C})$ be a finite abelian group. Show that A is cyclic.*

Exercise 46. *Let $G \leq \mathrm{SL}_2(\mathbb{C})$ be a finite group. Suppose that A is a non-trivial normal abelian subgroup of G. Pick $a \in A$ not equal to the identity that generates A as a group. Show that a can be diagonalized to be a diagonal matrix of the form $\mathrm{diag}(\omega, \overline{\omega})$. Show that if $x \in G$ then $xax^{-1} = a$ or $xax^{-1} = a^{-1}$. Use this to show that A must have index 1 or 2 in G.*

Exercise 47. *Show that there exists some N such that every finite subgroup of $\mathrm{SL}_2(\mathbb{C})$ of order $\geq N$ has a cyclic group of index at most 2.*

Exercise 48. *Is A_4 isomorphic to a subgroup of $\mathrm{SL}_2(\mathbb{C})$?*

Burnside's other theorem

While we're doing linear groups, we should cover another theorem of Burnside. Technically, this is a result due to Burnside and Schur. It states that if G is a finitely generated subgroup of $\mathrm{GL}_n(\mathbb{C})$ such that every element of G has finite order then G is a finite group. Burnside showed that this result holds if there exists some N such that every element of G has order at most N. Schur removed this hypothesis and replaced it by saying that every element of G has finite order (we call such a group a *torsion group*. This proof is again by induction on n with the case where $n = 1$ being trivial. We'll begin by proving Burnside's theorem and then give Schur's strengthening of the result.

Theorem 0.21. *(Burnside) Let N be a natural number and let $G \leq \mathrm{GL}_n(\mathbb{C})$ with the property that every element of G has order at most N. Then G is finite.*

We prove this by induction on n. When $n = 1$, G consists of all roots of unity of order at most N and so G is finite. Now assume the claim is true up to $n - 1$.

Let R denote the $\mathbb{C}$-subalgebra of $M_n(\mathbb{C})$ spanned by the elements of G. Then $V = \mathbb{C}^n$ is a left R-module. First suppose that V is a simple R-module. Then if $D = \mathrm{End}_R(V)$ then D is finite-dimensional over $\mathbb{C}$ and hence $D = \mathbb{C}$. Then we see that R is a dense subring of $\mathrm{End}_D(V) = M_n(\mathbb{C})$ and so $R = M_n(\mathbb{C})$. If follows that there exist $g_1, \ldots, g_{n^2}$ in G whose span is all of $M_n(\mathbb{C})$. Now let T_N equal the set of all j-th roots of unity with $j \leq N$ and let $S = \{w_1 + \cdots + w_n \colon w_i \in T_N \forall i\}$. Then S is a finite set. By assumption every eigenvalue of $g \in G$ is in T_N and so $\mathrm{Tr}(g) \in S$ for all $g \in G$. Now for $g \in G$, we have $gg_i \in G$ for $i = 1, \ldots, n^2$ and so $\mathrm{Tr}(g) \in S$. Consider the map $\phi : G \to S^{n^2}$ given by $g \mapsto (\mathrm{Tr}(gg_1), \ldots, \mathrm{Tr}(gg_{n^2}))$. Since S is finite, S^{n^2} is finite. Thus it suffices to show that ϕ is injective. Notice that if $\phi(g) = \phi(g')$ then $\mathrm{Tr}(gg_i) = \mathrm{Tr}(g'g_i)$ for $i = 1, \ldots, n^2$. Thus $h := g - g' \in M_n(\mathbb{C})$ satisfies $\mathrm{Tr}(hg_i) = 0$ for $i = 1, \ldots, n^2$. Since the g_i span $M_n(\mathbb{C})$, linearity of the trace function shows that $\mathrm{Tr}(hA) = 0$ for all $A \in M_n(\mathbb{C})$.

Exercise 49. *Show that if $A \in M_n(\mathbb{C})$ and $\mathrm{Tr}(AB) = 0$ for every $n \times n$ matrix B then $A = 0$.*

We then see that $h = 0$ and so $g = g'$. Thus G is finite. In fact, we even get a bound on G. Since T_N has size at most $1 + 2 + \cdots + N \leq N^2$ and S is sums of n elements of T_N, we see that S has size at most $\binom{N^2}{n} \leq N^{2n}$. Finally, $|G| \leq |S|^{n^2} \leq N^{2n^3}$. So we see more:

Theorem 0.22. *Let N be a natural number and let $G \leq \mathrm{GL}_n(\mathbb{C})$ with the property that every element of G has order at most N. Then $G \leq N^{2n^3}$.*

Schur extended this result a few years later by showing that any finitely generated torsion subgroup G of $\mathrm{GL}_n(\mathbb{C})$ is finite. He did so, by showing that there is some N such that all elements of G have order at most N and then invoking Burnside's theorem. We will give an "elementary" proof, but the "right" proof (which is harder) uses some results from Commutative algebra covered in 446 (Noether normalization, integral extensions, and the Nullstellensatz). You can see these results here:

http : //www.math.uwaterloo.ca/ ∼ jpbell/446lecturenotes.pdf

This harder proof is the "right" proof, because it is a standard set of tools that can often be used in extending results from number fields to arbitrary fields of characteristic zero. Let's begin with an observation due to Burnside.

Theorem 0.23. *(Burnside) Let G be a torsion subgroup of $\mathrm{GL}_n(\mathbb{Q})$. Then every element of G has order at most $(2n^2)!$ and so $|G| \leq (2n^2)!^{2n^3}$*

We note that the best bound is that $|G| \leq 2^n n!$, which occurs with the signed permutation matrices.

