FINITE GROUPS AND THEIR REPRESENTATIONS

JOHN PIKE

These notes were written at Bridgewater State University in the Spring of 2023. A solid background in undergraduate linear algebra and a decent amount of mathematical maturity is assumed. It is also expected that students have had some prior exposure to group theory, but a self-contained treatment of the necessary topics therein is provided in Section 1. Much of the material is taken from *Advanced Modern Algebra* by Joseph Rotman, *Linear Representations of Finite Groups* by Jean-Pierre Serre, and *Representation Theory of Finite Groups* by Benjamin Steinberg. There are likely typos and other mistakes. All such errors are mine and corrections are greatly appreciated.

1 Groups

Before diving into our main topic, we briefly review some basic group theory in order to fix notation, record a few useful results, and reacquaint ourselves with the general flavor of the subject.

A *group* is a set G equipped with a binary operation $(a, b) \mapsto ab$ that satisfies

- $(ab)c = a(bc)$ for all $a, b, c \in G$,
- there exists an *identity* element e such that $ea = ae = a$ for all $a \in G$
- for each $a \in G$, there exists an *inverse* $a^{-1} \in G$ with $aa^{-1} = a^{-1}a = e$

We say that G is *abelian* if one also has $ab = ba$ for all $a, b \in G$.

Note that the associative law means that the product abc is unambiguously defined, and thus, by induction, so is $a_1 \cdots a_n$ for any $a_1, \ldots, a_n \in G$.

Also, the group identity is unique since $fa = af = a$ for all $a \in G$ implies $e = ef = f$.

Similarly, if $ab = ba = e$, then $b = be = b(aa^{-1}) = (ba)a^{-1} = ea^{-1} = a^{-1}$, so inverses are unique as well.

Moreover, since $a^{-1}a = aa^{-1} = e$, we see that $(a^{-1})^{-1} = a$. Since $(b^{-1}a^{-1})(ab) = b^{-1}(a^{-1}a)b = b^{-1}b = e$ and $(ab)(b^{-1}a^{-1}) = a(bb^{-1})a^{-1} = aea^{-1} = aa^{-1} = e$, we have $(ab)^{-1} = b^{-1}a^{-1}$.

Note too that if $ab = ac$, left-multiplying both sides by a^{-1} shows that $b = c$, and if $ba = ca$, right-multiplying both sides by a^{-1} shows that $b = c$.

An immediate consequence is that if $ab = e$ or $ba = e$, then $b = a^{-1}$; one only needs to verify that a purported inverse is a one-sided inverse (provided that the group structure has already been established).

Likewise, if $ab = a$ for some $a \in G$, then $b = a^{-1}ab = a^{-1}a = e$; it is enough to check that a purported identity behaves appropriately at a single element.

We will use exponential notation to denote repeated multiplication so that for all $a \in G$, $n \in \mathbb{N}$, $a^0 = e$ and $a^n = aa^{n-1}$. This extends to negative exponents by writing $a^{-n} = (a^{-1})^n$. With this convention, we have $a^{i+j} = a^i a^j$ and $(a^i)^j = a^{ij}$ for all $a \in G$, $i, j \in \mathbb{Z}$.

If G has *order* $|G| = n$, then the pigeonhole principle shows that the $n+1$ elements $a^0, a^1, \ldots, a^n$ cannot all be distinct, so there must be some $0 \leq i < j \leq n$ with $a^i = a^j$ and thus $a^{j-i} = e$. In other words, for each $a \in G$, there is some $1 \leq k \leq n$ such that $a^k = e$. The smallest such k is called the *order of* a, denoted $o(a)$.

Observe that if $a^\ell = e$ for some $\ell > 0$, then we must have $o(a) \,|\, \ell$ since $\ell = m \cdot o(a) + r$ with $m \geq 0$ and $0 < r < o(a)$ implies $a^r = (a^{o(a)})^m a^r = a^\ell = e$, contradicting the minimality of $o(a)$.

Example 1.1. For $n \in \mathbb{N}$, the set $[n]_0 := \{0, 1, 2, \ldots, n-1\}$ endowed with the operation of addition modulo n is called the *cyclic group of order* n, written $\mathbb{Z}/n\mathbb{Z}$. The identity element is 0 and the inverse of j is $n - j$.

The cyclic group of order 10 is $\mathbb{Z}/10\mathbb{Z} = \{0, 1, 2, ...9\}$ with $1 + 2 = 3$, $5 + 5 = 0$, and $6 + 8 = 4$, for example.

Example 1.2. The set $[n]_0$ does not form a group under multiplication because the multiplicative identity is 1, and there is no j with the property that $j \cdot 0 = 1$.

However, $U_n = \left\{j \in [n]_0 : ij \equiv 1 \,(\mathrm{mod}\, n) \text{ for some } i \in [n]_0\right\}$ does form a group under multiplication. The identity is 1, each element has an inverse by construction, and associativity is inherited from $\mathbb{Z}$. Multiplicative closure follows by observing that if $j, k \in U_n$, then there exist $j^{-1}, k^{-1} \in U_n \subseteq [n_0]$ which necessarily satisfy $k^{-1}j^{-1} \in [n_0]$ and $(k^{-1}j^{-1})(jk) = 1$.

If $(i,j) = d$ is the largest positive integer that divides both i and j, then Bézout's lemma equivalently characterizes d as the smallest positive integer that can be expressed as $\alpha i + \beta j$ for some $\alpha, \beta \in \mathbb{Z}$. When $d = 1$, we say that i and j are *relatively prime*. Thus if j and n are relatively prime, then there exist integers α, β with $\alpha j + \beta n = 1$ and thus $\alpha j \equiv 1 \pmod{n}$. Taking $i \in [n_0]$ to be the congruence class representative of α certifies that $j \in U_n$. Conversely, if $(j,n) > 1$, then there can be no $i \in [n_0]$ with $ij \equiv 1 \pmod{n}$ as this would imply $ij + \beta n = 1$ for some $\beta \in \mathbb{Z}$.

When p is prime, we have $U_p = [p-1]$ where we are using the notation $[n] := \{1, 2, \ldots, n\}$.

The *group of units* for $n = 10$ is $U_{10} = \{1, 3, 7, 9\}$ with $3 \cdot 3 = 9$, $7 \cdot 9 = 3$, and $9 \cdot 9 = 1$, for example.

Example 1.3. Another abelian group is the *Klein four-group*, consisting of the symbols e, a, b, c with group law encoded in the *Cayley table*

$*$	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	e	a
c	c	b	a	e

Thus every element is its own inverse and the product of any two non-identity elements is the third.

One way to realize this group is to take $e = (0,0)$, $a = (0,1)$, $b = (1,0)$, and $c = (1,1)$ where the product of two elements is their coordinatewise sum modulo 2—e.g. $ab = (0,1) + (1,0) = (1,1) = c$ and $ac = (0,1) + (1,1) = (1,0) = b$.

Another is to set $e = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $a = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, $b = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$, $c = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$, and compute products by ordinary matrix multiplication.

For instance, we find that $bc = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = a$ and $a^2 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}^2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = e$.

Example 1.4. A standard nonabelian example is the *dihedral group* of order $2n$, which has *presentation* $D_n = \langle r, s \,|\, r^n = s^2 = (sr)^2 = e \rangle$. This says that D_n is generated by the symbols r and s subject to the relations $r^n = s^2 = (sr)^2 = e$.

A consequence is that $r^k s = sr^{-k}$ for all $k \in \mathbb{N}$ since $(sr)(sr) = e = s^2$ implies $rs = s^2(rs)(rr^{-1}) = s(sr)(sr)r^{-1} = sr^{-1}$, and if $r^k s = sr^{-k}$, then $r^{k+1}s = r(r^k s) = r(sr^{-k}) = (rs)r^{-k} = (sr^{-1})r^{-k} = sr^{-(k+1)}$. This shows that $r^k s$ is self-inverse since $(r^k s)(r^k s) = (r^k s)(sr^{-k}) = r^k s^2 r^{-k} = r^k r^{-k} = e$.

The elements of D_n are thus of the form r^k or $r^k s$, $k = 0, 1, \ldots, n-1$, with multiplication given by $r^i r^j = r^{i+j}$, $r^i(r^j s) = r^{i+j}s$, $r^j(sr^i) = sr^{-j}r^i = sr^{j-i} = r^{i-j}s$, and $(r^i s)(r^j s) = sr^{-i}r^j s = sr^{j-i}s = r^{i-j}s^2 = r^{i-j}$; the addition in the exponents is performed modulo n.

We think of D_n as encoding the symmetries of a regular n-gon under rotation and reflection: r^k rotates the figure by $2\pi k/n$ radians, and s reflects it about a fixed line of symmetry. There are n lines of symmetry in total and $r^k s$ corresponds to a reflection about the $(n-k)^{\text{th}}$ from that described by s. (When n is odd, these lines of symmetry run from a vertex to the midpoint of its opposing side. When n is even, there are $n/2$ connecting opposing vertices and $n/2$ connecting opposing edges.)

Example 1.5. Perhaps the most important finite group is S_n, the *symmetric group* on n symbols, which consists of all bijections from $[n]$ to itself with function composition as the group law. The order of S_n is $n!$ since a bijection $\sigma : [n] \to [n]$ is determined by specifying one of the n possibilities for $\sigma(1)$, one of the remaining $n-1$ for $\sigma(2)$, and so forth.

We can envision a *permutation* $\sigma \in S_n$ using the *two-line notation* $\sigma = \begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix}$.

This lets us think of the product $\sigma\tau$ as having the number below j that which appears below $\tau(j)$ in σ.

For instance, if $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 2 & 1 & 5 & 3 \end{pmatrix}$ and $\tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 3 & 1 & 2 & 4 \end{pmatrix}$, then $\sigma\tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 1 & 4 & 2 & 5 \end{pmatrix}$.

Indeed, τ sends 1 to $\tau(1) = 5$ and σ sends 5 to $\sigma(5) = 3$, so $\sigma\tau$ sends 1 to $\sigma\big(\tau(1)\big) = \sigma(5) = 3$, and so on.

Of course, the top row is always fixed, so a more succinct description is given by the *one-line notation* which just records the second row: $\sigma = \sigma(1)\,\sigma(2)\,\cdots\,\sigma(n)$.

This lets us think of permutations in terms of arrangements of a deck of cards labeled $1, \ldots, n$. Namely, σ is the arrangement with the card labeled $\sigma(k)$ in the k^{th} position from the top (and thus the card labeled ℓ in position $\sigma^{-1}(\ell)$).

A more dynamical picture is that, starting with the ordered deck $e = 1\,2\,\cdots\,n$, σ moves the card that was in position $\sigma(k)$ into position k to obtain the arrangement $\sigma = \sigma(1)\,\cdots\,\sigma(n)$. The product $\sigma\tau$ then corresponds to the arrangement obtained by 'shuffling' the deck ordered according to σ in the manner specified by τ. The card in position k in the arrangement $\sigma\tau$ is labeled $\sigma\big(\tau(k)\big)$ since τ moves the card that was in position $\tau(k)$—the one labeled $\sigma\big(\tau(k)\big)$—into position k. So τ-shuffling the σ-arrangement $4\,2\,1\,5\,3$ takes the card in position $\tau(1) = 5$ (labeled $\sigma(5) = 3$) and moves it to position 1, then takes the card in position $\tau(2) = 3$ (labeled $\sigma(3) = 1$) and moves it to position 2, etc., resulting in the $\sigma\tau$-arrangement $3\,1\,4\,2\,5$.

In many instances, yet another description is to be preferred, the *cycle notation*. If $i_1, \ldots, i_r$ are distinct elements of $[n]$ and $\pi \in S_n$ satisfies $\pi(i_1) = i_2,\ \pi(i_2) = i_3, \ldots, \pi(i_{r-1}) = i_r,\ \pi(i_r) = i_1$, and $\pi(j) = j$ for $j \notin \{i_1, \ldots, i_r\}$, then we say that $\pi = (i_1\, i_2 \ldots i_r)$ is an *r-cycle*. For instance, the permutation τ above is the 5-cycle $\tau = (15423)$.

Any permutation can be factored as a product of disjoint cycles by starting with 1 and then hopping from image to image before returning and starting the process anew with the smallest element not yet visited. By way of example, σ sends 1 to 4 to 5 to 3 to 1 and sends 2 to itself, so that $\sigma = (1453)(2)$. Likewise, the permutation $\pi = 4\,2\,5\,1\,3$ factors as $(14)(2)(35)$. In both cases, 2 is a *fixed point*, and we often suppress such 1-cycles, writing $\sigma = (1453)$ and $\pi = (14)(35)$, say, though the identity is sometimes denoted (1).

Observe that if $\alpha = (i_1 \ldots i_r)$ and $\beta = (j_1 \ldots j_s)$ with $\{i_1, \ldots, i_r\} \cap \{j_1, \ldots, j_s\} = \emptyset$, then $\alpha\beta = \beta\alpha$. Indeed, if $k \notin \{i_1, \ldots, i_r, j_1, \ldots, j_s\}$, then $\alpha\big(\beta(k)\big) = \alpha(k) = k = \beta(k) = \beta\big(\alpha(k)\big)$. Otherwise, we have $\alpha\big(\beta(i_k)\big) = \alpha(i_k) = i_{k+1} = \beta(i_{k+1}) = \beta\big(\alpha(i_k)\big)$ or $\alpha\big(\beta(j_\ell)\big) = \alpha(j_{\ell+1}) = j_{\ell+1} = \beta(j_\ell) = \beta\big(\alpha(j_\ell)\big)$ with the subscript addition performed modulo r and s, respectively. This shows that rearranging the cycles in a *complete factorization* has no effect, and of course, neither does cyclically shifting the terms within a cycle.

Modulo these operations, complete factorizations are unique because if $\alpha_1 \cdots \alpha_s = \sigma = \beta_1 \cdots \beta_t$ are decompositions of σ into disjoint cycles, then for any $k \in [n]$ with $\sigma(k) \neq k$, there is exactly one α_i and one β_j which do not fix k. By relabeling the cycles if need be (which is legitimate since they commute), we can assume that $i = s$ and $j = t$. It follows that $\alpha_s^m(k) = \sigma^m(k) = \beta_t^m(k)$ for all m, hence $\alpha_s = \beta_t$, so $\alpha_1 \cdots \alpha_{s-1} = \sigma = \beta_1 \cdots \beta_{t-1}$. Continuing thusly establishes the assertion.

A nice thing about factoring permutations into disjoint cycles is that inverses are particularly easy to compute. Namely, $(i_1\, i_2 \ldots i_r)^{-1} = (i_r\, i_{r-1} \ldots i_1)$. Since disjoint cycles commute, the inverse of a product of disjoint cycles is the product of their inverses. For instance, $[(145)(2736))]^{-1} = (541)(6372) = (154)(2637)$. (Another method of computing inverses is to swap the rows in the two-line notation and then sort the columns so the entries in the first row are increasing.)

We can also factor permutations into cycles that are not disjoint. In particular, $\alpha = (i_1 \ldots i_r)$ factors as $\alpha = \beta_r \cdots \beta_2$ with $\beta_k = (i_1\, i_k)$. Indeed, if $\beta_k \cdots \beta_2 = (i_1 \ldots i_k)$, then $[\beta_{k+1}(\beta_k \cdots \beta_2)](i_k) = \beta_{k+1}(i_1) = i_{k+1}$, $[\beta_{k+1}(\beta_k \cdots \beta_2)](i_{k+1}) = \beta_{k+1}(i_{k+1}) = i_1$, and $[\beta_{k+1}(\beta_k \cdots \beta_2)](j) = [\beta_k \cdots \beta_2](j)$ for $j \neq i_k, i_{k+1}$. Since permutations are products of disjoint cycles, we see that every permutation can be written as a product of 2-cycles, or *transpositions*.