Proof. Let $A \in G$ be a matrix of finite order and let $q(x)$ be its minimal polynomial over $\mathbb{Q}$. Then $q(x)$ is a polynomial with rational coefficients, all roots of $q(x)$ are roots of unity, and $\deg(q(x)) \leq n$ since $q(x)$ divides the characteristic polynomial. Thus if ω is a root of $q(x)$ then $[\mathbb{Q}(\omega) : \mathbb{Q}]$. Now we need the following elementary fact which is not done in this course: if ω is a primitive d-th root of unity then $[\mathbb{Q}(\omega) : \mathbb{Q}] = \phi(d)$, where ϕ is Euler's ϕ function. This is the only fact from outside of this course we will use. It is not hard to prove, either. Now we'll use an elementary estimate. We recall that $\phi(d) = d \cdot \prod_{p|d}(1 - 1/p)$. Now the number d can have at most $1 + \log_3(d/2)$ distinct prime factors, since each prime factor other than 2 is at least 3. Thus

$$\phi(d) = d \cdot \prod_{p|d}(1 - 1/p) \geq d(1 - 1/2)(1 - 1/3)^{\log_3(d/2)} \geq d/2 \cdot (2/3)^{\log_3(d/2)} = 2^{\log_3(d/2)} = (d/2)^{\alpha},$$

where $\alpha = \log_3(2) > 1/2$, since $\sqrt{3} < 2$. So in particular, since we must have $\phi(d) < n$ we need

$$(d/2)^{\alpha} \leq n \implies (d/2) \leq n^{1/\alpha} \implies d \leq 2n^{1/\alpha} < 2n^2.$$

It follows that every eigenvalue of $A \in G$ is a d-th root of unity for some $d \leq 2n^2$ and so every element of G has order at most $N := (2n^2)!$. □

Notice we have shown the following interesting fact:

FACT: there is a constant D_n, depending only on n, such that a matrix of finite order in $\mathrm{GL}_n(\mathbb{Q})$ has order dividing D_n.

Now Schur found a way to get the complex case (for finitely generated groups) from Burnside's work. We'll give a different proof than the "standard" one. This proof has the advantage of being more elementary, but the techniques used are less applicable to similar problems. Just for fun, we'll do it for all fields of characteristic zero. Before we begin, we recall the following important fact.

FACT: Every field K of characteristic zero contains a copy of $\mathbb{Q}$ and is thus a field extension of $\mathbb{Q}$.

The reason for this is that we see that we contain a copy of $\mathbb{N}$ by looking at 1_K, $2 := 1_K + 1_K, \ldots$. Then by looking at additive inverses, we get a copy of $\mathbb{Z}$; then we look at multiplicative inverses and products and we get a copy of $\mathbb{Q}$.

We recall that a field extension K of F is *finitely generated* if there exist $a_1, \ldots, a_m \in K$ such that the smallest subfield of K containing F and $a_1, \ldots, a_m$ is equal to K. For example, $\mathbb{Q}(x, y)$ is finitely generated over $\mathbb{Q}$, since we just need to adjoin x and y to $\mathbb{Q}$.

Lemma 0.36. *Let K be a field of characteristic zero and let G be a finitely generated subgroup of $\mathrm{GL}_n(K)$. Then there exists a finitely generated extension of $\mathbb{Q}$, K_0, with $K_0 \leq K$, such that $G \leq \mathrm{GL}_n(K_0)$.*

Proof. Let $A_1, \ldots, A_m$ be a set of generators for G. Let $a_{i,j}^l$ be the (i, j)-entry of A_l and let $b_{i,j}^l$ be the (i, j)-entry of A_l^{-1}. We let K_0 be the extension of $\mathbb{Q}$ generated by these $2mn^2$ entries. Then we have $A_l^{\pm 1} \in \mathrm{GL}_n(K_0)$ by construction. We claim that $G \leq \mathrm{GL}_n(K_0)$. To do this, we note that every element of G can be written as a finite product of elements from $A_1, \ldots, A_m, A_1^{-1}, \ldots, A_m^{-1}$. Since $\mathrm{GL}_n(K_0)$ is closed under multiplication we then get the result. □

Now we need to know a bit about finitely generated extensions of $\mathbb{Q}$. To do this, we say a few words about transcendence bases.

Transcendence bases

Let K be a field extension of F. We say that elements $a_1, \ldots, a_r \in K$ are *algebraically dependent* over F if there exists a nonzero polynomial $P(x_1, \ldots, x_r) \in F[x_1, \ldots, x_r]$ such that $P(a_1, \ldots, a_r) = 0$; if they are not algebraically dependent, we say they are *algebraically independent* over F. Now we'll assume that $K = F(a_1, \ldots, a_m)$ is a finitely generated extension of F. Then we'll pick a subset $S \subset \{a_1, \ldots, a_m\}$ that is maximal with respect to being algebraically independent over F. By relabelling, we may assume that $S = \{a_1, \ldots, a_r\}$ with $0 \leq r \leq m$ (and $r = 0$ meaning each a_i is algebraic over F). Then we let $E = F(a_1, \ldots, a_r)$. Since $a_1, \ldots, a_r$ are algebraically independent over F, by definition, we have $E \cong F(t_1, \ldots, t_r)$ where the t_i are indeterminates. Now $K = E(a_{r+1})(a_{r+2}) \cdots (a_m)$. Notice that by construction each a_i is algebraic over E for $i = r + 1, \ldots, m$. Hence

$$[K : E] = [K : E(a_{r+1})(a_{r+2}) \cdots (a_{m-1})][E(a_{r+1})(a_{r+2}) \cdots (a_{m-1}) : E(a_{r+1}) \cdots (a_{m-2})] \cdots [E(a_{r+1}) : E] \leq \prod_{i>r}[E(a_i) : E] < \infty.$$

Thus we have the following theorem.

Theorem 0.24. *Let K_0 be a finitely generated extension of $\mathbb{Q}$. Then K_0 is isomorphic to a finite extension of $E = \mathbb{Q}(t_1, \ldots, t_r)$ for some nonnegative integer r.*

Now we have the following observation.

Proposition 0.37. *Let K_0 be a finitely generated extension of $\mathbb{Q}$ and let $G \leq \mathrm{GL}_n(K_0)$. Then there exists a natural number N and a natural number r such that G is isomorphic to a subgroup of $\mathrm{GL}_N(\mathbb{Q}(t_1, \ldots, t_r))$.*

Proof. We gave that G is a subgroup of $\mathrm{GL}_n(K_0)$. This means that $V := K_0^{n\times 1}$ is a faithful left $K_0[G]$-module. Now we know that K_0 is isomorphic to a finite extension of $E = \mathbb{Q}(t_1, \ldots, t_r)$ for some r. Now since $E[G]$ is a subring of $K_0[G]$, we see that V is also a left $E[G]$-module and we may regard V as an E-vector space. Since $[K_0 : E] < \infty$, we see that E has dimension $N := n[K_0 : E] < \infty$ as an E vector space and so $V \cong E^{N\times 1}$ as an E-vector space. This gives us a map from G into $\mathrm{GL}_N(E)$. Now one might worry that this map is not injective, but this is not a problem: if g acts trivially on $E^{N\times 1}$ then really this is the same as acting trivially on $V = K_0^{N\times 1}$, and so the action is faithful. □

So from the results we have shown, we see that it suffices to prove that a finitely generated torsion subgroup of $\mathrm{GL}_N(\mathbb{Q}(t_1, \ldots, t_r))$ to get the result for linear groups over all fields of characteristic zero. The way this goes is that we first reduce to finitely generated extensions of $\mathbb{Q}$ and then from there we use the preceding proposition to reduce to purely transcendental extensions of $\mathbb{Q}$. Now we're getting close. The last piece of the puzzle is the following lemma, which requires the following exercise.

Exercise 50. *Let $c(t_1, \ldots, t_r), b(t_1, \ldots, t_r) \in \mathbb{Q}[t_1, \ldots, t_r]$ with $b \neq 0$. Suppose that there is a constant C such that for every $(\alpha_1, \ldots, \alpha_r) \in \mathbb{Q}^r$ with $b(\alpha_1, \ldots, \alpha_r) \neq 0$ we have $c(\alpha_1, \ldots, \alpha_r)/b(\alpha_1, \ldots, \alpha_r) = C$. Show that $c(t_1, \ldots, t_r)/b(t_1, \ldots, t_r) = C$. (Hint: prove this by induction on r—when $r = 1$, we have that $c(t) - Cb(t)$ has infinitely many roots and so it is identically zero.)*

Lemma 0.38. *Let $A \in \mathrm{GL}_N(\mathbb{Q}(t_1, \ldots, t_r))$ be a matrix of finite order. Then there is some constant D_N, depending only on N, such that $A^{D_N} = I$.*

Proof. Suppose that A has order M. By picking a common denominator for the entries of A, we write the entries of A as $a_{i,j}(t_1, \ldots, t_r)/b(t_1, \ldots, t_r)$, where the $a_{i,j}$ and b are in $\mathbb{Q}[t_1, \ldots, t_r]$. Now we think of A as a matrix-valued function in the variables $t_1, \ldots, t_r$. Thus given $(\alpha_1, \ldots, \alpha_r) \in \mathbb{Q}^r$ with $b(\alpha_1, \ldots, \alpha_r) \neq 0$ we can *specialize* the variables t_i at α_i and obtain a matrix in $\mathrm{GL}_N(\mathbb{Q})$. $A(\alpha_1, \ldots, \alpha_r) := (a_{i,j}(\alpha_1, \ldots, \alpha_r)/b(\alpha_1, \ldots, \alpha_r))$. Notice that if A has finite order, then so will $A(\alpha_1, \ldots, \alpha_r)$. By the above fact, there is a constant D_N such that $A(\alpha_1, \ldots, \alpha_r)^{D_N} = I$ for all valid specializations $(\alpha_1, \ldots, \alpha_r) \in \mathbb{Q}^r$. We claim that we must have $A^{D_N} = I$. To see this, observe that $A^{D_N} = \left(c_{i,j}(t_1, \ldots, t_r)/b(t_1, \ldots, t_r)^{D_N}\right)$ for some polynomials $c_{i,j}$ and b as before. Then given $(\alpha_1, \ldots, \alpha_r) \in \mathbb{Q}^r$ with $b(\alpha_1, \ldots, \alpha_r) \neq 0$ we have shown that

$$c_{i,j}(\alpha_1, \ldots, \alpha_r)/b(\alpha_1, \ldots, \alpha_r)^{D_N} = \delta_{i,j}.$$

The result now follows from the exercise, because we see that $A^{D_N} = I$. □

Now we get Schur's result.

Theorem 0.25. *Let K be a field of characteristic zero and let G be a finitely generated torsion subgroup of $\mathrm{GL}_n(K)$. Then G is finite.*

Proof. We first saw that we could reduce to the case where $G \leq \mathrm{GL}_n(K_0)$ with K_0 a finitely generated extension of $\mathbb{Q}$. From there, we saw we could reduce to $G \leq \mathrm{GL}_N(\mathbb{Q}(t_1, \ldots, t_r)$. But here we saw that there was some constant D_N such that every element of G has order at most D_N. The result then follows from Burnside's original result. □

Unipotent groups and the Lie-Kolchin theorem

We'll now give an important result in the understanding of linear groups. This is the Lie-Kolchin theorem, which states that unipotent linear groups are simultaneously triangularizable. This can be quickly proved using the Borel fixed point theorem, but we will give a more algebraic proof using a beautiful result of Wedderburn. We've already seen two theorems of Wedderburn: the theorem on finite division rings and the theorem (due to Artin and Wedderburn) on the structure of Artinian rings with no nonzero nil ideals. Now we'll do another great theorem about simultaneous triangularization of nilpotent semigroups. We recall that a matrix $A \in M_n(\mathbb{C})$ is *nilpotent* if $A^n = 0$. Let's have some fun.