In contrast with complete factorizations, the order generally does matter in a transposition decomposition. For example, $(12)(23) = (123) \neq (321) = (23)(12)$. Also, such decompositions are not unique: In S_4 we can write $(123) = (12)(23) = (23)(13) = (13)(24)(12)(14) = \ldots$ However, the number of factors in a transposition decomposition of a given permutation always has the same parity.

One way to see this is to define an *inversion* of π as a pair (i, j) with $i < j$ and $\pi(i) > \pi(j)$. Let $N(\pi)$ be the number of inversions of π and define $\operatorname{sgn}(\pi) = (-1)^{N(\pi)}$. If $\tau = (k\,\ell)$ with $k < \ell$, then $\operatorname{sgn}(\tau\pi) = -\operatorname{sgn}(\pi)$ because $\tau\pi$ is obtained from π by swapping k and ℓ in the one-line notation, so every pair (i, j) with $i, j \notin \{k, \ell\}$ has the same inversion status as before; if $i < k < \ell < j$, then $(i, k), (i, \ell), (k, j), (\ell, j)$ do as well; if $k < i < \ell$, then (k, i) and (i, ℓ) both switch their inversion status for a net change of $(-1)^2$; and (k, ℓ) has opposite inversion status, contributing the claimed factor of -1. As $N(e) = 0$, we see that $\operatorname{sgn}(\pi) = 1$ if and only if π can be expressed as the product of an even number of transpositions.

In light of the foregoing, we say that π is *even* if $\operatorname{sgn}(\pi) = 1$ and *odd* if $\operatorname{sgn}(\pi) = -1$. If σ can be written as a product of r transpositions and π can be written as a product of s transpositions, then $\sigma\pi$ can be written as a product of $r + s$ transpositions and thus $\operatorname{sgn}(\sigma\pi) = (-1)^{r+s} = (-1)^r(-1)^s = \operatorname{sgn}(\sigma)\operatorname{sgn}(\pi)$.

Since $(i_1 \ldots i_r) = \beta_r \cdots \beta_2$ with $\beta_k = (i_1\, i_k)$, we see that every r-cycle contributes a factor of $(-1)^{r-1}$, so we can write $\operatorname{sgn}(\pi) = (-1)^{E(\pi)}$ with $E(\pi)$ the number of even length cycles in the complete factorization of π.

Example 1.6. Though our focus here is primarily on finite groups, there is a class of infinite groups that will be crucial to our study of representation theory: The *general linear group* associated with a vector space V, denoted $GL(V)$, consists of all bijective linear transformations from V to itself, and the group law is composition of mappings.

We will primarily be concerned with finite-dimensional vector spaces over $\mathbb{C}$, and if V is a complex vector space with basis $\{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$, we can identify it with $\mathbb{C}^n$ via the map $(\alpha_1\mathbf{b}_1 + \cdots + \alpha_n\mathbf{b}_n) \mapsto \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}$.

Under this identification, $GL(V) \cong GL_n(\mathbb{C})$ is the set of invertible $n \times n$ complex matrices under ordinary matrix multiplication.

(If $T : V \to V$ is linear, then $T(\alpha_1\mathbf{b}_1 + \cdots + \alpha_n\mathbf{b}_n) = \alpha_1 T(\mathbf{b}_1) + \cdots + \alpha_n T(\mathbf{b}_n) = \begin{bmatrix} T(\mathbf{b}_1) & \cdots & T(\mathbf{b}_n) \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}$.)

Since a square matrix is invertible if and only if it has a nonzero determinant and $\det(AB) = \det(A)\det(B)$, this does indeed define a group. Of course, matrix multiplication is not commutative in general, so $GL(V)$ is nonabelian when $\dim(V) > 1$.

If $\emptyset \neq H \subseteq G$ satisfies the group axioms under the inherited operation, we say that H is a *subgroup* of G, written $H \leq G$.

To check that $H \leq G$, it suffices to show that if $a, b \in H$, then $ab^{-1} \in H$.

Indeed, given $a \in H$, taking $b = a$ shows that $e = aa^{-1} \in H$; taking $a = e$, $b = a$ shows that $a^{-1} \in H$; and for any $c \in H$, taking $b = c^{-1}$ shows that $ac \in H$.

If H is finite, it's enough to check that $ab \in H$ for all $a, b \in H$ since closure ensures that $b^{-1} = b^{o(b)-1} \in H$.

Note that we always have the trivial subgroups $\{e\}$ and G itself.

Given a subset $S \subseteq G$, we write $\langle S \rangle$ for the smallest subgroup of G which contains S. (By definition, the intersection of subgroups of G is itself a subgroup, so one can unambiguously define $\langle S \rangle$ to be the intersection of all subgroups containing S.)

Example 1.7. An important example is the subgroup generated by a single element $g \in G$, defined by $\langle g \rangle = \{g, g^2, \cdots, g^m\}$ with $m = o(g)$. Note that $g^i g^j = g^k$ where $i + j = k$ in $\mathbb{Z}/m\mathbb{Z}$. As such, we say that $\langle g \rangle$ is a *cyclic subgroup* of order m.

Observe that if $(r, m) = d$, then $o(g^r) = m/d$ since $g^{ro(g^r)} = (g^r)^{o(g^r)} = e$ implies $m \,|\, ro(g^r)$ and thus $m/d \,|\, o(g^r)$, and $(g^r)^{m/d} = (g^m)^{r/d} = e$ implies $o(g^r) \,|\, m/d$. The group $\langle g \rangle$ is thus generated by any element of the form g^t with $(m, t) = 1$.

Example 1.8. Recall that permutations in S_n can be classified according to their parity. If we set $A_n = \{\sigma \in S_n : \operatorname{sgn}(\sigma) = 1\}$, then the fact that $\operatorname{sgn}(\sigma\pi) = \operatorname{sgn}(\sigma)\operatorname{sgn}(\pi)$ shows that $\operatorname{sgn}(\sigma)\operatorname{sgn}(\sigma^{-1}) = \operatorname{sgn}(e) = 1$, hence $\operatorname{sgn}(\sigma^{-1}) = \operatorname{sgn}(\sigma)$. Accordingly, if $\sigma, \tau \in A_n$, then $\operatorname{sgn}(\sigma\tau^{-1}) = \operatorname{sgn}(\sigma)\operatorname{sgn}(\tau) = 1$, hence $\sigma\tau^{-1} \in A_n$ and we conclude that $A_n \leq S_n$. (We call A_n the *alternating group*.)

Example 1.9. Subgroups of $GL_n(\mathbb{C})$ include the *unitary group* $U_n(\mathbb{C}) = \{A \in GL_n(\mathbb{C}) : A^*A = I\}$, the *special linear group* $SL_n(\mathbb{C}) = \{A \in GL_n(\mathbb{C}) : \det(A) = 1\}$, and the group of upper-triangular matrices $T_n(\mathbb{C}) = \{A \in GL_n(\mathbb{C}) : A_{ij} = 0 \text{ for } i > j\}$.

Indeed, if $A, B \in U_n(\mathbb{C})$, then $(AB^{-1})^*(AB^{-1}) = (BA^*)(AB^*) = I$. That $SL_n(\mathbb{C}) \leq GL_n(\mathbb{C})$ follows from $\det(AB) = \det(A)\det(B)$, and it is straightforward to check that products and inverses of upper-triangular matrices with nonzero diagonal terms are upper-triangular and invertible.

When $n = 1$, we often write $\mathbb{C}^* = GL_1(\mathbb{C})$ for the group of nonzero complex numbers under multiplication and $\mathbb{T} = U_1(\mathbb{C}) = \{e^{i\theta} : \theta \in [0, 2\pi)\}$ for the 'circle group' of complex numbers with unit modulus.

One can also consider matrix groups over subfields of $\mathbb{C}$. For instance, $SL_n(\mathbb{R})$ is the group of linear transformations from $\mathbb{R}^n$ to itself that preserve volume and orientation.

Given a subgroup $H \leq G$ and an element $a \in G$, we define the *left coset* $aH = \{ah : h \in H\}$ and the *right coset* $Ha = \{ha : h \in H\}$. The following example shows that left and right cosets do not necessarily agree.

Example 1.10. In S_3, the subgroup $H = \langle (12) \rangle$ has 3 left cosets:

$$H = \{(1), (12)\} = (12)H,$$
$$(13)H = \{(13), (123)\} = (123)H,$$
$$(23)H = \{(23), (132)\} = (132)H.$$

The right cosets are

$$H = \{(1), (12)\} = (12)H,$$
$$H(13) = \{(13), (132)\} = H(132),$$
$$H(23) = \{(23), (123)\} = H(123).$$

A fundamental observation about cosets is that they partition the ambient group into equally sized parts.

To see this, note that $a \sim b$ if $a^{-1}b \in H$ is an equivalence relation since $a^{-1}a = e \in H$, $a^{-1}b \in H$ implies $b^{-1}a = (a^{-1}b)^{-1} \in H$, and $a^{-1}b \in H$, $b^{-1}c \in H$ implies $a^{-1}c = (a^{-1}b)(b^{-1}c) \in H$.

The equivalence classes are the left cosets of H since $a^{-1}b \in H$ iff $b \in aH$, and they all have the same size because the map $h \mapsto ah$ is a bijection from H to aH.

Clearly $g \in gH$ for all $g \in G$ and $gH = H$ if and only if $g \in H$. (Right coset analogues of all these results hold by parallel arguments, but unless otherwise specified, "coset" means "left coset" henceforth.)

If G is finite and the distinct cosets of H are $a_1H, \ldots, a_tH$ (in which case we say that the set $\{a_1, \ldots, a_t\}$ forms a *transversal* of H in G), then the preceding shows that $|G| = |a_1H| + \cdots + |a_tH| = t\,|H|$.

The number t of distinct left cosets of H is called the *index* of H in G, denoted $[G : H] = |G| \,/\, |H|$.

(The index is defined for infinite groups as well. For instance, the additive groups $2\mathbb{Z} \leq \mathbb{Z} \leq \mathbb{R}$ have $[\mathbb{Z} : 2\mathbb{Z}] = 2$ and $[\mathbb{R} : \mathbb{Z}] = \infty$.)

An upshot of this observation is *Lagrange's theorem* that the order of any subgroup must divide the order of the group. Specializing to the subgroup $\langle g \rangle$ shows that $o(g) \,\big|\, |G|$ for all $g \in G$.

Theorem 1.2 shows that the converse of Lagrange's theorem holds for abelian groups, and Example 1.25 shows that it does not hold in general.

Example 1.11. If $|G| = p$ with p prime, then the only possible orders of subgroups of G are 1 and p. Thus for any $g \in G \setminus \{e\}$, $\langle g \rangle = G$, hence G is cyclic.

Another useful way to partition a group is conjugacy: For any group G and any $g \in G$, the map $a \mapsto gag^{-1}$ is called *conjugation* by g. This defines an equivalence relation on G via $h \sim k$ if $h = gkg^{-1}$ for some $g \in G$. Indeed, $x = exe^{-1}$; if $x = gyg^{-1}$, then $y = g^{-1}xg$; and if $x = gyg^{-1}$, $y = hzh^{-1}$, then $x = (gh)z(gh)^{-1}$.
We write $\mathrm{cl}(g)$ for the *conjugacy class* containing g.

Example 1.12. The *cycle type* of a permutation $\pi \in S_n$ is $(\lambda_1, \ldots, \lambda_k)$ if its complete factorization into cycles of nonincreasing length consists of a λ_1-cycle, followed by a λ_2-cycle, etc. For instance, $(15)(28)(3496) \in S_9$ has cycle type $(4, 2, 2, 1)$. Equivalently, we can define the cycle type of π as $[1^{\epsilon_1}2^{\epsilon_2}\cdots n^{\epsilon_n}]$ where ϵ_k is the number of k-cycles. For the sake of conciseness, we generally drop terms of the form k^{ϵ_k} whenever $\epsilon_k = 0$, so that $(15)(28)(3496)$ has cycle type $[1^1 2^2 4^1]$.
Now suppose $\beta = (i_1 \ldots i_m)$ is an m-cycle in $S_n \ni \sigma$. If $\sigma^{-1}(j) = i_r$, then $(\sigma\beta\sigma^{-1})(j) = \sigma\big(\beta(i_r)\big) = \sigma(i_{r+1})$, and if $\sigma^{-1}(j) = k \notin \{i_1, \ldots, i_m\}$, then $(\sigma\beta\sigma^{-1})(j) = \sigma\big(\beta(k)\big) = \sigma(k) = j$. It follows that $\sigma\beta\sigma^{-1}$ maps $\sigma(i_r)$ to $\sigma(i_{r+1})$ for $r = 1, \ldots, m$ and fixes all other elements of $[n]$, hence $\sigma(i_1 \cdots i_m)\sigma^{-1} = \big(\sigma(i_1) \ldots \sigma(i_m)\big)$.
As any $\pi \in S_n$ can be completely factored as a product of disjoint cycles, $\pi = \beta_1\beta_2\cdots\beta_k$ has the same cycle type as $\sigma\pi\sigma^{-1} = \sigma\beta_1\sigma^{-1}\sigma\beta_2\sigma^{-1}\cdots\sigma\beta_k\sigma^{-1}$.

Conversely, suppose that π and γ both have cycle type $(\lambda_1, \ldots, \lambda_k)$, and let $\widehat{\pi}$ and $\widehat{\gamma}$ be the permutations whose one-line notations are given by dropping the parentheses in the complete factorizations of π and γ into cycles of nonincreasing length. (To avoid ambiguity in this definition, one could adopt the convention that cycles begin with their least element and cycles of equal length are ordered lexicographically.) Then the preceding analysis shows that $\gamma = \sigma\pi\sigma^{-1}$ with $\sigma = \widehat{\gamma}\widehat{\pi}^{-1}$.

The conjugacy classes of S_n thus consist of all elements having the same cycle type, so the number of conjugacy classes is the number of partitions of n. The size of the conjugacy class consisting of all elements of cycle type $[1^{\epsilon_1}2^{\epsilon_2}\cdots n^{\epsilon_n}]$ is

$$\frac{n!}{\prod_{k=1}^{n} \epsilon_k! k^{\epsilon_k}}.$$

(There are $n!$ ways to order the numbers in $[n]$ and the cycle type determines placement of parentheses. But this overcounts since there are $\epsilon_k!$ ways to permute the k-cycles amongst themselves and k ways to cyclically shift the terms within each of the ϵ_k k-cycles.)

Suppose that G is a group with multiplication denoted by $\cdot$ and K is a group with multiplication denoted by $*$. If $\varphi : G \to K$ has the property that for all $g, h \in G$, $\varphi(g \cdot h) = \varphi(g) * \varphi(h)$, then we say that φ is a *homomorphism*.

Writing e_G and e_K for the identity elements in G and K, we have that $\varphi(e_G) = \varphi(e_G \cdot e_G) = \varphi(e_G) * \varphi(e_G)$, hence $\varphi(e_G) = e_K$. Also, for any $g \in G$, $e_K = \varphi(e_G) = \varphi(g \cdot g^{-1}) = \varphi(g) * \varphi(g^{-1})$ thus $\varphi(g^{-1}) = \varphi(g)^{-1}$, the inverse of $\varphi(g)$ in K. By induction, we see that $\varphi(g^n) = \varphi(g)^n$ for all $n \in \mathbb{Z}$.

If φ is also bijective, then it is called an *isomorphism*, and we say that G and K are *isomorphic*, written $G \cong K$. Isomorphic groups may differ as sets and in terms of other structural properties like ordering or topology, but from a group theoretic perspective, they are the same.

For instance, if $\varphi : G \to K$ is an isomorphism and $H \leq G$, then for all $k_1, k_2 \in \varphi(H)$, there are $h_1, h_2 \in H$ with $\varphi(h_i) = k_i$, so $k_1 k_2^{-1} = \varphi(k_1)\varphi(k_2)^{-1} = \varphi(h_1 h_2^{-1}) \in \varphi(H)$. That is, isomorphisms map subgroups to subgroups, necessarily of the same size.