Theorem 0.26. *Let K be a field of characteristic zero. Then $A \in M_n(K)$ is nilpotent if and only if $\mathrm{Tr}(A^d) = 0$ for $d = 1, 2, \ldots$.*

Proof. Let's do a fun proof first. Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of A with multiplicity. Suppose that $\lambda_1, \ldots, \lambda_r \neq 0$ and the rest are zero. Then as formal power series, we have

$$\sum_{i=1}^{r} 1/(1 - \lambda_i x) = r,$$

which we get by expanding geometric series. Now multiply by $(1 - \lambda_1 x) \cdots (1 - \lambda_r x)$. Notice that if $r > 0$ then the left-hand side is a polynomial of degree $r - 1$ and the right-hand side has degree r. This is impossible.

(Now ... here is the proof from Class.) If A is nilpotent then its characteristic polynomial is x^n and so its eigenvalues are all zero. Since the trace of A^d is the sum of the d-th powers of the eigenvalues, we see that the trace of A^d is zero for all d. Conversely, let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of A with multiplicity. Then if the Trace of A^d is zero for all d then we have

$$\sum_i \lambda_i^d = 0$$

for $d = 1, 2, \ldots$. Now if the λ_i are all zero, then the characteristic polynomial of A is x^n and so A is nilpotent by Cayley-Hamilton; otherwise, let $\alpha_1, \ldots, \alpha_r$ be the distinct nonzero eigenvalues of A and let $C_1, \ldots, C_r \geq 1$ denote their respective multiplicities.

Then we have

$$\sum_{j=1}^{r} C_j \alpha_j^d = 0$$

for $d = 1, 2, \ldots$. Now if you like Vandermonde arguments, then you know what to do. Otherwise, we just need to prove the following claim:

Claim: If $\alpha_1, \ldots, \alpha_r$ are distinct nonzero numbers and $C_1, \ldots, C_r$ are nonzero then

$$\sum_{j=1}^{r} C_j \alpha_j^d$$

is nonzero for some $d \geq 1$.

To prove the claim, we argue by induction on r. If $r = 1$, then it is immediate. So now suppose it is true up to $r-1$ and suppose that the above sum is zero for all d. Then we can use the equations with d and $d+1$ to get

$$\sum_{j=1}^{r} C_j(\alpha_1 \alpha_j^d - \alpha_j^{d+1}) = \sum_{j=2}^{r} C_j(\alpha_1 - \alpha_j)\alpha_j^d = 0$$

for $d = 1, 2, \ldots$. But this then gives us a contradiction. □

Let $\mathcal{S}$ be a set of nilpotent matrices in $M_n(\mathbb{C})$ that is closed under multiplication. An obvious example of such a set is all strictly upper-triangular matrices. Wedderburn's theorem says in fact that any such $\mathcal{S}$ must be conjugate to a subset of the set of strictly upper-triangular matrices.

Back when Wedderburn was proving this, he had to be particular ingenious, but today with the help of Jacobson's density theorem we can give a short proof. We remark that Jacobson was Wedderburn's student. Here is how we will do this. We'll prove this by induction on n. When $n = 1$, the only nilpotent matrix is 0, so there is nothing to prove. Now assume that $n > 1$. We'll now let $R \subseteq M_n(\mathbb{C})$ be the set of all matrices of the form $\alpha I + B$ where B is a $\mathbb{C}$-linear combination of elements of $\mathcal{S}$. Then we observe that R is a ring. Now let $V = \mathbb{C}^{n \times 1}$. Then R, being a subring of $M_n(\mathbb{C})$, acts on V and so V is a left R-module. Now if V is not simple, then we are done! We have a submodule W of V and we can look at the actions of $\mathcal{S}$ on W and V/W and then triangularize $\mathcal{S}$ on V. If V is simple then as we saw earlier we have $\Delta = \mathrm{End}_R(V)$ must be $\mathbb{C}$ and R is dense in $\mathrm{End}_{\mathbb{C}}(V)$ and so $R = M_n(\mathbb{C})$. But this is a problem, because this means that we can find $S_1, \ldots, S_{n^2}$ in $\mathcal{S}$ such that $\mathcal{X} = \{I, S_1, \ldots, S_{n^2}\}$ span all of $M_n(\mathbb{C})$. But notice that for $S \in \mathcal{S}$ we have $SX \in \mathcal{S}$ and so $\mathrm{Tr}(SX) = 0$ for all $S \in \mathcal{S}$ and all $X \in \mathcal{X}$. We then see as before that $S = 0$ for all $S \in \mathcal{S}$, which is a contradiction.

Unipotent linear groups

A matrix A is called *unipotent* if all of its eigenvalues are 1. Equivalently, $A = I + N$, where N is nilpotent. A group of matrices is *unipotent* if every matrix in the group is unipotent. Unipotent groups are one of the basic building blocks in the theory of linear groups and a fundamental result in understanding them is the Lie-Kolchin theorem.

Theorem 0.27. *(Lie-Kolchin) Let U be a unipotent subgroup of $\mathrm{GL}_n(\mathbb{C})$ then there is an invertible matrix T such that TUT^{-1} is a subgroup of the group of upper-triangular matrices with ones along the diagonal.*

Proof. We can give a quick proof using Wedderburn's theorem—this is a bit different from the classical proof. We let $\mathcal{T}$ denote the set of nilpotent matrices N such that $I + N$ is in the group U. We now let $\mathcal{S}$ denote the set of all products of elements of $\mathcal{T}$. We claim that if $S \in S$ then $\mathrm{Tr}(S) = 0$. To see this, we have $S = N_1 N_2 \cdots N_r$ for some $r \geq 1$ with $N_i \in \mathcal{T}$. We prove this claim by induction on r. When $r = 1$, the claim is immediate. So suppose that the claim holds for products of length $< r$. Then $(I + N_1)(I + N_2) \cdots (I + N_r) \in U$ and so $(I + N_1)(I + N_2) \cdots (I + N_r) - I$ is nilpotent and so it has trace zero. But notice that when we expand this out, we get that this of the form $N_1 \cdots N_r + B$, where B is a sum of products of length $< r$ of elements in $\mathcal{T}$. Thus by the claim, B has trace zero and since the sum has trace zero, we see that $\mathrm{Tr}(N_1 \cdots N_r)$ has trace zero too. We now see that every element of $\mathcal{S}$ is nilpotent since $\mathcal{S}$ is closed under multiplication and so all powers of elements have trace zero; this means that every element is nilpotent. It follows that there is some matrix T such that $T\mathcal{S}T^{-1}$ consists of strictly upper-triangular matrices. In particular, if $N \in \mathcal{T}$ then TNT^{-1} is strictly upper-triangular and so for $A \in U$ we have $A = I + N$ with $N \in \mathcal{T}$ and so we have $TAT^{-1} = T(I + N)T^{-1}$ is upper-triangular with 1s along the diagonal. □