Similarly, suppose that G is abelian and let $x, y \in K$. If $\varphi : G \to K$ is an isomorphism, then there exist $g, h \in G$ with $\varphi(g) = x$ and $\varphi(h) = y$, hence $xy = \varphi(g)\varphi(h) = \varphi(gh) = \varphi(hg) = \varphi(h)\varphi(g) = yx$, so K is abelian as well. (We generally don't bother emphasizing the different group operations, identities, etc. when it is clear from context.)

Example 1.13. If $g \in G$ has $o(g) = m$, then the map $\varphi : \mathbb{Z}/m\mathbb{Z} \to \langle g \rangle$ defined by $\varphi(j) = g^j$ is an isomorphism; the latter group is the former in disguise. Similarly, $x \mapsto e^{2\pi i x}$ shows that $[0, 1) \cong \mathbb{T}$.

If $\varphi : G \to K$ is a homomorphism, then it is routine to show that the *kernel* $\ker(\varphi) = \{g \in G : \varphi(g) = e\}$ is a subgroup of G and the *image* $\mathrm{Im}(\varphi) = \{k \in K : k = \varphi(g) \text{ for some } g \in G\}$ is a subgroup of K.

Note that φ is injective if and only if $\ker(\varphi) = \{e\}$ since $\varphi(g) = \varphi(h)$ if and only if $e = \varphi(gh^{-1})$.

By definition, surjectivity of φ is equivalent to $\mathrm{Im}(\varphi) = K$.

Example 1.14. Example 1.8 shows that sgn is a homomorphism from S_n to $\mathbb{C}^*$. The image is $U_1(\mathbb{R})$ and the kernel is A_n. Likewise, det defines a surjective homomorphism from $GL_n(\mathbb{C})$ to $\mathbb{C}^*$ with kernel $SL_n(\mathbb{C})$.

Observe that if $\varphi : G \to K$ is a homomorphism, $g \in G$, and $h \in \ker(\varphi)$, then $\varphi(ghg^{-1}) = \varphi(g)e\varphi(g)^{-1} = e$, so $ghg^{-1} \in \ker(\varphi)$ as well.

We say a subgroup $N \leq G$ is *normal* (written $N \lhd G$) if it is closed under conjugation by any element of G.

The preceding observation shows that kernels of homomorphisms are always normal subgroups.

Clearly all subgroups of abelian groups are normal as well, but Example 1.10 demonstrates that not all subgroups are normal. Indeed, if $N \lhd G$ and $h \in gN$, then there is some $y \in N$ with $h = gy = gyg^{-1}g = zg$ where $z = gyg^{-1} \in N$. This shows that $gN \subseteq Ng$, and since both have cardinality $|N|$, they must be equal. Conversely, if $gN = Ng$ for all $g \in G$, then for any $n \in N$, $g \in G$, $gn = n'g$ for some $n' \in G$, hence $gng^{-1} = n' \in N$. Thus an equivalent characterization of normality is that $gN = Ng$ for all $g \in G$.

Example 1.15. If $H \leq G$ has $[G : H] = 2$, then $G = H \bigsqcup aH$ for any $a \notin H$. Thus for any $h \in H$, $g \in G$, either $g \in H$, hence $ghg^{-1} \in H$, or $g = ak$ for some $k \in H$, hence $ghg^{-1} = (ak)h(ak)^{-1} = ah'a^{-1}$ for $h' = khk^{-1} \in H$. It follows that $ghg^{-1} \in H$ in this case as well since the alternative implies $ah'a^{-1} = ax$ for some $x \in H$, giving the contradiction that $a = x^{-1}h' \in H$. Thus we see that index 2 subgroups are necessarily normal. As a concrete example, $A_n \lhd S_n$ since it has two cosets, the even an odd permutations. (This also follows from the fact that $A_n = \ker(\text{sgn})$.)

Example 1.16. If $K \leq H \leq G$ and $K \lhd G$, then $K \lhd H$ since $gKg^{-1} = K$ for all $g \in G$ and thus for all $g \in H$. However, we may have that $K \lhd H$ and $H \lhd G$ without K being normal in H.

For instance, in A_4, the subgroup $V = \{(1), (12)(34), (13)(24), (14)(23)\}$, which is readily seen to be isomorphic to the Klein four-group, is normal since conjugation preserves cycle structure. Also, $W = \{(1), (12)(34)\}$ is normal in V because it has index 2. However, W is not normal in A_4 since, for example, $(123)W(123)^{-1} = \{(1), (14)(23)\} \neq W$.

If H happens to be a cyclic normal subgroup of G, then every subgroup $K \leq H$ is also normal in G. To see this, note that our assumptions imply that there is an $h \in G$ such that $H = \langle h \rangle$ and $K = \langle h^k \rangle$ where k is the smallest positive integer with $h^k \in K$. Normality of H shows that for any $g \in G$, $h^j \in H$, $gh^jg^{-1} = h^m$ for some $m \in \mathbb{N}_0$. Thus for any $h^{jk} \in K$, $gh^{jk}g^{-1} = (gh^jg^{-1})^k = h^{mk} \in K$.

To further explore the notion of normality, define a product on the collection of nonempty subsets of G by $AB = \{ab : a \in A, b \in B\}$. (When $A = \{a\}$ and $B \leq G$, $AB = aB$ is the left coset of B containing a.)

This operation is associative by definition of multiplication in G, and it satisfies $HH = H$ for all $H \leq G$.

If $H, K \leq G$, we might hope that $HK \leq G$ as well, but this turns out to be overly optimistic in general. For instance, in S_3, if $H = \langle (12) \rangle$ and $K = \langle (23) \rangle$, then $HK = \{(1), (12), (23), (123)\}$, which cannot be a subgroup since $4 \nmid 6$.

However, if subgroups satisfy $HK = KH$, then for all $h_1, h_2 \in H$, $k_1, k_2 \in K$, we have in obvious notation $(h_1k_1)(h_2k_2)^{-1} = h_1k_1k_2^{-1}h_2^{-1} = h_1k_3h_2^{-1} = h_1h_3k_4 = h_4k_4 \in HK$, so $HK = KH$ is a subgroup of G.

In particular, if $H, K \leq G$ with $K \lhd G$, then for any $h \in H$, $k \in K$, $k' = hkh^{-1}, k'' = h^{-1}kh \in K$, thus $hk = hkh^{-1}h = k'h$ and $kh = hh^{-1}kh = hk''$. This shows that $HK \subseteq KH$ and $KH \subseteq HK$, hence $HK \leq G$ by the previous observation.

If $H \lhd G$ too, then for any $g \in G$, $h \in H$, $k \in K$, $g(hk)g^{-1} = (ghg^{-1})(gkg^{-1}) = h'k' \in HK$, so $HK \lhd G$.

For $N \lhd G$, let G/N denote the family of left cosets of N. Then for all $a, b \in G$, $(aN)(bN) = (ab)N$, and G/N forms a group under this product.

The first claim is because $(aN)(bN) = a(Nb)N = a(bN)N = (ab)NN = (ab)N$. Note that the product does not depend on the choice of coset representatives: If $aN = a'N$ and $bN = b'N$, then $(ab)N = (a'b')N$.

Now this product formula shows G/N is closed under multiplication, which we already know to be associative. Since $(aN)N = aN$ and $N(aN) = N(Na) = Na = aN$, we see that $N = eN$ serves as the identity. Since $(aN)(a^{-1}N) = (aa^{-1})N = N$ and $(a^{-1}N)(aN) = (a^{-1}a)N = N$, the inverse of aN is $a^{-1}N$.

The group G/N is called a *quotient group*. Its order is $[G : N]$, so if G is finite, then $|G/N| = |G| / |N|$.

Observe that if $N \lhd G$, then the *natural map* $\pi : G \to G/N$ defined by $\pi(g) = gN$ defines a homomorphism with $\ker(\pi) = \{g \in G : gN = N\} = N$, providing a converse to our previous observation that kernels of homomorphisms are normal subgroups.

In fact, up to isomorphism, the homomorphic images of G are precisely its quotients by normal subgroups.

Theorem 1.1 (First Isomorphism Theorem)**.** *If $\varphi : G \to K$ is a homomorphism, then $\ker(\varphi) \lhd G$ and*

$$G/\ker(\varphi) \cong \mathrm{Im}(\varphi).$$

Proof. Normality of $N = \ker(\varphi)$ has already been established. The map $\psi : G/N \to \mathrm{Im}(\varphi)$ given by $\psi(gN) = \varphi(g)$ is well-defined because $\psi(gN) = \varphi(g) \in \mathrm{Im}(\varphi)$ for all $g \in G$, and if $gN = g'N$, then $g = g'n$ for some $n \in N$, hence $\psi(gN) = \varphi(g) = \varphi(g'n) = \varphi(g')\varphi(n) = \varphi(g') = \psi(g'N)$.

Since φ is a homomorphism and N is normal, $\psi(gNhN) = \psi(ghN) = \varphi(gh) = \varphi(g)\varphi(h) = \psi(gN)\psi(hN)$, thus ψ is a homomorphism. It is surjective because if $k \in \mathrm{Im}(\varphi)$, then $k = \varphi(g) = \psi(gN)$ for some $g \in G$, and it is injective because $e_K = \psi(gN) = \varphi(g)$ implies $g \in \ker(\varphi) = N$, hence $gN = N = e_{G/N}$. □

Remark 1.1. See the appendix for isomorphism theorems two, three, and four.

Example 1.17. Suppose that $n = 2m$ for some integer $m \geq 2$. Then the dihedral group D_n has 2 conjugacy classes of size 1, $\{1\}$ and $\{r^m\}$; $m-1$ of size 2, $\{r^{\pm k}\}$ for $k = 1, \ldots, m-1$; and 2 of size m, $\{r^{2k}s : 1 \leq k \leq m\}$, $\{r^{2k-1}s : 1 \leq k \leq m\}$. This follows from $r^i r^j = r^{i+j}$ and $r^k s = (r^k s)^{-1} = sr^{-k}$, hence

$$r^j r^k r^{-j} = r^k, \qquad (r^j s)r^k(r^j s) = r^{-k},$$

$$r^j s r^{-j} = r^{2j}s = (r^j s)s(r^j s),$$

$$r^j(rs)r^{-j} = r^{2j+1}s = (r^j s)(rs)(r^j s).$$

The proper normal subgroups of D_n are thus the cyclic groups $\langle r^d \rangle$ with $d \,|\, n$ and the dihedral groups $\langle r^2, s \rangle$ and $\langle r^2, rs \rangle$. The former is because $\langle r \rangle$ is a subgroup of index 2 and all subgroups of a cyclic normal subgroup are normal. For the latter, note that if a normal subgroup N contains a reflection, then it contains all reflections of the same parity by the conjugacy calculations above. As N also contains the identity, we must have $|N| > n/2$, hence $[D_n : N] = 2n/|N| < 4$. Since N is proper and $\langle s, rs \rangle = D_n$, it can only contain half of the reflections, and since any reflection outside of N has order 2 in D_n/N, $[D_n : N]$ must be even. It follows that $|N| = n$. Because $n/2$ of the elements are reflections, the remaining must be rotations, and taken together, they generate a cyclic subgroup of order $n/2$, the only one of which is $\langle r^2 \rangle$. As $\langle r^2, s \rangle$ and $\langle r^2, rs \rangle$ are the only subgroups satisfying these requirements and both have index 2, the assertion follows.

The fourth isomorphism theorem establishes a nice correspondence between subgroups of G/N and subgroups of G containing N that can be harnessed to prove a converse to Lagrange's theorem for abelian groups.

Theorem 1.2. *If G is a finite abelian group with order divisible by d, then G has a subgroup of order d.*

Proof. We first show that if $|G| = mp$ with p prime, then G has an element of order p. Example 1.11 establishes this claim when $m = 1$. Assume for the sake of induction that the statement holds for all integers $1 \leq m < n$, and let G be an abelian group of order np. Choose an element $a \neq e$ so that $k = o(a) > 1$. If $k = jp$, then $o(a^j) = p$. Otherwise, $H = \langle a \rangle$ is a normal subgroup and $|G/H| = n/k$ is divisible by p, so the inductive hypothesis guarantees the existence of some $bH \in G/H$ with order p. If $o(b) = m$, then we must have that $(bH)^m = b^m H = H$, so $m = \ell p$ and $o(b^\ell) = p$.

Now suppose that d divides $|G|$ and let p be a prime divisor of d. Then there is a normal subgroup $S = \langle g \rangle$ of order p so that $|G/S| = n/p$. By induction on $|G|$, G/S has a subgroup H' of order d/p and Theorem 7.5 shows that $H' = H/S$ for some $S \leq H \leq G$. Since $d/p = |H'| = |H|/p$, we have proved the claim. □

Observe that for any group G, the map $\varphi_g : G \to G$ given by $\varphi_g(h) = ghg^{-1}$ is an isomorphism for each $g \in G$ since for any $h, k \in G$, $\varphi_g(hk) = ghkg^{-1} = ghg^{-1}gkg^{-1} = \varphi_g(h)\varphi_g(k)$, $\varphi_g(g^{-1}hg) = h$, and $e = \varphi_g(h) = ghg^{-1}$ implies $h = g^{-1}eg = e$.

Isomorphisms from a group to itself are termed *automorphisms*, and the collection $\mathrm{Aut}(G)$ of automorphisms of G forms a group under function composition. Indeed, if $\varphi, \theta \in \mathrm{Aut}(G)$, then for any $x, y \in G$, there exist $g, h \in G$ with $\varphi(g) = x$ and $\varphi(h) = y$, so $\varphi^{-1}(xy) = \varphi^{-1}\big(\varphi(g)\varphi(h)\big) = \varphi^{-1}\big(\varphi(gh)\big) = gh = \varphi^{-1}(x)\varphi^{-1}(y)$ and $\varphi\big(\theta(xy)\big) = \varphi\big(\theta(x)\theta(y)\big) = \varphi\big(\theta(x)\big)\varphi\big(\theta(y)\big)$. The claim follows since inverses and compositions of bijections are bijections, and the identity map serves as $e_{\mathrm{Aut}(G)}$.

Because the conjugation maps satisfy $\varphi_g\big(\varphi_h(x)\big) = \varphi_g(hxh^{-1}) = g(hxh^{-1})g^{-1} = (gh)x(gh)^{-1} = \varphi_{gh}(x)$, we see that $\mathrm{Inn}(G) = \{\varphi_g : g \in G\}$ forms a subgroup of $\mathrm{Aut}(G)$ called the group of *inner automorphisms*. (By construction, $\varphi_g^{-1} = \varphi_{g^{-1}}$ and φ_e is the identity.) It's normal since for any $\theta \in \mathrm{Aut}(G)$, we have $\theta\varphi_g\theta^{-1}(x) = \theta(g\theta^{-1}(x)g^{-1}) = \theta(g)x\theta(g^{-1}) = \varphi_{\theta(g)}(x)$.

In general, a homomorphism is a structure preserving function. It's an isomorphism if bijective and an automorphism if the domain and range coincide.

Given a set X, we can define $\mathrm{Aut}_0(X)$ to be the collection of bijections from X to itself, which is easily seen to form a group under function composition. If $|X| = n$, we can label the elements $x_1, \ldots, x_n$, and one can check that the map $f : S_n \to \mathrm{Aut}_0(X)$ given by $f(\pi)(x_k) = x_{\pi(k)}$ is an isomorphism.

Note that if G is a group, then $\mathrm{Aut}(G) \leq \mathrm{Aut}_0(G)$ and the inclusion is generally strict.

Now a group G is said to *act* on a set X if there is a homomorphism $\Phi : G \to \mathrm{Aut}_0(X)$.

Alternatively, we can define a *group action* as a function $\varphi : G \times X \to X$ that satisfies $\varphi(e, x) = x$ and $\varphi(gh, x) = \varphi\big(g, \varphi(h, x)\big)$ for all $x \in X$, $g, h \in G$.

These definitions are seen to be equivalent under the identification $\varphi(g, \cdot) = \Phi(g)$.

(Verification of this claim is routine. For instance, $\varphi(g, \cdot) : X \to X$ is 1-1 since $\varphi(g, x) = \varphi(g, y)$ implies $x = \varphi(g^{-1}g, x) = \varphi\big(g^{-1}, \varphi(g, x)\big) = \varphi\big(g^{-1}, \varphi(g, y)\big) = \varphi(g^{-1}g, y) = y$. Similarly, if $\varphi(g, x) = y$, then $\varphi(g^{-1}, y) = \varphi\big(g^{-1}, \varphi(g, x)\big) = \varphi(g^{-1}g, x) = x$, hence $\varphi(g, \cdot)^{-1} = \varphi(g^{-1}, \cdot)$.)

In addition to abstracting fundamental properties of many common mathematical constructs in order to suggest analogies, establish sweeping results, and illuminate aspects otherwise obscured by extraneous details, a key feature of groups is that they encode symmetries via their actions on sets.

To give some examples, S_n acts on $[n]$ by permuting its elements, $\varphi(\sigma, n) = \sigma(n)$; $GL(V)$ acts linearly on V by $\varphi(T, \mathbf{x}) = T(\mathbf{x})$; and D_n acts on the vertices of a regular n-gon by rotation and reflection.

If $\Phi : G \to \mathrm{Aut}_0(X)$ is injective, we say that the action is *faithful*. Equivalently, the action $\varphi : G \times X \to X$ is faithful if $\varphi(g, x) = x$ for all $x \in X$ implies $g = e$.

A stronger property is that $\varphi(g, x) = x$ for some $x \in X$ implies $g = e$, in which case the action is called *free*.

Finally, if for every $x, y \in X$, there is a $g \in G$ with $\varphi(g, x) = y$, then the action is said to be *transitive*.

Example 1.18. Every group acts on itself by conjugation since the map $\varphi : G \times G \to G$ defined by $\varphi(g, h) = ghg^{-1}$ satisfies $\varphi(e, h) = h$ and $\varphi(gh, k) = (gh)k(gh)^{-1} = g(hkh^{-1})g^{-1} = \varphi\big(g, \varphi(h, k)\big)$.

This action is not transitive if $|G| > 1$ because, for instance, the identity is not conjugate to anything else.

It is also not free in this case since $\varphi(h, h) = h$ for all $h \in G$.

It is faithful if and only if $Z(G) = \{g \in G : gh = hg \text{ for all } h \in G\}$ consists only of the identity—for example, when $G = S_n$ with $n > 2$.

Example 1.19. G also acts on itself by (left) translation since the map $\tau : G \times G \to G$ defined by $\tau(g, h) = gh$ clearly satisfies the definition.

(Note that if $g \neq id$, then $\tau(g, hk) = ghk \neq ghgk = \tau(g, h)\tau(g, k)$, so $\tau(g, \cdot)$ does not define a homomorphism from G to itself; it is an automorphism in the category of sets, but not groups.)

This action is transitive since for any $g, h \in G$, $\tau(hg^{-1}, g) = h$, and it is free (and thus faithful) since $h = \tau(g, h) = gh$ implies $g = e$. '

If $|G| = n$, writing τ_g for the automorphism $\tau_g(k) = \tau(g, k)$, we have $\tau_g\big(\tau_h(k)\big) = \tau_g(hk) = (gh)k = \tau_{gh}(k)$, so $g \mapsto \tau_g$ is a homomorphism from G to $\mathrm{Aut}_0(G) \cong S_n$. Since the kernel is trivial, this shows that G is isomorphic to a subgroup of S_n, a fact known as *Cayley's theorem*.

Example 1.20. An action $\varphi : G \times X \to X$ induces an action of G on $Y^X = \{\text{functions from } X \text{ to } Y\}$ defined by $\widetilde{\varphi}(a, f)(x) = f\big(\varphi(a^{-1}, x)\big)$. This follows from $\widetilde{\varphi}(e, f)(x) = f\big(\varphi(e, x)\big) = f(x)$ and

$$\widetilde{\varphi}(ab, f)(x) = f\big(\varphi(b^{-1}a^{-1}, x)\big) = f\left(\varphi\big(b^{-1}, \varphi(a^{-1}, x)\big)\right) = \widetilde{\varphi}\big(b, f\big)\big(\varphi(a^{-1}, x)\big) = \widetilde{\varphi}\big(a, \widetilde{\varphi}(b, f)\big)(x).$$

Given a group action $\varphi : G \times X \to X$, the *orbit* of $x \in X$ is the subset of X defined by

$$\mathcal{O}(x) = \{y \in X : y = \varphi(g, x) \text{ for some } g \in G\}$$

Since $x \sim y$ if $y \in \mathcal{O}(x)$ is easily seen to be an equivalence relation, the orbits partition X. The action is transitive if and only if there is a single orbit.

The *stabilizer* of $x \in X$ is the subset of G defined by

$$G_x = \{g \in G : \varphi(g, x) = x\}.$$

If $g, h \in G_x$, then $x = \varphi(h^{-1}h, x) = \varphi\big(h^{-1}, \varphi(h, x)\big) = \varphi(h^{-1}, x)$ and $\varphi(gh, x) = \varphi\big(g, \varphi(h, x)\big) = \varphi(g, x) = x$, so we see that $G_x \leq G$. The action is free precisely when all stabilizers are tivial.

Example 1.21. If $\varphi : G \times X \to X$ is a group action and $y \in \mathcal{O}(x)$, then there is some $g \in G$ with $\varphi(g,x) = y$. It follows that for any $h \in G_x$, $\varphi(ghg^{-1}, y) = \varphi\big(gh, \varphi(g^{-1}, y)\big) = \varphi(gh, x) = \varphi\big(g, \varphi(h,x)\big) = \varphi(g,x) = y$, thus $ghg^{-1} \in G_y$. Likewise, if $h' \in G_y$, then $\varphi(g^{-1}h'g, x) = \varphi(g^{-1}h', y) = \varphi(g^{-1}, y) = x$.

We conclude that the stabilizer subgroups of elements in a common orbit are conjugate.

Example 1.22. If G acts on itself by conjugation, then the orbit of $x \in G$ is $\mathcal{O}(x) = \{gxg^{-1} : g \in G\} = \mathrm{cl}(x)$, and the stabilizer is $G_x = \{g \in G : gxg^{-1} = x\} = C_G(x)$, the set of elements that commute with x.

More generally, given $S \subseteq G$, we can define the *centralizer* $C_G(S) = \{g \in G : gsg^{-1} = s \text{ for all } s \in S\}$ and the *normalizer* $N_G(S) = \{g \in G : gSg^{-1} = S\}$.

One readily checks that $C_G(S) \leq N_G(S) \leq G$ for all $S \subseteq G$. If $S \leq G$, then for any $a \in N_G(S)$, $b \in C_G(S)$, $s \in S$, writing $t = a^{-1}sa \in S$ gives $(aba^{-1})s(aba^{-1})^{-1} = (ab)(a^{-1}sa)(ab)^{-1} = a(btb^{-1})a^{-1} = ata^{-1} = s$, hence $C_G(S) \lhd N_G(S)$.

We say that $Z(G) = C_G(G)$ is the *center* of G. Since $xg^{-1} = g^{-1}x$ and thus $gxg^{-1} = x$ for all $x \in Z(G)$, $g \in G$, we see that $H \leq Z(G)$ implies $H \lhd G$.

To simplify notation going forward, we write $gx = \varphi(g,x)$ when no confusion is likely to arise. With this convention, the defining properties are $ex = x$ and $(gh)x = g(hx)$.

Let G/G_x be the family of left cosets of G_x and consider the map $\phi : G/G_x \to \mathcal{O}(x)$ given by $\phi(aG_x) = ax$. This is well-defined since $aG_x = bG_x$ implies $a^{-1}b \in G_x$, so $\phi(bG_x) = bx = a(a^{-1}b)x = ax = \phi(aG_x)$, and it is surjective since $y \in \mathcal{O}(x)$ implies $y = ax = \phi(aG_x)$ for some $a \in G$. It's injective since $\phi(aG_x) = \phi(bG_x)$ implies $(a^{-1}b)x = x$ and thus $aG_x = bG_x$.

Since ϕ is a bijection, we have $|\mathcal{O}(x)| = |G/G_x| = [G : G_x]$. If G is finite, then $|\mathcal{O}(x)| = |G| \,/\, |G_x|$, a result known as the *orbit-stabilizer theorem*.

Example 1.23. Suppose that a finite group G acts on some set X. Write X/G for the collection of disjoint orbits in X, define $X^g = \big\{x \in X : gx = x\big\}$, and set $F = \big\{(g,x) \in G \times X : gx = x\big\}$. On the one hand, $|F| = \sum_{g \in G} |X^g|$, and on the other,

$$|F| = \sum_{x \in X} |G_x| = \sum_{x \in X} \frac{|G|}{|\mathcal{O}(x)|} = |G| \sum_{A \in X/G} \sum_{x \in A} \frac{1}{|\mathcal{O}(x)|} = |G| \sum_{A \in X/G} 1 = |G| \, |X/G|$$

since X is the disjoint union of its orbits and $x \in A$ implies $|\mathcal{O}(x)| = |A|$.

Combining these observations yields the *lemma that is not Burnside's*, $|X/G| = \frac{1}{|G|} \sum_{g \in G} |X^g|$. In words, the number of orbits is equal to the average number of fixed points.

We have seen that when G acts on itself by conjugation, the orbits are the conjugacy classes and the stabilizers are the centralizers. The orbit-stabilizer theorem thus implies that $|\mathrm{cl}(x)| = [G : C_G(x)]$ divides $|G|$ if the latter is finite.

Now $x \in Z(G)$ iff $\mathrm{cl}(x) = \{x\}$, so if $\{x_i\}_{i \in I}$ are representatives of the conjugacy classes having size greater than one, expressing G as the disjoint union of its conjugacy classes yields the *class equation*

$$|G| = |Z(G)| + \sum_{i \in I} |\mathrm{cl}(x_i)| \,.$$

Example 1.24. If $|G| = p^k$ for some prime p and positive integer k, then for each $x \in G \setminus Z(G)$, $C_G(x)$ is a proper subgroup of G, so $|\mathrm{cl}(x_i)| = [G : C_G(x)]$ is divisible by p. The class equation then implies that $|Z(G)|$ is divisible by p, so the center is nontrivial.

As such, when $k = 2$, we must have $|Z(G)| \in \{p, p^2\}$. If $|Z(G)| = p$, $G/Z(G)$ has order p and thus is cyclic, say $G/Z(G) = \langle hZ(G)\rangle$. This means that for any $x \in G$, $xZ(G) = h^m Z(G)$ for some m, hence $x = h^m z$ for some $z \in Z$. Thus if $x, x' \in G$, then $xx' = h^m z h^{m'} z' = z' h^m z h^{m'} = z' h^{m'} h^m z = h^{m'} z' h^m z = x' x$, so G is abelian, contradicting $Z(G)$ a proper subgroup of G. We conclude that $|Z(G)| = p^2$, hence $G = Z(G)$.

Note that while groups of prime or prime-squared order are necessarily abelian, the dihedral group D_4 is nonabelian of order 2^3.

We can also use the class equation to give a nice proof of Sylow's theorem on the existence of subgroups of prime power order. The argument is similar to and relies upon Theorem 1.2.

Theorem 1.3 (Sylow I)**.** *If p is prime and p^k divides $|G|$, then G has a subgroup of order p^k.*

Proof. We proceed by strong induction on $|G|$, observing that the $G = \{e\}$ case is vacuously true. For the inductive step, suppose first that p divides $|Z(G)|$. Then Theorem 1.2 shows that $Z(G)$ and thus G has a (necessarily normal) subgroup H of order p, and the inductive hypothesis ensures that G/H has a subgroup K of order p^{k-1}. Theorem 7.5 guarantees the existence of some $H \leq K' \leq G$ with $K'/H = K$ and $|K'| = |G| / [G : K'] = |G/H| \, |H| / [G/H : K] = |H| \, |K| = p^k$.

If p does not divide $|Z(G)|$, then p does not divide some $|\mathrm{cl}(x_i)|$, so $C_G(x_i)$ is a subgroup of order $|G| / |\mathrm{cl}(x_i)|$, which is less than $|G|$ and divisible by p^k, thus $C_G(x_i)$ contains a subgroup of order p^k by the inductive hypothesis. □

Remark 1.2. The $k = 1$ case of Theorem 1.3 shows that any group with order divisible by a prime p has an element of order p, a result known as *Cauchy's theorem*.

A (sub)group in which the order of every element is a power of a prime p is called a *p-(sub)group*.

For finite (sub)groups, this is equivalent to having order a power of p: Lagrange's theorem ensures that the order of every element divides the order of the group, and if q is a different prime divisor of the order, Cauchy's theorem gives an element of order q.

If p^k divides $|G|$, but p^{k+1} does not, we say that a subgroup of order p^k is a *p-Sylow subgroup* of G. The first Sylow theorem guarantees that such subgroups always exist, and Sylow theorems two and three give further details about their nature and number.

Our next example establishes the important fact that if $n \geq 5$, then A_n is *simple*—that is, it has no nontrivial proper normal subgroups. This implies in particular that it has no subgroup of order $|A_n| / 2$, showing that Theorem 1.2 does not extend to arbitrary groups.

Example 1.25. Let $n \geq 5$. We will establish simplicity of A_n by showing that it is generated by 3-cycles, all of which are conjugate in A_n. We then argue that a nontrivial normal subgroup of A_n contains a 3-cycle. It follows that it contains all 3-cycles and thus all of A_n.

Our proof makes use of the general observation that if $N \lhd G$, $n \in N$, and $g \in G$, then N necessarily contains the *commutator* $[g, n] := gng^{-1}n^{-1}$.

We first note that every $\sigma \in A_n$ can be factored as $\sigma = \tau_1\tau_2\cdots\tau_{2r}$ for some adjacently distinct transpositions $\tau_1, \ldots, \tau_{2r}$. For each $s \in [r]$, either $\tau_{2s-1} = (i\,j)$, $\tau_{2s} = (j\,k)$ for distinct i, j, k and thus $\tau_{2s-1}\tau_{2s} = (i\,j\,k)$, or $\tau_{2s-1} = (i\,j)$, $\tau_{2s} = (k\,\ell)$ for distinct i, j, k, ℓ and thus $\tau_{2s-1}\tau_{2s} = (i\,j)(j\,k)(j\,k)(k\,\ell) = (i\,j\,k)(j\,k\,\ell)$. This shows that every permutation in A_n can be expressed a product of 3-cycles.

To see that the 3-cycles are all conjugate in A_n, observe that for any 3-cycle β, there is some $\sigma \in S_n$ with $\sigma\beta\sigma^{-1} = (123)$ because all elements of a given cycle type are conjugate in S_n. If $\sigma \in A_n$, we are done. Otherwise, $\widetilde{\sigma} = (45)\sigma \in A_n$ and $\widetilde{\sigma}\beta\widetilde{\sigma}^{-1} = (45)(123)(45) = (123)$. (Note the we needed $n \geq 5$ for this argument to work.) The assertion follows since conjugacy is symmetric and transitive.

We now embark on an induction proof of the claim. For the base case, suppose that N is a nontrivial normal subgroup of A_5 and π is a nonidentity element of N. If π is a 3-cycle, then we are done. Otherwise, evenness implies that π is of the form $(i\,j)(k\,\ell)$ or $(i\,j\,k\,\ell\,m)$ for distinct i, j, k, ℓ, m. In the first case, N must contain $\big[(i\,j\,m), (i\,j)(k\,\ell)\big] = (i\,j\,m)(i\,j)(k\,\ell)(m\,j\,i)(i\,j)(k\,\ell) = (i\,m\,j)$, and in the second case N must contain $\big[(i\,j\,k), (i\,j\,k\,\ell\,m)\big] = (i\,j\,k)(i\,j\,k\,\ell\,m)(k\,j\,i)(m\,\ell\,k\,j\,i) = (i\,j\,\ell)$.