Nilpotent groups

One of the nice consequences of the Lie-Kolchin theorem is that unipotent groups are nilpotent.

Definition 0.39. *We say that a group G is* nilpotent *if there exists a series $G = G_0 \geq G_1 \geq G_2 \geq \cdots \geq G_r = \{1\}$, where for $i = 1, \ldots, r$, $G_i = [G_{i-1}, G]$, where for subgroups H and K of G we define $[H, K]$ to be the smallest subgroup containing all $hkh^{-1}k^{-1}$ with $h \in H$ and $k \in K$.*

We contrast this with a solvable group, which says that there is a chain $G = G^{(0)} \geq G^{(1)} \geq \cdots \geq G^{(r)} = \{1\}$, where $G^{(i)} = [G^{(i-1)}, G^{(i-1)}]$. In particular, it is straightforward to show that a nilpotent group is solvable. In fact, we have the containments

$$\{\text{abelian groups}\} \subseteq \{\text{nilpotent groups}\} \subseteq \{\text{solvable groups}\}.$$

Let's show that these containments are proper.

Example 0.40. The quaternion group Q_8 is nilpotent but not abelian, the group S_3 is solvable but not nilpotent.

Proof. To see that Q_8 is nilpotent, notice that $Q_8' = [Q_8, Q_8] = \{\pm 1\}$. We already saw this. Now $\{\pm 1\}$ is central, so $[Q_8, Q_8'] = \{1\}$. To see that S_3 is not nilpotent, notice that $S_3' = [S_3, S_3] = A_3$; next $[A_3, S_3] = A_3$ because $(123)(12)(123)^{-1}(12) = (132)$, which generates A_3. On the other hand, A_3 is abelian and so $[A_3, A_3] = \{1\}$ and so S_3 is solvable. □

One of the basic results in finite group theory is that a finite group is nilpotent if and only if it is a direct product of groups of prime power order.

Proposition 0.41. *If G is nilpotent and N is normal in G then G/N is nilpotent. If H is a subgroup then H is nilpotent. We do not have that if G/N and N are nilpotent then G is nilpotent; example, S_3.*

Theorem 0.28. *A unipotent subgroup U of $\mathrm{GL}_n(\mathbb{C})$ is nilpotent.*

Proof. We let $V = \mathbb{C}^{n \times 1}$. By Lie-Kolchin, our group is simultaneously triangularizable with ones along the diagonal. This means there is a basis $v_1, \ldots, v_n$ of V such that for each $A \in U$ we have $Av_i \in v_i + \sum_{j<i} \mathbb{C}v_j$. Notice that in particular, $Av_1 = v_1$ for all $A \in U$. Now we claim that $Av_j = v_j$ for $j = 1, \ldots, i$ for each $A \in U_i$ where $U_1 = U$ and $U_j = [U_{j-1}, U]$ for $j \geq 2$. Notice that if we write $A \in U$ as $A = I + N$ then we have Nv_i is in the span of v_j for $j < i$.

We prove this by induction on i. We have seen that it is true for $i = 1$. Now suppose it holds for $i < k$. Then U_k is generated by things of the form $ABA^{-1}B^{-1}$ with $A \in U_1$ and $B \in U_{k-1}$. We write $A = I + N$ and $B = I + N'$. Then by assumption $N'v_j = 0$ for $j < k$ and $N'v_k$ is in the span of v_1. Now $ABA^{-1}B^{-1} = (I+N)(I+N')(I-N+N^2-N^3+\cdots)(1-N'+(N')^2-\cdots)$. Now $N'v_j = 0$ for $j < k$ and since $N'v_k$ is in the span of v_j with $j < k$ we see that $(N')^r v_j = 0$ for $r \geq 1$ and $j \leq k$. Similarly, since $N'v_k$ is in the span of v_1 we see that $NN'v_j = 0$ for $j \leq k$. In particular, for $j \leq k$, we have $(I + N)(I + N')(I - N + N^2 - N^3 + \cdots)(1 - N' + (N')^2 - \cdots)v_j = (I + N)(I + N')(I - N + N^2 - N^3 + \cdots)(I - N')v_j$. This becomes

$$(I + N + N' - N - N')v_j = v_j.$$

The result follows. □

Application of Artin-Wedderburn: units in rings

Let's have some more fun. Did you know that if F is a finite field then its multiplicative group F^* is a cyclic group? Let's see how to prove this. The idea is as follows. How many elements of order exactly d does F^* have? Every such element is a solution to $x^d - 1 = 0$ in F and since a polynomial of degree d in $F[x]$ has at most d roots, we see that there are at most d solutions. On the other hand, if $\alpha \in F^*$ has order exactly d then $1, \alpha, \alpha^2, \ldots, \alpha^{d-1}$ are all distinct and all roots of $x^d - 1 = 0$ and so they must be all of them. But notice that only $\phi(d)$ of these elements have order exactly d. So we conclude that for every d there are at most $\phi(d)$ elements of order d in F^*. Notice that we can only have an element of order d if $d||F^*| =: n$, so we see that $n = |F^*| \leq \sum_{d|n} \phi(d)$. Notice that $\sum_{d|n} \phi(d) = n$ (Exercise!) and so we see that we must have equality. In particular, F^* has an element of order n.