Now suppose that A_n is simple for some fixed $n \geq 5$. For each $i \in [n+1]$, define $H_i = \{\sigma \in A_{n+1} : \sigma(i) = i\}$ so that H_i is a subgroup of A_{n+1} that is isomorphic to A_n and thus is simple by the induction hypothesis. Let N be a nontrivial normal subgroup of A_{n+1} and choose some $\pi \in N \setminus \big\{(1)\big\}$. It suffices to show that there is an $i \in [n+1]$ with $\pi(i) = i$ since this implies that $N \cap H_i$ is a nontrivial normal subgroup of the simple group H_i and thus equals H_i. As $H_i \cong A_n$ contains a 3-cycle, N does as well, hence $N = A_{n+1}$.

To this end, let $j, k, \ell \in [n+1]$ be distinct with $\pi(j) = k$ and $\pi(k) \neq \ell$, and set $\sigma = (j\,k\,\ell) \in A_{n+1}$. Then $[\pi, \sigma] = (\pi\sigma\pi^{-1})\sigma^{-1} = \big(\pi(j)\,\pi(k)\,\pi(\ell)\big)(\ell\,k\,j)$ is not the identity and fixes all points outside of $\big\{i, j, k, \pi(i), \pi(j), \pi(k)\big\}$. Since $\pi(j) = k$, this set has cardinality less than $6 \leq n+1$, hence $[\pi, \sigma] \in N$ has a fixed point and the proof is complete.

Simple groups are important as they may be regarded as the basic building blocks for all finite groups via the Jordan-Hölder theorem, and the simplicity of A_n is at the heart of the insolvability of general polynomials by radicals addressed in Galois theory.

While we do not need to go into these details here, the process of building new groups from old will be important in what follows.

If H and K are groups, we define their *(external) direct product* $H \times K$ to be the set of all ordered pairs (h, k) with $h \in H$ and $k \in K$ equipped with the operation $(h_1, k_1)(h_2, k_2) = (h_1h_2, k_1, k_2)$. This is easily seen to define a group with identity (e_H, e_K) and inverses $(h, k)^{-1} = (h^{-1}, k^{-1})$.

Clearly $|H \times K| = |H|\,|K|$, and the map $(h, k) \mapsto (k, h)$ shows that $H \times K \cong K \times H$.

Moreover, the projection maps $\pi_1\big((h, k)\big) = h$ and $\pi_2\big((h, k)\big) = k$ define isomorphisms from the subgroups $\big\{(h, e_K) : h \in H\big\}$ and $\big\{(e_H, k) : k \in K\big\}$ to H and K, respectively. These subgroups are normal since, for instance, $(h, k)(h_0, e_K)(h, k)^{-1} = (hh_0h^{-1}, e_K)$.

Note too that if $\varphi : H \to H'$ and $\phi : K \to K'$ are isomorphisms, then the map $(h, k) \mapsto \big(\varphi(h), \phi(k)\big)$ shows that $H \times K \cong H' \times K'$.

If $N \lhd G$ and $N' \lhd G'$, then the function $\varphi : G \times G' \to (G/N) \times (G'/N')$ defined by $\varphi\big((g, g')\big) = (gN, g'N')$ is readily seen to be a surjective homomorphism with kernel $N \times N'$, so Theorem 1.1 tells us that $N \times N' \lhd G \times G'$ with $(G \times G')/(N \times N') \cong (G/N) \times (G'/N')$.

There is a related construction that allows one to factor a group as an *internal direct product* of normal subgroups: If $H, K \lhd G$ with $H \cap K = \{e\}$ and $HK = G$, then $G \cong H \times K$.

The assumption that $HK = G$ ensures that every $g \in G$ can be written as $g = hk$ for some $h \in H$, $k \in K$, and this decomposition is unique since $h_1k_1 = h_2k_2$ implies $h_2^{-1}h_1 = k_2k_1^{-1} \in H \cap K = \{e\}$. As such, the map $\varphi : H \times K \to G$ defined by $\varphi\big((h,k)\big) = hk$ is bijective. To see that it is a homomorphism, note that if $h \in H$, $k \in K$, then $[h,k] = (hkh^{-1})k^{-1} \in K$ and $[h,k] = h(kh^{-1}k^{-1}) \in H$, hence $hkh^{-1}k^{-1} = e$ or $hk = kh$. Accordingly, $\varphi\big((h_1,k_1)(h_2,k_2)\big) = \varphi\big((h_1h_2, k_1k_2)\big) = h_1h_2k_1k_2 = h_1k_1h_2k_2 = \varphi\big((h_1,k_1)\big)\varphi\big((h_2,k_2)\big)$.

Direct products can be regarded as a partial inverse to taking quotients: If G is the internal direct product of $H, K \lhd G$, then it must be the case that $K \cong G/H$ since the map $hk \mapsto k$ is a surjective homomorphism from G to K with kernel H.

Example 1.26. Suppose that $(m,n) = 1$ and denote the congruence class of $a \in \mathbb{Z}$ modulo k by $[a]_k$. Since $[ab]_k = [a]_k[b]_k$, the map $\phi : \mathbb{Z} \to (\mathbb{Z}/m\mathbb{Z}) \times (\mathbb{Z}/n\mathbb{Z})$ given by $\phi(a) = \big([a]_m, [a]_n\big)$ is a surjective homomorphism. The kernel is $mn\mathbb{Z}$ because $\phi(a) = \big([0]_m, [0]_n\big)$ if and only if $m, n \,\big|\, a$, if and only if $mn \,\big|\, a$. Therefore, the first isomorphism theorem tells us that $(\mathbb{Z}/m\mathbb{Z}) \times (\mathbb{Z}/n\mathbb{Z}) \cong \mathbb{Z}/mn\mathbb{Z}$.

Observe that if $(m,n) = d > 1$, then $(\mathbb{Z}/m\mathbb{Z}) \times (\mathbb{Z}/n\mathbb{Z})$ is not isomorphic to $\mathbb{Z}/mn\mathbb{Z}$ since every element of the former has order at most mn/d and the latter contains an element of order mn.

Example 1.27. Suppose that $H, K \leq G$ with G finite. Then $H \times K$ acts on $HK \subseteq G$ by $(h,k)x = hxk^{-1}$. This action is transitive since for any $x = h_1k_1$, $y = h_2k_2$, $(h_2h_1^{-1}, k_2^{-1}k_1)x = y$, thus there is a single orbit. Since the stabilizer of the identity is $G_e = \big\{(h,k) \in H \times K : h = k\big\} = \big\{(x,x) : x \in H \cap K\big\}$, the orbit-stabilizer theorem shows that $|HK| = |H \times K| \,/\, |G_e| = |H|\,|K| \,/\, |H \cap K|$.

One can extend the external direct product construction to more than two groups by defining $G_1 \times \cdots \times G_n$ to be the Cartesian product of $G_1, \ldots, G_k$ with coordinatewise multiplication.

Since $\big((g_1, \ldots, g_n), g_{n+1}\big) \mapsto (g_1, \ldots, g_{n+1})$ is an isomorphism from $(G_1 \times \cdots \times G_n) \times G_{n+1}$ to $G_1 \times \cdots \times G_{n+1}$, we can generalize the many of the preceding results by induction.

For example, if $H_k \lhd G_k$ for $k = 1, \ldots, n+1$, then

$$\begin{aligned}(G_1 \times \cdots \times G_{n+1})/(H_1 \times \cdots \times H_{n+1}) &\cong [(G_1 \times \cdots \times G_n) \times G_{n+1}]/[(H_1 \times \cdots \times H_n) \times H_{n+1}] \\ &\cong [(G_1 \times \cdots \times G_n)/(H_1 \times \cdots \times H_n)] \times (G_{n+1}/H_{n+1}) \\ &\cong (G_1/H_1) \times \cdots \times (G_{n+1}/H_{n+1}).\end{aligned}$$

Similarly, if $n = n_1 \cdots n_k$ with $n_1, \ldots, n_k$ pairwise coprime, then $\mathbb{Z}/n\mathbb{Z} \cong (\mathbb{Z}/n_1\mathbb{Z}) \times \cdots \times (\mathbb{Z}/n_k\mathbb{Z})$.

We say that G is the internal direct product of $H_1, \ldots, H_k \leq G$ if the map $\varphi : H_1 \times \cdots \times H_k \to G$ defined by $\varphi\big((h_1, \ldots, h_k)\big) = h_1 \cdots h_k$ is an isomorphism.

An induction argument shows that this is true if and only if $G = H_1 \cdots H_k$ and for each $i \in [k]$ we have $H_i \lhd G$ and $H_i \cap (H_1 \cdots H_{i-1}) = \{e\}$.

Note that normality of the H_k's guarantees that the product of any of them is itself a subgroup.

Our immediate concern is with abelian groups, in which all subgroups are normal, and it is straightforward to establish the above assertion directly in this case:

Suppose that φ is an isomorphism. Then $G = H_1 \cdots H_k$ by surjectivity. If $h \in H_i \cap (H_1 \cdots H_{i-1}) \leq G$, then $h \in H_i$ and $h^{-1} \in H_1 \cdots H_{i-1}$, so there exist $h_1, \ldots, h_{i-1}$ with $h_j \in H_j$ and $h_1 \cdots h_{i-1} = h^{-1}$, hence $e = h_1 \cdots h_{i-1} h e \cdots e$ and we conclude that $h = e$ by injectivity of φ.

Conversely, suppose $G = H_1 \cdots H_k$ and $H_i \cap (H_1 \cdots H_{i-1}) = \{e\}$ for all i. The first condition implies φ is surjective. Now observe that if $h_1 \cdots h_k = e$, then $h_1 \cdots h_{k-1} = h_k^{-1}$, hence $h_k \in H_k \cap (H_1 \cdots H_{k-1})$, so the second condition implies $h_k = e$. This in turn implies $e = h_1 \cdots h_{k-1}$, thus $h_1 \cdots h_{k-2} = h_{k-1}^{-1}$, so $h_{k-1} \in H_{k-1} \cap (H_1 \cdots H_{k-2}) = \{e\}$. Continuing thusly gives $h_1 = \cdots = h_k = e$ and we conclude that φ is injective. Finally, abelianity gives $\varphi\big((h_1, \ldots, h_k)(h_1', \ldots, h_k')\big) = \varphi\big((h_1 h_1', \ldots, h_k h_k')\big) = h_1 h_1' \cdots h_k h_k' = h_1 \cdots h_k h_1' \cdots h_k' = \varphi\big((h_1, \ldots, h_k)\big)\varphi\big((h_1', \ldots, h_k')\big)$.

Example 1.28. Let G be abelian of order $n = p_1^{e_1} \cdots p_k^{e_k}$ with $p_1 < \cdots < p_k$ prime and $e_1, \ldots, e_k \geq 1$. Sylow's theorem ensures that for each $j \in [k]$, there is a subgroup $A_j \leq G$ of order $p_j^{e_j}$. In fact, A_j is the only subgroup of this order since it's normal and all p_j-Sylow subgroups are conjugate. Writing $n_j = p_1^{e_1} \cdots p_{j-1}^{e_{j-1}}$, we have $|A_i| = p_i^{e_i} \,\big|\, n_j$ for each $i < j$, so if $g \in A_1 \cdots A_{j-1}$, then $g^{n_j} = h_1^{n_j} \cdots h_{j-1}^{n_j} = e$ and thus $o(g) \,\big|\, n_j$. Since any $h \in A_j \setminus \{e\}$ has order $p_j^{\alpha(h)}$ for some $\alpha(h) \geq 1$, $o(h) \nmid n_j$. This shows that $A_j \cap (A_1 \cdots A_{j-1}) = \{e\}$.

Now $|A_1 A_2| = |A_1|\,|A_2| \,/\, |A_1 \cap A_2| = |A_1|\,|A_2|$ and if $|A_1 \cdots A_{j-1}| = |A_1| \cdots |A_{j-1}|$, then $|A_1 \cdots A_j| = |A_1 \cdots A_{j-1}|\,|A_j| \,/\, |(A_1 \cdots A_{j-1}) \cap A_j| = |A_1 \cdots A_{j-1}|\,|A_j| = |A_1| \cdots |A_j|$. It follows that $A_1 \cdots A_k \leq G$ has order $|A_1 \cdots A_k| = |A_1| \cdots |A_k| = |G|$, hence $A_1 \cdots A_k = G$.

We have thus proved that G is the internal direct product of $A_1, \ldots, A_k$.

If G is a finite abelian group and p is a prime divisor of $|G|$, we call $G_p = \{x \in G : x^{p^n} = e \text{ for some } n \geq 1\}$ the *p-primary component* of G. Clearly G_p is a p-subgroup of G containing the unique p-Sylow subgroup A_p. Since the latter is a p-subgroup of maximal order, we have $G_p = A_p$. Accordingly, the factorization of an abelian group as a direct product of its Sylow subgroups is often called the *primary decomposition* of G.

Note that if G and G' are abelian groups and $\varphi : G \to G'$ is an isomorphism, then for each primary subgroup $G_p \leq G$, $\varphi(G_p)$ is a subgroup of G' having the same order, hence the restriction of φ to G_p provides an isomorphism between G_p and G_p'. Conversely, if $G_p \cong G_p'$ for each p dividing $|G| = |G'|$, then $G \cong G_{p_1} \times \cdots \times G_{p_k} \cong G_{p_1}' \times \cdots \times G_{p_k}' \cong G'$. That is, primary decompositions of finite abelian groups are unique up to isomorphism.

Example 1.29. Suppose that G is abelian of order p^n for some prime p and positive integer n. We will prove by induction on n that if $a \in G$ is an element of maximal order, then there is some $K \lhd G$ with $G \cong \langle a \rangle \times K$. When $n = 1$, G is cyclic so we can take a to be a generator and $K = \{e\}$. Suppose then that the claim holds for all $1 \leq k < n$, and let $a \in G$ be of maximal order, say $o(a) = p^m$. This guarantees that $g^{p^m} = e$ for all $g \in G$, and we can assume that $G \neq \langle a \rangle$ as the result is immediate in this case.

Let h be an element of minimal order in $G \setminus \langle a \rangle$ and set $H = \langle h \rangle$. We will prove that $\langle a \rangle \cap H = \{e\}$ by showing that $|H| = p$ because then $\langle a \rangle \cap H \leq H$ must have order 1 or p and the latter is precluded by $h \notin \langle a \rangle$. Since G is a p-group and $h \neq e$, $o(h^p) = o(h)/p < o(h)$, hence $h^p \in \langle a \rangle$ by our minimality assumption on $o(h)$. Accordingly, $h^p = a^r$ for some $r \geq 1$ and thus $(a^r)^{p^{m-1}} = (h^p)^{p^{m-1}} = h^{p^m} = e$. It follows that $o(a^r) \leq p^{m-1}$, so a^r does not generate $\langle a \rangle$, so $r = ps$ for some $s \in \mathbb{N}$. Now $g = a^{-s} h$ does not belong to $\langle a \rangle$ since this would imply that $h = a^s g \in \langle a \rangle$. Moreover, $g^p = a^{-ps} h^p = (a^r)^{-1} h^p = e$. As we have exhibited an element of order p outside of $\langle a \rangle$ and H was generated by an element outside of $\langle a \rangle$ having minimal order, we are forced to conclude that $|H| = o(h) = p$.

The next step is to show that the order of aH in G/H is $o(a) = p^m$. Indeed, G/H is a p-group, so if $|aH| < p^m$, we must have $H = (aH)^{p^{m-1}} = a^{p^{m-1}}H$, so that $a^{p^{m-1}} \in H \cap \langle a \rangle = \{e\}$, a contradiction. Since we have seen that every $g \in G$ satisfies, $g^{p^m} = e$ and thus $(gH)^{p^m} = g^{p^m}H = H$, aH has maximal order in G/H. The induction hypothesis and Theorem 7.5 thus give a subgroup K of G containing H with $G/H \cong \langle aH \rangle \times (K/H)$. It follows that for any $g \in G$, there exist $1 \leq j \leq p^m$, $k \in K$ with $gH = (aH)^j(kH) = (a^j k)H$, hence $g = a^j kh$ for some $h \in H$. As $H \leq K$, $kh \in K$, which shows that $G = \langle a \rangle K$. Also, if $b \in \langle a \rangle \cap K$, then $bH \in \langle aH \rangle \cap (K/H) = \{H\}$, so $b \in \langle a \rangle \cap H = \{e\}$, and we conclude that $G \cong \langle a \rangle \times K$.

Applying this result again gives $K \cong \langle a' \rangle \times K'$. Since the order of the second factor decreases to 1 as the procedure is iterated, we see that G is an internal direct product of cyclic subgroups of prime power order.

The orders of these subgroups are called *elementary divisors* and they are uniquely determined by G because if $H_1 \times \cdots \times H_m \cong G \cong K_1 \times \cdots \times K_n$ are decompositions into cyclic subgroups of nonincreasing prime-power orders, then it must be the case that $|H_1| = |K_1|$ is the largest order of any element in G, hence $H_1 \cong K_1$, and the other factors likewise agree by induction on $|G|$.

Since we can uniquely factor an arbitrary finite abelian group as a direct product of its Sylow subgroups and then uniquely factor each of those p-groups as a direct product of cyclic groups of prime power order, we see that every finite abelian group is isomorphic to a direct product of cyclic subgroups of prime power order, and any such decomposition consists of cyclic groups of the same size and multiplicity.

Alternatively, an abelian group G of order $p_1^{e_1} \cdots p_k^{e_k}$ can be expressed as the direct product of subgroups $A_1, \ldots, A_k$ where $|A_i| = p_i^{e_i}$. These primary subgroups in turn factor as direct products of cyclic groups $A_i \cong C_{i,1} \times \cdots \times C_{i,\ell(i)}$ where $|C_{i,j}| = p_i^{\alpha_{i,j}}$ with $\alpha_{i,1} \geq \cdots \geq \alpha_{i,\ell(i)}$ and $\alpha_{i,1} + \cdots + \alpha_{i,\ell(i)} = e_i$.

Let $\ell = \max_i \ell(i)$, set $\alpha_{i,j} = 0$ for $\ell(i) < j \leq \ell$, and form the $k \times \ell$ matrix E having (i,j)-entry $E_{i,j} = p_i^{\alpha_{i,j}}$.

Define the *invariant factors* $c_j = \prod_{i=1}^k E_{i,j}$ for $j = 1, \ldots, \ell$. Then $c_\ell \,|\, c_{\ell-1} \,|\, \cdots \,|\, c_1$ and, by Example 1.26, $(\mathbb{Z}/c_j\mathbb{Z}) \cong (\mathbb{Z}/p_1^{\alpha_{1,j}}\mathbb{Z}) \times \cdots \times (\mathbb{Z}/p_k^{\alpha_{k,j}}\mathbb{Z})$ for each j, hence $G \cong (\mathbb{Z}/c_1\mathbb{Z}) \times \cdots \times (\mathbb{Z}/c_\ell\mathbb{Z})$.

Since the elementary divisors determine the invariant factors and can also be recovered from them via prime factorization, we see that this decomposition is also unique.

We record this observation as the *fundamental theorem of finite abelian groups*.

Theorem 1.4. *If G is a finite abelian group, then there exist integers $c_1, \ldots, c_\ell$ such that $c_j \,|\, c_{j-1}$ for $j = 2, \ldots, \ell$ and $G \cong (\mathbb{Z}/c_1\mathbb{Z}) \times \cdots \times (\mathbb{Z}/c_\ell\mathbb{Z})$. Furthermore, ℓ and $c_1, \ldots, c_\ell$ are uniquely determined by G.*

Before moving on to representation theory, we briefly mention an extension of the direct product construction that is useful for factoring groups and building new ones from old.

To wit, if $K, Q \leq G$ with $K \cap Q = \{e\}$ and $KQ = G$, then we say that Q is a *complement* of K in G.

If $K \lhd G$ and Q is a complement of K in G, we say that G is a *semidirect product* of K by Q, written $G = K \rtimes Q$. (Note that we may not have $Q \lhd G$; if so, G is a direct product of K and Q.)

Since Q is a complement of K, each $g \in G$ can be uniquely expressed as $g = kq$ for some $k \in K$, $q \in Q$, and normality of K enables us to preserve this structure under multiplication by writing $(k_1q_1)(k_2q_2) = (k_1 \cdot q_1k_2q_1^{-1})(q_1q_2)$.

If $\iota : Q \to G$ is the embedding $\iota(q) = q$ and $\pi : G \to G/K$ is the natural map, then $\pi \circ \iota : Q \to G/K$ is a composition of homomorphisms and thus is a homomorphism. It is injective since $q_1K = q_2K$ implies there is some $k \in K$ with $q_1 = q_2k$ and thus $q_2^{-1}q_1 \in K \cap Q = \{e\}$. Since $|QK| = |Q|\,|K| \,/\, |Q \cap K| = |Q|\,|K|$ and thus $|G/K| = |G| \,/\, |K| = |QK| \,/\, |K| = |Q|$, $\pi \circ \iota$ must be surjective as well, hence $Q \cong G/K$.

Example 1.30. The dihedral group D_n is easily seen to be a semidirect product of the normal subgroup $\langle r \rangle \cong \mathbb{Z}/n\mathbb{Z}$ and its complement $Q = \langle s \rangle \cong \mathbb{Z}/2\mathbb{Z}$. It is not a direct product of these cyclic groups since it is nonabelian, but it is very natural to factor it into rotations and reflections.

When $n = 4$, we also have the normal subgroup $K = \langle r^2, s \rangle$ (which consists of four self-inverse elements and is thus isomorphic to the Klein four-group $V \cong (\mathbb{Z}/2\mathbb{Z})^2$), and it is easy to see that $Q = \langle rs \rangle \cong \mathbb{Z}/2\mathbb{Z}$ serves as a complement. Thus D_4 can be written as a semidirect product of $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ by $\mathbb{Z}/2\mathbb{Z}$ as well as a semidirect product of $\mathbb{Z}/4\mathbb{Z}$ by $\mathbb{Z}/2\mathbb{Z}$.

Example 1.31. Similarly, $S_n = A_n \rtimes \langle (12) \rangle$. When $n = 3$, we have $A_n \cong \mathbb{Z}/3\mathbb{Z}$, hence $S_3 \cong (\mathbb{Z}/3\mathbb{Z}) \rtimes (\mathbb{Z}/2\mathbb{Z})$. $\mathbb{Z}/6\mathbb{Z}$ is a direct product and thus a semidirect product of $(\mathbb{Z}/3\mathbb{Z})$ and $(\mathbb{Z}/2\mathbb{Z})$ as well. These groups are not isomorphic as the latter is abelian and the former is not, so we see again that semidirect products are not determined by the isomorphism classes of the factors.

Example 1.32. The *discrete Heisenberg group* consists of all integer matrices of the form $\begin{bmatrix} 1 & x & z \\ 0 & 1 & y \\ 0 & 0 & 1 \end{bmatrix}$.

Such a matrix is more succinctly represented by the 3-tuple $(x, y, z) \in \mathbb{Z}^3$, in which case the group law reads $(x, y, z)(x', y', z') = (x+x', y+y', z+z'+xy')$. Clearly $(0, 0, 0)$ is the identity and $(x, y, z)^{-1} = (-x, -y, xy-z)$. Keeping this notation, consider the subgroups $M = \big\{(x, 0, 0) : x \in \mathbb{Z}\big\}$ and $N = \big\{(0, y, z) : y, z \in \mathbb{Z}\big\}$. Since $(x, 0, 0)(x', 0, 0) = (x + x', 0, 0)$ and $(0, y, z)(0, y', z') = (0, y + y', z + z')$, we see that $M \cong \mathbb{Z}$ and $N \cong \mathbb{Z}^2$. Moreover, $M \cap N = \{(0, 0, 0)\}$, and for any $x, y, z \in \mathbb{Z}$, $(x, 0, 0)(0, y, z - xy) = (x, y, z)$ and $(x, 0, 0)(0, y, z)(x, 0, 0)^{-1} = (x, y, z + xy)(-x, 0, 0) = (0, y, z + xy)$, so $H \cong \mathbb{Z}^2 \rtimes \mathbb{Z}$.

If $G = K \rtimes Q$, then the map $\theta : Q \to \mathrm{Aut}(K)$ defined by $\theta(q) = \theta_q$ with $\theta_q(k) = qkq^{-1}$ is clearly a homomorphism, and we have $(k_1q_1)(k_2q_2) = (k_1 \cdot q_1k_2q_1^{-1})(q_1q_2) = \big(k_1\theta_{q_1}(k_2)\big)(q_1q_2)$.

This suggests that a semidirect product depends not only on the normal subgroup K and its complement Q, but also on the 'way in which K is normal in G' as determined by the conjugation action of Q on K, which helps to demystify the observation that semidirect products of isomorphic groups need not be isomorphic.

It also suggests a way to combine general groups into *external semidirect products*:

If G and H are groups and $\theta : H \to \mathrm{Aut}(G)$ is a homomorphism, define the semidirect product $G \rtimes_\theta H$ to be the set $\big\{(a, x) : a \in G, x \in H\big\}$ equipped with the group law $(a, x)(b, y) = (a\theta_x(b), xy)$ where $\theta_x = \theta(x)$. (When θ maps every $x \in H$ to the identity map on G, we recover the direct product as a special case.)

$G \rtimes_\theta H$ is certainly closed under this product, and we compute

$$[(a, x)(b, y)](c, z) = \big(a\theta_x(b), xy\big)(c, z) = \big(a\theta_x(b)\theta_{xy}(c), xyz\big),$$
$$(a, x)[(b, y)(c, z)] = (a, x)\big(b\theta_y(c), yz\big) = \big(a\theta_x\big(b\theta_y(c)\big), xyz\big),$$

so associativity follows from $\theta_x\big(b\theta_y(c)\big) = \theta_x(b)\theta_x\big(\theta_y(c)\big) = \theta_x(b)\theta_{xy}(c)$.

Now (e_G, e_H) is evidently the group identity, so the inverse of (a, x) must be the element (b, y) with $(e_G, e_H) = (a, x)(b, y) = (a\theta_x(b), xy)$. Examining the second coordinate shows that $y = x^{-1}$ and examining the first shows that $e_G = a\theta_x(b)$, so that $b = \theta_x^{-1}(a^{-1}) = \theta_{x^{-1}}(a^{-1})$; one easily checks that $\big(\theta_{x^{-1}}(a^{-1}), x^{-1}\big)(a, x) = \big(\theta_{x^{-1}}(a^{-1})\theta_{x^{-1}}(a), x^{-1}x\big) = \big(\theta_{x^{-1}}(e_G), e_H\big) = (e_G, e_H)$ as well.

Moreover, it is routine to show that $(a, x) \mapsto x$ is a surjective homomorphism from $G \rtimes_\theta H$ to H with kernel $\big\{(a, e_H) : a \in G\big\}$. We identify this kernel with G via the isomorphism $(a, e_H) \mapsto a$, and we have $\big\{(e_G, x) : x \in H\big\} \cong H$ by a parallel argument. The two subgroups clearly have only (e_G, e_H) in common, and for any $(a, x) \in G \rtimes_\theta H$, we have $(a, e_H)(e_G, x) = \big(a\theta_{e_H}(e_G), e_H x\big) = (a, x)$. (We can also write $(a, x) = \big(e_G\theta_x\big(\theta_x^{-1}(a)\big), xe_H\big) = (e_G, x)\big(\theta_x^{-1}(a), e_H\big)$.)

Finally, observe that $(e_G, x)(a, e_H)(e_G, x)^{-1} = (e_G, x)(a, e_H)(e_G, x^{-1}) = (e_G, x)(a, x^{-1}) = \big(e_G\theta_x(a), xx^{-1}\big) = \big(\theta_x(a), e_H\big)$, so θ_x corresponds to conjugation by (e_G, x) in $G \rtimes_\theta H$.

Remark 1.3. The Schur-Zassenhaus theorem asserts that if $N \lhd G$ with $|N|$ and $[G : N]$ relatively prime, then $G = N \rtimes Q$ with $Q \cong G/N$.

2 Group Representations

The general idea of representation theory is that one can study a group by letting it act linearly on a vector space. In this course, we will work exclusively with finite-dimensional vector spaces over $\mathbb{C}$.

Formally, a *representation* of a finite group G is a pair (ρ, V) where V is a vector space and ρ is a homomorphism from G to $GL(V)$, the group of automorphisms of V.

Thus for every $s, t \in G$, we have $\rho(st) = \rho(s)\rho(t)$. Writing I for the identity map on V, this implies that $\rho(id) = I$ and $\rho(s^{-1}) = \rho(s)^{-1}$.

(Here and henceforth we write id for the group identity to avoid confusion with the natural exponent.)

Since the codomain is part of the definition of a function, we will often just speak of the representation ρ. We call V the *representation space* and say that $d_\rho = \dim(V)$ is the *degree* or *dimension* of ρ.
Also, we will occasionally find it convenient to employ the notation $\rho_s := \rho(s)$.

When $V \cong \mathbb{C}^n$ comes equipped with a basis $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$, we can represent $a \in GL(V)$ by the $n \times n$ matrix having j^{th} column $a(\mathbf{e}_j)$. In this view, a representation of G is a rule that associates an invertible matrix to each group element in a manner that respects the underlying structure.

We won't be doing anything too fancy in this class, so we can generally just think of the representation space as $\mathbb{C}^n$ with the standard basis and treat our representations as matrices.

Of course, the choice of basis is arbitrary, so let's say that representations (ρ, V) and (ρ', V') of G are *equivalent* if there is a linear bijection $\tau : V \to V'$ which satisfies

$$\tau \circ \rho_s = \rho'_s \circ \tau \quad \text{for all } s \in G.$$

Example 2.1. We always have the *trivial representation* $\rho_0(s) = 1$ for all $s \in G$

When $G = S_n$, another one-dimensional representation is the *sign representation* $\rho_\pm(\sigma) = \operatorname{sgn}(\sigma)$.

Example 2.2. Suppose that $W = \operatorname{span}(\mathbf{w})$ is a one-dimensional subspace of $\mathbb{C}^d$ and let $\eta_0(s) = I_d$, the $d \times d$ identity matrix, for all $s \in G$. The map $\tau : W \to \mathbb{C}$ defined by $\tau(c\mathbf{w}) = c$ is a linear bijection satisfying

$$\rho_0(s)\tau(c\mathbf{w}) = \rho_0(s)c = c = \tau(c\mathbf{w}) = \tau\big(\eta_0(s)c\mathbf{w}\big),$$

so (η_0, W) is equivalent to $(\rho_0, \mathbb{C})$. Similarly, $\eta_\pm(\sigma) = \operatorname{sgn}(\sigma) I_d$ is equivalent to $\rho_\pm$.

In general, $\rho(s)\mathbf{v} = \mathbf{v}$ and $\rho(\sigma)\mathbf{v} = \operatorname{sgn}(\sigma)\mathbf{v}$ are valid representations for any vector space V. However, when $\dim(V) > 1$, these are direct sums of trivial/sign representations; see below.

Example 2.3. If $|G| = m$, the *left regular representation* is (λ, V) where V is an m-dimensional vector space with basis $\{\mathbf{e}_g\}_{g \in G}$ and λ satisfies $\lambda(g)\mathbf{e}_h = \mathbf{e}_{gh}$ for all $g, h \in G$.