Units in rings

This leads us to an interesting question. When can a group be realized as the group of units in a ring? Notice the trivial group is the set of units $\mathbb{Z}_2$, $\mathbb{Z}_2$ is the group of units of $\mathbb{F}_3$; $\mathbb{Z}_3$ is the group of units of $\mathbb{F}_4$; $\mathbb{Z}_4$ is the group of units of $\mathbb{F}_5$; $\mathbb{Z}_2 \times \mathbb{Z}_2$ is the group of units of $\mathbb{F}_3^2$. How about $\mathbb{Z}_5$? Well...

Proposition 0.42. *There does not exist a ring R (with identity) whose group of units is isomorphic to $\mathbb{Z}_5$.*

Proof. Notice that R must have characteristic 2 or else -1 would be an element of order 2. Thus R contains a copy of $\mathbb{F}_2$. Suppose that R^* has order 5 and let $1, u, u^2, u^3, u^4$ be its elements. So $u^5 = 1$. Let $S = \mathbb{F}_2 + \mathbb{F}_2 u + \cdots + \mathbb{F}_2 u^4 \subseteq R$. Then $S^* \subseteq R^*$ but it contains u so they are equal. Notice that we have a homomorphism from $\mathbb{F}_2[x]/(x^5 - 1)$ onto S given by sending $x \mapsto u$. It is onto but it need not be 1-to-1. I claim that S has no nonzero nilpotent elements: if $P = c_0 + c_1 u + \cdots + c_4 u^4$ is nilpotent then there is some d such that $P^{16^d} = 0$. But $P^{2^d} = c_0 + c_1 u + \cdots + c^4 u^4$ by Frobenius property and the fact that $u^{16} = u$. Thus S is a commutative Artinian (finite, in fact!) ring with no nonzero nil ideals. By Artin-Wedderburn, it

is a finite direct product of finite fields. Let's write $S = \mathbb{F}_1 \times \cdots \times \mathbb{F}_r$. Then $S^* \cong \mathbb{F}_1^* \times \cdots \times \mathbb{F}_r^*$. Since S^* is cyclic of prime order, we see that $F_i^* \cong \mathbb{Z}_5$ for some i. But then $|F_i| = 6$, which is impossible. □

In fact, if we study this proof, it shows us more.

Theorem 0.29. *Let G be a group of odd order. Then G is the group of units of a ring R if and only if G is a finite product of cyclic groups of order $2^d - 1$.*

Proof. As before, R has characteristic 2 and contains $\mathbb{F}_2$. Let G be the group of units. Then we let S be the algebra generated by $\mathbb{F}_2$ and G. As before, $S^* = R^* = G$. Since the elements of G might be linearly dependent, we only know that S is a homomorphic image of the group algebra $\mathbb{F}_2[G]$. By Maschke's theorem $\mathbb{F}_2[G]$ has no nonzero nil ideals and so by Artin-Wedderburn, $\mathbb{F}_2[G]$ is isomorphic to a finite product of matrix rings over finite fields. Moreover, each such field has characteristic two since S has characteristic two. Next, any homomorphic image of a product of matrix rings over fields is again of this form (by simplicity of matrix rings) and so $S \cong \prod M_{n_i}(\mathbb{F}_{2^{d_i}})$. Then S^* is the product of the groups of units. But notice that if some $n_i > 1$ then $\mathrm{GL}_{n_i}(\mathbb{F}_{2^{d_i}})$ has an element of order 2. Namely, pick a permutation matrix coming from a transposition. Since G has odd order, we see that S is a product of fields and we get the result from earlier work on groups of units of fields. □

For groups of even order it is harder. Can we think of an example of an even group that doesn't occur? S_3? It occurs: it is the group of units of $M_2(\mathbb{F}_2)$. Q_8? It occurs too: it is the units of the quaternions over $\mathbb{Z}$. How about A_5? Now we can show something. If $R^* = A_5$ then since A_5 is centerless, $-1 = 1$ in the units group. This means that R has characteristic 2. Now let S be the subring of R generated by $\mathbb{F}_2$ and A_5. This is a finite ring. Notice that if it has a nonzero nil ideal N then all elements of the form $1 + n$ with $n \in N$ is a non-trivial normal subgroup of the units group. Since A_5 is simple, every unit must then be of the form $1 + n$. But notice that there is some k such that $n^{2^k} = 0$ and so by Frobenius, we have $(1 + n)^{2^k} = 1 + n^{2^k} = 1$; i.e., every element of the group of units of the form $1 + n$ has order a power of 2, so it cannot be A_5. We conclude that S has no nonzero nil ideals and so by Artin-Wedderburn and Wedderburn's theorem on finite division rings, it is a finite product of matrix rings over finite fields. It follows that A_5 must be isomorphic to a finite direct product of groups of the form $\mathrm{GL}_m(\mathbb{F})$, where F is a finite field of characteristic two. Since A_5 is simple and direct products of ≥ 2 non-trivial groups are not, we see that we must have $A_5 \cong \mathrm{GL}_m(\mathbb{F})$. But this cannot occur! Why? Let's say that it were. Since A_5 is nonabelian, we must have $m \geq 2$. Notice that the nonzero diagonal matrices form a proper central (and hence normal) subgroup of $\mathrm{GL}_m(\mathbb{F})$ of size $|F| - 1$. In particular, since A_5 has trivial centre, we see that $|F| - 1 = 1$ and so $F = \mathbb{F}_2$. On one of the assignments, we showed that the size of $\mathrm{GL}_m(\mathbb{F}_p)$ has size $(p^m - 1)(p^m - p) \cdots (p^m - p^{m-1})$. In particular, if $A_5 \cong \mathrm{GL}_m(\mathbb{F})$ we must have

$$60 = (2^m - 1)(2^m - 2)(2^m - 4) \cdots (2^m - 2^{m-1}).$$

Notice that 4 is the biggest power of 2 that divides 60; on the other hand $2^{\binom{m}{2}}$ is the largest power of 2 dividing the right-hand side. Thus we must have $\binom{m}{2} = 2$; i.e., $m^2 - m = 4$. But this has no solutions with $m \geq 2$.

OK. I really think we've done enough at this point. To conclude the course, let's just have some fun.

My examples of fun group theory facts

Everyone should know the Sylow theorems. It's a bit weird that this course doesn't assume them, but I guess we did all right. Let's recall what they say. A finite group G has a Sylow p-subgroup; they're all conjugate. The number of them is 1 mod p; and the number divides the order of G. We can use this for our first fun proof. Presumably, you know that A_5 is simple. But let's show it's the only simple group of order 60. Since all Sylow p-subroups are conjugate, if there is a unique Sylow p-subgroup for some prime p, it is necessarily normal. The following is what I regard as the perfect group theory exercise. It's certainly not the most important result in the theory of finite groups, but its utility lies in the fact that it shows one how to do arguments with finite groups using all the material one should know before learning representation theory.

Theorem 0.30. *Let G be a simple group of order 60. Then $G \cong A_5$.*

Proof. Let n_2, n_3, n_5 denote respectively the number of Sylow 2-, 3-, and 5-subgroups. Then $n_2 \in \{1, 3, 5, 15\}$, $n_3 \in \{1, 4, 10\}$, $n_5 \in \{1, 6\}$. Now since G is simple, we cannot have $n_5 = 1$ or else the Sylow 5-subgroup would be normal. Thus $n_5 = 6$. What does this mean? Since two distinct subgroups of order 5 must intersect trivially, it means we have exactly $24 = 6(5 - 1)$ elements of order 5. By the same argument, we see that $n_3 \in \{4, 10\}$. If $n_3 = 4$ then we have 4 Sylow 3-groups. Notice G acts transitively on this set of Sylow 3-subgroups by conjugation and this gives a non-trivial homomorphism from G into S_4. Since $|G| = 60 > 24$, there must be a non-trivial proper kernel, contradiction! G is simple! So $n_3 = 10$. Then we see that we have $10(3 - 1) = 20$ elements of order 3. That means we have accounted for the orders of 44 elements of our group. Now by the same reasoning as before n_2 cannot be 1 or 3, so it is 5 or 15.

Case I. $n_2 = 15$. Notice that if all of our Sylow 2-subgroups have pairwise trivial intersection then we have at least $15 \cdot (4 - 1) = 45$ elements of order 2 or 4, which is a problem given that we've used up 44 elements already. This means that there must be two Sylow 2-subgroups Q_1 and Q_2—both of size 4—whose intersection has size 2. Now pick $z \neq 1$ in $Q_1 \cap Q_2$. Since Q_1 and Q_2 are abelian, let $H = \{a \in G \colon [a, z] = 1\}$. Then H contains Q_1 and Q_2 and so $|H| \geq |Q_1 \cup Q_2| \geq 6$. On

the other hand, Q_1 is a subgroup of H and so 4 divides the order of H. And $|H|$ divides 60. So $|H| \in \{12, 20, 60\}$. If H has size 60 then $G = H$ and H has a non-trivial centre, so it cannot be simple. If $|H| = 20$ then G acts on left cosets of H by left multiplication and this gives a non-trivial homomorphism to S_3—the kernel is a proper normal non-trivial subgroup. So $|H| = 12$ and so H has index 5. Then we get a homomorphism from G into S_5 by looking at the action on left cosets. Since G is simple, this is injective; thus G is isomorphic to a subgroup of S_5 of size 60. Use second isomorphisms theorem to show $G \cong A_5$! (We identify G with a subgroup of S_5 of size 60. If $G \neq A_5$ then $GA_5 = S_5$ and so by second isomorphism theorem, we have $S_5/G \cong A_5/G \cap A_5$. But A_5 is simple, so $G \cap A_5 = A_5$.)

Case II. $n_2 = 5$. Act by conjugation on Sylow 2-subgroups. Same argument as before using second isomorphism theorem! □

Let's do a few more exercises. We all know a subgroup of index 2 is normal in the bigger group. Here is a generalization. Recall that we showed that if H is a subgroup of G of index d then we get a non-trivial homomorphism $G \to S_d$ by looking at the action of G on the collection of left cosets coming from left multiplication. We showed that the kernel of this map is a normal subgroup that is contained in H.

Proposition 0.43. *Let G be a finite group and suppose that p is the smallest prime that divides $|G|$. If H is a subgroup of index p then H is normal.*

Proof. We have a homomorphism from $G \to S_p$. Let N be the kernel of this map. Then G/N is isomorphic to a subgroup of S_p. Since $|G/N|$ divides $|S_p| = p!$ and it divides $|G|$ and all prime factors of $|G|$ are at least p, we see that $|G/N|$ divides $\gcd(|G|, p!) = p$. Thus N must have index p or 1 in G. Since $N \leq H$, we see that the index is p and so $N = H$. □

Let's see another result. First, let's remark that if $n \geq 5$ and H is a non-trivial proper normal subgroup of S_n then $H = A_n$. We showed this already, but it doesn't hurt to remember why. We use the second isomorphism theorem. If $H \neq A_n$ then $S_n = A_n H$ and so by the second isomorphism theorem we have $S_n/H = A_n H/H \cong A_n/(A_n \cap H)$. Now A_n is simple and $A_n \cap H$ is a proper normal subgroup of A_n and so it must be trivial. Thus $S_n/H \cong A_n$. The only way this can occur is if H has size 2. But a normal subgroup is a union of conjugacy classes. Since the identity is its own conjugacy class, we see that H must have a non-identity element z that is equal to its own conjugacy class; in particular, $z \in S_n$ is a non-trivial central element. But we know the centre of S_n is trivial for $n \geq 3$.