The *right regular representation* on V is given by $\rho(g)\mathbf{e}_h = \mathbf{e}_{hg^{-1}}$, and the map defined by $\tau(\mathbf{e}_g) = \mathbf{e}_{g^{-1}}$ shows that λ and ρ are equivalent.

Example 2.4. If $G = S_n$, the *permutation representation* is defined by taking V to be a vector space with basis $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ and letting $\rho : S_n \to V$ be given by $\rho(\sigma)\mathbf{e}_k = \mathbf{e}_{\sigma(k)}$.
This associates to each $\sigma \in S_n$ the permutation matrix R_σ having (i,j)-entry $1\{\sigma(j) = i\}$. For any $\mathbf{x} \in \mathbb{C}^n$, $R_\sigma \mathbf{x}$ has k^{th} coordinate $x_{\sigma^{-1}(k)}$.

More generally if $\varphi : G \times X \to X$ defines an action of a group G on a finite set X and V is a vector space with basis $\{\mathbf{e}_x\}_{x \in X}$, the associated permutation representation (ρ, V) is defined by $\rho(g)\mathbf{e}_x = \mathbf{e}_{\varphi(g,x)}$.

The (left) regular representation is the special case $X = G$, $\varphi(g,h) = gh$.

If (ρ, V) is a representation and W is a subspace of V which is *stable* under ρ (so $\rho(s)\mathbf{w} \in W$ for every $s \in G$, $\mathbf{w} \in W$), the restriction of ρ to W gives a *subrepresentation*. We always have the subrepresentations corresponding to $W = V$ and $W = \{\mathbf{0}\}$. If ρ admits no other subrepresentations, we say that it is *irreducible*.

Now recall that V is said to be the *direct sum* of $W_1, W_2 \leq V$ (written $V = W_1 \oplus W_2$) if every $\mathbf{v} \in V$ can be uniquely expressed as $\mathbf{v} = \mathbf{w}_1 + \mathbf{w}_2$ with $\mathbf{w}_1 \in W_1$ and $\mathbf{w}_2 \in W_2$.
This is equivalent to requiring that $W_1 \cap W_2 = \{\mathbf{0}\}$ and $\dim(V) = \dim(W_1) + \dim(W_2)$.

(We can also form the *external direct sum* of vector spaces U and V as the vector space consisting of ordered pairs in $U \times V$ with all operations performed componentwise.)

The *direct sum of representations* (ρ^1, W_1) and (ρ^2, W_2) is the representation $(\rho^1 \oplus \rho^2, W_1 \oplus W_2)$ defined by $(\rho^1 \oplus \rho^2)_s(\mathbf{w}_1 + \mathbf{w}_2) = \rho^1_s(\mathbf{w}_1) + \rho^2_s(\mathbf{w}_2)$.

(For external direct sums, the analogous definition is $(\rho^1 \oplus \rho^2)_s\big(\mathbf{w}_1, \mathbf{w}_2\big) = \big(\rho^1_s(\mathbf{w}_1), \rho^2_s(\mathbf{w}_2)\big)$.)

By construction, $\rho^1 \oplus \rho^2$ has degree $d_{\rho^1 \oplus \rho^2} = d_{\rho^1} + d_{\rho^2}$.

The direct sum of more than two representations is defined by $\rho^1 \oplus \cdots \oplus \rho^{k+1} = (\rho^1 \oplus \cdots \oplus \rho^k) \oplus \rho^{k+1}$.

If we think of $\rho^1(s), \ldots, \rho^k(s)$ as matrices, then we can express the direct sum as the block diagonal matrix

$$\rho^1 \oplus \cdots \oplus \rho^k(s) = \begin{bmatrix} \rho^1(s) & & O \\ & \ddots & \\ O & & \rho^k(s) \end{bmatrix}.$$

Here we are assuming that a basis of $\bigoplus_{i=1}^k W_i$ is given by $\{\mathbf{e}^1_1, \ldots, \mathbf{e}^1_{d_1}, \ldots, \mathbf{e}^k_1, \ldots, \mathbf{e}^k_{d_k}\}$ with $\{\mathbf{e}^i_1, \ldots, \mathbf{e}^i_{d_i}\}$ the corresponding basis for W_i.

Example 2.5. Let $G = S_3$ and $W = \big\{\mathbf{x} \in \mathbb{C}^3 : x_1 + x_2 + x_3 = 0\big\}$. A basis for W is given by $\mathbf{w}_1 = \mathbf{e}_1 - \mathbf{e}_2$ and $\mathbf{w}_2 = \mathbf{e}_2 - \mathbf{e}_3$ where $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ are the standard basis vectors in $\mathbb{C}^3$. W is stable under the permutation representation $(\rho, \mathbb{C}^3)$ since permuting the coordinates of a vector does not change their sum.

I claim that the *standard representation* (ρ, W) is irreducible. Indeed every nontrivial subspace of W is of the form $W' = \text{span}(\mathbf{w})$ for some nonzero $\mathbf{w} = (x, y, z)$ in W.

Without loss of generality, assume that $x \neq 0$ so that $(1, y', z') \in W'$. If W' were stable under ρ, then we would also have $(y', 1, z')$ and thus $(1 - y', y' - 1, 0)$ in W'.

If $y' \neq 1$, this implies that $\mathbf{e}_1 - \mathbf{e}_2$ and thus $\mathbf{e}_2 - \mathbf{e}_3$ are in W', hence $W' = W$.
If $y' = 1$, we would have $(1, 1, -2) \in W'$ (as the coordinates must sum to 0), so $(1, -2, 1)$ and thus $(0, 3, -3)$ are in W', which implies that $\mathbf{e}_2 - \mathbf{e}_3$ and thus $\mathbf{e}_1 - \mathbf{e}_2$ are in W'.

We can express $\rho(\pi)$ in matrix form by computing

π	$\rho(\pi)\mathbf{w}_1$	$\rho(\pi)\mathbf{w}_2$	$\rho(\pi)$
id	$(1,-1,0) = \mathbf{w_1}$	$(0,1,-1) = \mathbf{w_2}$	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$
$(1\,2)$	$(-1,1,0) = -\mathbf{w}_1$	$(1,0,-1) = \mathbf{w}_1 + \mathbf{w}_2$	$\begin{bmatrix} -1 & 1 \\ 0 & 1 \end{bmatrix}$
$(1\,3)$	$(0,-1,1) = -\mathbf{w}_2$	$(-1,1,0) = -\mathbf{w}_1$	$\begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$
$(2\,3)$	$(1,0,-1) = \mathbf{w}_1 + \mathbf{w}_2$	$(0,-1,1) = -\mathbf{w}_2$	$\begin{bmatrix} 1 & 0 \\ 1 & -1 \end{bmatrix}$
$(1\,2\,3)$	$(0,1,-1) = \mathbf{w}_2$	$(-1,0,1) = -\mathbf{w}_1 - \mathbf{w_2}$	$\begin{bmatrix} 0 & -1 \\ 1 & -1 \end{bmatrix}$
$(3\,2\,1)$	$(-1,0,1) = -\mathbf{w}_1 - \mathbf{w_2}$	$(1,-1,0) = \mathbf{w_1}$	$\begin{bmatrix} -1 & 1 \\ -1 & 0 \end{bmatrix}$

Observe that the orthogonal complement of W in $\mathbb{C}^3$ is $W^\perp = \text{span}(\mathbf{1})$. This one-dimensional subspace carries the trivial representation and we can form the direct sum $\rho_0 \oplus \rho$.

Relative to the basis $\mathcal{B} = \{\mathbf{1}, \mathbf{w}_1, \mathbf{w}_2\}$, this has matrix form

$$\rho_0 \oplus \rho\big((1\,3)\big) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{bmatrix}, \; \rho_0 \oplus \rho\big((3\,2\,1)\big) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & -1 & 0 \end{bmatrix}, \ldots$$

To express these in the standard basis $\mathcal{E} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$, we must conjugate with the change of basis matrix $P_{\mathcal{E}\leftarrow\mathcal{B}}$ whose j^{th} column is the standard coordinates of the j^{th} vector in $\mathcal{B}$.

This gives the equivalent matrix representations

$$\begin{aligned}
(\rho_0 \oplus \rho)'\big((1\,3)\big) &= \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 1 \\ 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 1 \\ 1 & 0 & -1 \end{bmatrix}^{-1} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \\
(\rho_0 \oplus \rho)'\big((3\,2\,1)\big) &= \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 1 \\ 1 & 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & -1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 1 & -1 & 1 \\ 1 & 0 & -1 \end{bmatrix}^{-1} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}, \\
&\vdots
\end{aligned}$$

which we recognize as the permutation representation!

Our next order of business is to show that the irreducible representations are the building blocks of all others in the sense that every representation is a direct sum of irreducible representations.

To this end, we record the following proposition.

Proposition 2.1. *Let $\rho : G \to GL(V)$ be a representation and suppose that $W \leq V$ is stable under ρ. Then there exists a complement $W' \leq V$ such that $V = W \oplus W'$ and W' is stable under ρ.*

Proof. Let $\langle \cdot, \cdot \rangle$ be an inner product on V and define a new inner product $\langle \cdot, \cdot \rangle_\rho$ by

$$\langle \mathbf{x}, \mathbf{y} \rangle_\rho = \sum_{s \in G} \langle \rho(s)\mathbf{x}, \rho(s)\mathbf{y} \rangle .$$

This is indeed conjugate-symmetric, linear in the first argument, and positive-definite since $\langle \cdot, \cdot \rangle$ is an inner product and $\rho(s)$ is invertible. Moreover, it is invariant under ρ in the sense that

$$\begin{aligned} \langle \rho(t)\mathbf{x}, \rho(t)\mathbf{y} \rangle_\rho &= \sum_{s \in G} \langle \rho(s)\rho(t)\mathbf{x}, \rho(s)\rho(t)\mathbf{y} \rangle = \sum_{s \in G} \langle \rho(st)\mathbf{x}, \rho(st)\mathbf{y} \rangle \\ &= \sum_{u \in G} \langle \rho(u)\mathbf{x}, \rho(u)\mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{y} \rangle_\rho . \end{aligned}$$

Let W' be the orthogonal complement of W with respect to this inner product. Then $V = W \oplus W'$ and W' is stable under ρ because for any $\mathbf{x} \in W'$, $\mathbf{y} \in W$, $t \in G$, we have $\mathbf{z} = \rho(t^{-1})\mathbf{y} \in W$ and thus

$$\begin{aligned} \langle \rho(t)\mathbf{x}, \mathbf{y} \rangle_\rho &= \sum_{s \in G} \langle \rho(s)\rho(t)\mathbf{x}, \rho(s)\mathbf{y} \rangle = \sum_{s \in G} \left\langle \rho(st)\mathbf{x}, \rho(st)\rho(t^{-1})\mathbf{y} \right\rangle \\ &= \sum_{u \in G} \langle \rho(u)\mathbf{x}, \rho(u)\mathbf{z} \rangle = \langle \mathbf{x}, \mathbf{z} \rangle_\rho = 0. \end{aligned}$$

□

Remark 2.1. Note that the invariance of $\langle \cdot, \cdot \rangle_\rho$ means that if $\{\mathbf{f}_1, \ldots, \mathbf{f}_n\}$ is an orthonormal basis of V with respect to $\langle \cdot, \cdot \rangle_\rho$, then $\langle \rho(s)\mathbf{f}_i, \rho(s)\mathbf{f}_j \rangle_\rho = \delta_{ij}$ for all $s \in G$, $i, j \in [n]$.

Also, if $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is an orthonormal basis of V with respect to $\langle \cdot, \cdot \rangle$ and M is the linear transformation defined by $M\mathbf{e}_i = \mathbf{f}_i$, then $\langle M\mathbf{e}_i, M\mathbf{e}_j \rangle_\rho = \langle \mathbf{f}_i, \mathbf{f}_j \rangle_\rho = \delta_{ij} = \langle \mathbf{e}_i, \mathbf{e}_j \rangle$, hence $\langle M\mathbf{u}, M\mathbf{v} \rangle_\rho = \langle \mathbf{u}, \mathbf{v} \rangle$ by linearity.

It follows that the equivalent representation $\tau = M^{-1}\rho M$ satisfies

$$\langle \tau(s)\mathbf{e}_i, \tau(s)\mathbf{e}_j \rangle = \langle M\tau(s)\mathbf{e}_i, M\tau(s)\mathbf{e}_j \rangle_\rho = \langle \rho(s)M\mathbf{e}_i, \rho(s)M\mathbf{e}_j \rangle_\rho = \langle \rho(s)\mathbf{f}_i, \rho(s)\mathbf{f}_j \rangle_\rho = \delta_{ij}$$

and thus is unitary with respect to $\langle \cdot, \cdot \rangle$.

As such, we can always assume that our representations are unitary.

We are now able to prove the following extremely powerful result which enables us to study representations by breaking them up into their irreducible components.

Theorem 2.1 (Maschke's Theorem)**.** *Every representation is a direct sum of irreducible representations.*

Proof. If $d_\rho = 1$, then ρ is irreducible since V has no nontrivial subspaces. Now assume that the result holds for all representations of degree at most k and let $d_\rho = k + 1$. If (ρ, V) is irreducible, then we are done. Otherwise, there is a stable subspace $W < V$ and Proposition 2.1 gives $V = W \oplus W'$ with $\dim(W), \dim(W') \leq k$. The induction hypothesis shows that W and W' are direct sums of irreps and the result follows by the principle of induction. □

The direct sum construction gives us a means of constructing new representations of G from old ones. The other main way of doing this is by taking tensor products.

The *tensor product* of vector spaces U and V is the space $U \otimes V$ consisting of formal linear combinations of symbols of the form $\mathbf{u} \otimes \mathbf{v}$ (with $\mathbf{u} \in U$, $\mathbf{v} \in V$) subject to the relations

$$(\alpha \mathbf{u}_1 + \beta \mathbf{u}_2) \otimes \mathbf{v} = \alpha \mathbf{u}_1 \otimes \mathbf{v} + \beta \mathbf{u_2} \otimes \mathbf{v},$$
$$\mathbf{u} \otimes (\alpha \mathbf{w}_1 + \beta \mathbf{w}_2) = \alpha \mathbf{u} \otimes \mathbf{w}_1 + \beta \mathbf{u} \otimes \mathbf{w}_2.$$

If $\{\mathbf{e}_i\}_{i \in [m]}$ and $\{\mathbf{f}_j\}_{j \in [n]}$ are bases for U and V, then a basis for $U \otimes V$ is given by $\{\mathbf{e}_i \otimes \mathbf{f}_j\}_{i \in [m], j \in [n]}$.

The *tensor product of representations* (ρ, U) and (η, V) is the representation $(\rho \otimes \eta, U \otimes V)$ defined by $(\rho \otimes \eta)_s(\mathbf{u} \otimes \mathbf{v}) = \rho_s(\mathbf{u}) \otimes \eta_s(\mathbf{v})$, having degree $d_{\rho \otimes \eta} = d_\rho d_\eta$.

If $\rho(s)$ and $\eta(s)$ are in matrix form relative to $\{\mathbf{e}_i\}_{i \in [m]}$ and $\{\mathbf{f}_j\}_{j \in [n]}$, then relative to $\{\mathbf{e}_i \otimes \mathbf{f}_j\}_{i \in [m], j \in [n]}$, their tensor product has (block) matrix form

$$\rho \otimes \eta(s) = \begin{bmatrix} \rho(s)_{1,1}\eta(s) & \rho(s)_{1,2}\eta(s) & \cdots & \rho(s)_{1,d_\rho}\eta(s) \\ \rho(s)_{2,1}\eta(s) & \rho(s)_{2,2}\eta(s) & \cdots & \rho(s)_{2,d_\rho}\eta(s) \\ \vdots & & \ddots & \vdots \\ \rho(s)_{d_\rho,1}\eta(s) & \rho(s)_{d_\rho,2}\eta(s) & \cdots & \rho(s)_{d_\rho,d_\rho}\eta(s) \end{bmatrix}.$$

3 Characters

Recall that if V is a vector space with basis $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$, then $a \in GL(V)$ can be represented as the matrix $[a_{i,j}]$ having j^{th} column $a(\mathbf{e}_j)$. This enables us to define the *trace* of a as $\operatorname{Tr}(a) = \sum_{k=1}^{n} a_{k,k}$.