Proposition 0.44. *Let $n \geq 5$ and let G be a proper subgroup of S_n. If G is not A_n then $[S_n : G] \geq n$.*

Proof. Suppose that $[S_n : G] = d < n$. Then by action on left cosets we get a non-trivial map from $S_n \to S_d$ whose kernel is contained in G. Now the kernel cannot be trivial since $d < n$ and so by the above remark, it must be A_n and so A_n is a subset of G. Since G is proper, we see $G = A_n$. □

And who can forget this old chestnut?

Exercise 51. *Let G be a finite group and suppose that $\phi : G \to G$ is an automorphism of G that is its own inverse. Show that if G is non-abelian then ϕ must have a non-trivial fixed point.*

Let's see the proof. We let $f(x) = \phi(x)x^{-1}$. This map doesn't have any nice group theoretic properties, but let's observe that if ϕ has no non-trivial fixed points then it is 1-to-1. If $f(x) = f(y)$ then $\phi(x)x^{-1} = \phi(y)y^{-1}$ and so $\phi(y^{-1})\phi(x) = y^{-1}x$. That is, $\phi(y^{-1}x) = y^{-1}x$ and so $y^{-1}x = 1$ by our hypothesis. This gives $x = y$. Since G is finite that means that f is onto. Now notice that $\phi(f(x)) = x\phi(x)^{-1} = f(x)^{-1}$; that is, $\phi(f(x)) = f(x)^{-1}$. Since f is onto, we then see $\phi(t) = t^{-1}$ for all $t \in G$. But ϕ is an automorphism of G so $s^{-1}t^{-1} = (ts)^{-1} = \phi(ts) = \phi(t)\phi(s) = t^{-1}s^{-1}$ for all $s, t \in G$; that is, s and t commute for all $s, t \in G$ and so G is abelian.

OK. One more.

Proposition 0.45. *Let G be a (possibly infinite) group of order > 2. Then G has a non-trivial automorphism.*

Proof. If G is not abelian then there is some g not in the centre of G. Notice that the map $x \mapsto gxg^{-1}$ is an automorphism of G and since g is not central it is non-trivial. If G is abelian the map $x \mapsto x^{-1}$ is an automorphism. It is trivial if and only if every element of G has order 1 or 2. So we've reduced to the case where G is an abelian group with every non-identity element of G having order 2. Notice if we write G additively then we can regard G as a vector space over $\mathbb{Z}/2\mathbb{Z}$ and an automorphism of G is just an invertible linear transformation from G to itself. So now $|G| > 2$ so the dimension of G is at least 2. Now let $a, b \in G$ be linearly independent over $\mathbb{Z}/2\mathbb{Z}$. We may extend $\{a, b\}$ to a basis for G. Let f be the automorphism of G induced by sending $a \to b$, $b \to a$, and fixing all other elements of our basis. Then f is non-trivial. □

Just for fun! We've just see that only finitely many groups have trivial automorphism group. Let's try the next biggest case. Try to guess which finite groups have automorphism group isomorphic to C_2, the cyclic group of order 2. Well, I don't know the answer, but we'll figure it out now. Let's first note that if G is a group with centre Z then G/Z is isomorphic to a subgroup of $\mathrm{Aut}(G)$ with the map $G \to \mathrm{Aut}(G)$ given by $g \mapsto \Phi_g$, where $\Phi_g(x) = gxg^{-1}$. The kernel of this homomorphism is Z. So if $\mathrm{Aut}(G)$ is C_2 then either $G = Z$; i.e., G is abelian, or G/Z is cyclic of order 2. In the first case notice that by the fundamental theorem of abelian groups we must have that G is a finite product of cyclic groups, say $C_{m_1} \times \cdots \times C_{m_d}$ with each

$m_i > 1$. Now a cyclic group of order n has $\phi(n)$ automorphisms since we can send any generator to another generator. Since ϕ is multiplicative and $\phi(p^i) = p^{i-1}(p-1) > 2$ for $p \geq 5$ or $p = 3$ and $i \geq 2$ or $p = 2$ and $i > 2$, we see that $\phi(n) \geq 3$ unless $n \in \{1, 2, 3, 4, 6\}$. So since every automorphism of C_{m_i} extends to an automorphism of G we see that each $m_i \in \{2, 3, 4, 6\}$ and the number of $m_i > 2$ is at most 1. Thus $G \cong C_2^i \times C_m^j$ with $i \geq 0$, $j \in \{0, 1\}$ and $m \in \{3, 4, 6\}$. Notice that C_2^i is a $\mathbb{Z}_2$-vector space of dimension i and so $\mathrm{GL}_i(\mathbb{Z}_2)$ embeds in the automorphism group of G. We saw that this group has size $(2^i - 1)(2^i - 2) \cdots (2^i - 2^{i-1}) > 2$ if $i > 1$. Thus $G \cong C_2^i \times C_m^j$ with $i, j \in \{0, 1\}$ and $m \in \{3, 4, 6\}$. Let's look at these cases now. If $i = 0$ then $G \cong C_3, C_4, C_6$—all of these have automorphism group C_2. If $i = 1$ then $G \cong C_2 \times C_3, C_2 \times C_4, C_2 \times C_6$; since $C_2 \times C_3 = C_6$, there is nothing new here. Notice $C_2 \times C_4 = \langle x, y \rangle$ with $x^2 = y^4 = 1$. Notice that we have automorphisms from $x \mapsto xy^2$, $y \mapsto y$ and one from $x \mapsto x$, $y \mapsto y^3$. So there are at least two non-trivial automorphisms. Similarly we can do the same for $C_2 \times C_6$. So in the abelian case there are just the groups C_3, C_4, C_6.

Next case: What if $G/Z \cong C_2$? But maybe you recall the famous fact: if G/Z is cyclic then G is abelian. In particular $G = Z$ and so we've done all the cases already.

Question: (I do not know the answer ... maybe no one does). Let G be a finite group. Is the collection of finite groups H (up to isomorphism) with $\mathrm{Aut}(H) \cong G$ finite?