Some basic properties of the trace are established in the appendix. In particular, it is shown to be independent of the choice of basis, so one can speak unambiguously of the trace of a finite-dimensional representation.

We define the *character* of a representation $\rho : G \to GL(V)$ as the function $\chi_\rho : G \to \mathbb{C}$ given by

$$\chi_\rho(s) = \operatorname{Tr}\big(\rho(s)\big).$$

If ρ is an irreducible representation, we call χ_ρ an *irreducible character*. If $d_\rho = 1$, then $\chi_\rho = \rho$ is called a *linear character*.

Characters are extremely useful objects. One reason for this is that they retain a lot of information about the associated representation even though they are scalar- rather than matrix-valued.

Example 3.1. If ρ is the n-dimensional permutation representation of S_n from Example 2.4, then

$$\chi_\rho(\sigma) = \sum_{k=1}^{n} \rho(\sigma)_{k,k} = \sum_{k=1}^{n} 1\{\sigma(k) = k\}$$

gives the number of fixed points of σ.

Example 3.2. If ρ and λ are representations of G, then the block matrix constructions of the direct sum and tensor product show that $\chi_{\rho\oplus\lambda}(g) = \chi_\rho(g) + \chi_\lambda(g)$ and $\chi_{\rho\otimes\lambda}(g) = \chi_\rho(g)\chi_\lambda(g)$.

Proposition 3.1. *If χ is the character of a representation ρ having degree d, then*

(1) $\chi(id) = d$

(2) $\chi(s^{-1}) = \overline{\chi(s)}$

(3) $\chi(sts^{-1}) = \chi(t)$

Proof. The first assertion follows from the fact that $\rho(id) = I_d$.

For the second, if $o(s) = m$, then $\rho(s)^m = \rho(s^m) = \rho(id) = I_d$, hence the eigenvalues of $\rho(s)$ must be m^{th} roots of unity. (This is also a consequence of the fact that we can choose a basis in which our representations are unitary.) It follows that

$$\begin{aligned}\overline{\chi(s)} = \operatorname{Tr}\big(\overline{\rho(s)}\big) &= \sum_{k=1}^{d} \overline{\lambda_k} = \sum_{k=1}^{d} \lambda_k^{-1} \\ &= \operatorname{Tr}\big(\rho(s)^{-1}\big) = \operatorname{Tr}\big(\rho(s^{-1})\big) = \chi(s^{-1}).\end{aligned}$$

Finally since $\operatorname{Tr}(AB) = \operatorname{Tr}(BA)$,

$$\begin{aligned}\chi(sts^{-1}) = \operatorname{Tr}\big(\rho(sts^{-1})\big) &= \operatorname{Tr}\big(\rho(s)\rho(t)\rho(s)^{-1}\big) \\ &= \operatorname{Tr}\big(\rho(s)^{-1}\rho(s)\rho(t)\big) = \operatorname{Tr}\big(\rho(t)\big) = \chi(t).\end{aligned}$$

□

Since the trace is preserved under cyclic shifts of the argument—e.g. $\operatorname{Tr}(ABC) = \operatorname{Tr}(CAB)$—we likewise see that if $\rho' = \tau\rho\tau^{-1}$, then $\operatorname{Tr}(\rho') = \operatorname{Tr}(\rho)$. That is, equivalent representations have identical characters.

For our next results about characters, we need to establish the following surprisingly useful fact known as *Schur's Lemma*.

Lemma 3.1. *Let (ρ^1, V_1) and (ρ^2, V_2) be irreducible representations of G, and suppose that $f : V_1 \to V_2$ is a linear map that satisfies*

$$f \circ \rho_s^1 = \rho_s^2 \circ f \text{ for all } s \in G.$$

(1) If ρ^1 and ρ^2 are not equivalent, then $f \equiv \mathbf{0}$.

(2) If $V_1 = V_2$ and $\rho^1 = \rho^2$, then f is a scalar multiple of the identity.

Proof. Note that $\ker(f) = \{\mathbf{v} \in V_1 : f(\mathbf{v}) = \mathbf{0}\}$ is stable under ρ^1 since $\mathbf{v} \in \ker(f)$ implies

$$f\big(\rho_s^1(\mathbf{v})\big) = \rho_s^2\big(f(\mathbf{v})\big) = \rho_s^2(\mathbf{0}) = \mathbf{0}.$$

Similarly, $\mathrm{Im}(f) = \{\mathbf{w} \in V_2 : \mathbf{w} = f(\mathbf{v}) \text{ for some } \mathbf{v} \in V_1\}$ is stable under ρ^2 since $\mathbf{w} = f(\mathbf{v})$ implies that

$$\rho_s^2(\mathbf{w}) = \rho_s^2\big(f(\mathbf{v})\big) = f\big(\rho_s^1(\mathbf{v})\big).$$

Thus by irreducibility, $\ker(f)$ is either $\{\mathbf{0}\}$ or V_1 and $\mathrm{Im}(f)$ is either $\{\mathbf{0}\}$ or V_2.

It follows that if $f \not\equiv \mathbf{0}$, then $\ker(f) = \{\mathbf{0}\}$ and $\mathrm{Im}(f) = V_2$, so f is a bijection and the representations are equivalent.

Now suppose that $V_1 = V_2$ and $\rho^1 = \rho^2$. The claim certainly holds if $f \equiv \mathbf{0}$. Otherwise, f has a nonzero eigenvalue λ. In this case, the map $f_\lambda = f - \lambda I$ has a nontrivial kernel and satisfies $f_\lambda \circ \rho_s^1 = \rho_s^2 \circ f_\lambda$, hence $f_\lambda \equiv \mathbf{0}$. (Observe that it is important here that we are working over an algebraically closed field.) □

Corollary 3.1. *Let (ρ^1, V_1) and (ρ^2, V_2) be irreducible representations of G, write $d = \dim(V_1)$, and let $h : V_1 \to V_2$ be a linear map. Define*

$$\widetilde{h} = \frac{1}{|G|} \sum_{s \in G} (\rho_s^2)^{-1} \circ h \circ \rho_s^1. \tag{3.1}$$

(1) If ρ^1 and ρ^2 are not equivalent, then $\widetilde{h} \equiv \mathbf{0}$.

(2) If $V_1 = V_2$ and $\rho^1 = \rho^2$, then $\widetilde{h} = \lambda I$ with $\lambda = \mathrm{Tr}(h)/d$.

Proof. For any $t \in G$,

$$\begin{aligned}(\rho_t^2)^{-1} \circ \widetilde{h} \circ \rho_t^1 &= \frac{1}{|G|} \sum_{s \in G} (\rho_t^2)^{-1} (\rho_s^2)^{-1} \circ h \circ \rho_s^1 \rho_t^1 \\ &= \frac{1}{|G|} \sum_{s \in G} (\rho_{st}^2)^{-1} \circ h \circ \rho_{st}^1 = \widetilde{h},\end{aligned}$$

hence $\widetilde{h} \circ \rho_t^1 = \rho_t^2 \circ \widetilde{h}$.

If ρ^1 and ρ^2 are not equivalent, then Schur's lemma implies that $\widetilde{h} \equiv \mathbf{0}$.

If $V_1 = V_2$ and $\rho^1 = \rho^2$, Schur's lemma ensures that $\widetilde{h} = \lambda I$, and taking the trace of both sides in Equation (3.1) shows that $\lambda = \mathrm{Tr}(h)/d$. □

Let us now suppose that our representations are given in the matrix form $\rho^1_s = [r_{i,j}(s)]$, $\rho^2_s = [q_{i,j}(s)]$. Writing the linear maps from Corollary 3.1 as $h = [x_{i,j}]$, $\widetilde{h} = [\widetilde{x}_{i,j}]$, Equation (3.1) can be expressed entrywise as

$$\widetilde{x}_{i,\ell} = \frac{1}{|G|} \sum_{s,j,k} q_{i,j}(s^{-1}) x_{j,k} r_{k,\ell}(s). \tag{3.2}$$

In the first case, $\widetilde{x}$ is the zero matrix for every choice of x—such as those with a single entry equal to 1 and all others 0—so we must have

$$\frac{1}{|G|} \sum_{s} q_{i,j}(s^{-1}) r_{k,\ell}(s) = 0.$$

In the second case, $\widetilde{x}_{i,\ell} = \lambda \delta_{i\ell}$ with $\lambda = \frac{1}{d} \sum_{j,k} x_{j,k} \delta_{jk}$. Substituting this into Equation (3.2) yields

$$\frac{1}{d} \sum_{j,k} x_{j,k} \delta_{jk} \delta_{i\ell} = \frac{1}{|G|} \sum_{s,j,k} r_{i,j}(s^{-1}) x_{j,k} r_{k,\ell}(s).$$

As this holds for all choices of x, we can equate coefficients to obtain

$$\frac{1}{|G|} \sum_{s} r_{i,j}(s^{-1}) r_{k,\ell}(s) = \frac{1}{d} \delta_{jk} \delta_{i\ell}.$$

Recalling that we can choose bases so that our representations are unitary and thus satisfy $r_{i,j}(s^{-1}) = \overline{r_{j,i}(s)}$ (and employing the reindexing $s \mapsto s^{-1}$, $k \leftrightarrow \ell$ for the sake of aesthetics), we record the foregoing as

Corollary 3.2. *Let (ρ^1, V_1) and (ρ^2, V_2) be irreducible representations of G having (unitary) matrix form $\rho^1_s = [r_{i,j}(s)]$, $\rho^2_s = [q_{i,j}(s)]$, and write $d = \dim(V_1)$. Then for all valid indices i, j, k, ℓ*

(1) If ρ^1 and ρ^2 are not equivalent,

$$\frac{1}{|G|} \sum_{s} q_{i,j}(s) \overline{r_{k,\ell}(s)} = 0.$$

(2) If $V_1 = V_2$ and $\rho^1 = \rho^2$,

$$\frac{1}{|G|} \sum_{s} r_{i,j}(s) \overline{r_{k,\ell}(s)} = \begin{cases} \frac{1}{d}, & i = k \text{ and } j = \ell \\ 0, & \text{otherwise} \end{cases}.$$

That is, the matrix entries of the irreducible representations are orthogonal with respect to the inner product

$$(f \,|\, g) = \frac{1}{|G|} \sum_{s \in G} f(s) \overline{g(s)}, \quad f, g : G \to \mathbb{C}.$$

One immediate consequence of this observation is that there are only finitely many irreducible representations of a finite group G since $\dim\left(\mathbb{C}^G\right) = |G|$.

Another is the *first orthogonality relation* given below.

Theorem 3.1. *The irreducible characters are orthonormal with respect to $(\cdot \,|\, \cdot)$.*

Proof. Let ρ be an irreducible representation of degree d with $[r_{i,j}(t)] = \rho(t)$ a unitary matrix. The associated character is $\chi_\rho(t) = \sum_{k=1}^{d} r_{k,k}(t)$, and Corollary 3.2 implies

$$(\chi_\rho \,|\, \chi_\rho) = \frac{1}{|G|} \sum_{s \in G} \chi_\rho(s) \overline{\chi_\rho(s)} = \frac{1}{|G|} \sum_{s \in G} \sum_{k=1}^{d} r_{k,k}(s) \sum_{\ell=1}^{d} \overline{r_{\ell,\ell}(s)} = \sum_{k,\ell} (r_{k,k} \,|\, r_{\ell,\ell}) = \sum_{k,\ell} \delta_{k\ell}/d = 1.$$

Similarly, if η is an inequivalent irrep with $[q_{i,j}(t)] = \eta(t)$ a unitary matrix, then

$$(\chi_\rho \,|\, \chi_\eta) = \sum_{k,\ell} (r_{k,k} \,|\, q_{\ell,\ell}) = 0. \qquad \square$$

We can now say a bit more about the direct sum decomposition from Theorem 2.1.

Proposition 3.2. *Let (ρ, V) be a representation of G, and suppose that $V = W_1 \oplus \cdots \oplus W_k$ is a decomposition of V into irreducible components. If (η, W) is an irreducible representation of G, then the number of W_i which are equivalent to W is $(\chi_\rho \,|\, \chi_\eta)$.*

Proof. Since the character of a direct sum is the sum of the constituent characters (see Homework 4),

$$(\chi_\rho \,|\, \chi_\eta) = (\chi_1 \,|\, \chi_\eta) + \cdots + (\chi_k \,|\, \chi_\eta)$$

with χ_i the character of W_i. The result follows since $(\chi_i \,|\, \chi_\eta)$ is 1 if $W_i \cong W$ and 0 otherwise. $\square$

An upshot of this result is that the multiplicity of W in V does not depend on the chosen decomposition.

Corollary 3.3. *Representations with the same character are equivalent.*

Proof. Both contain the same irreps with the same multiplicity. $\square$

Corollary 3.4. *For any representation (ρ, V), $(\chi_\rho \,|\, \chi_\rho)$ is a positive integer which equals 1 iff ρ is irreducible.*

Proof. Let $V = n_1 V_1 \oplus \cdots \oplus n_m V_m$ be a direct sum decomposition of V into irreducible components. Here $V_1, \ldots, V_m$ is a complete list of the irreps and $n_i \in \mathbb{N}_0$ is the number of copies of V_i in V.
Writing χ_i for the character corresponding to V_i, we have

$$(\chi_\rho \,|\, \chi_\rho) = \left(\sum_{i=1}^m n_i \chi_i \,\middle|\, \sum_{j=1}^m n_j \chi_j \right) = \sum_{i,j} n_i n_j \, (\chi_i \,|\, \chi_j) = \sum_{i=1}^m n_i^2.$$

This is a positive integer that equals 1 if and only if some n_i is 1 and the rest are 0. $\square$

Example 3.3. If ρ is an irreducible representation of G and λ is a one-dimensional representation of G, then $\lambda \otimes \rho$ is irreducible because $\lambda(g)\overline{\lambda(g)} = 1$ for all g, hence

$$(\chi_{\lambda\otimes\rho} \,|\, \chi_{\lambda\otimes\rho}) = \frac{1}{|G|} \sum_{g\in G} \lambda(g)\chi_\rho(g)\overline{\lambda(g)\chi_\rho(g)} = \frac{1}{|G|} \sum_{g\in G} \chi_\rho(g)\overline{\chi_\rho(g)} = 1.$$

Example 3.4. Let ρ be the n-dimensional permutation representation of S_n. Arguing as in Example 2.5, the subspaces $W = \{\mathbf{x} \in \mathbb{C}^n : x_1 + \cdots + x_n = 0\}$ and $W^\perp = \operatorname{span}(\mathbf{1})$ are stable under ρ, so ρ is the direct sum of the standard representation and the trivial representation. The latter is irreducible since it is one-dimensional. If we are able to show that $(\chi_\rho \,|\, \chi_\rho) = 2$, then we can conclude that the standard representation is irreducible as well.
To this end, let $X = \sum_{i=1}^n 1\{\sigma(i) = i\}$ be the random variable that records the number of fixed points in a permutation σ drawn uniformly from S_n. We saw in Example 3.1 that $\chi_\rho(\sigma)$ gives the number of fixed points of σ. Since σ and σ^{-1} have the same number of fixed points,

$$(\chi_\rho \,|\, \chi_\rho) = \frac{1}{n!} \sum_{\sigma\in S_n} \chi_\rho(\sigma)\chi_\rho(\sigma^{-1}) = \mathbb{E}[X^2].$$

The desired result follows since $X^2 = \sum_{i=1}^n 1\{\sigma(i) = i\} + \sum_{i \neq j} 1\{\sigma(i) = i, \sigma(j) = j\}$ has expectation

$$\begin{aligned} E[X^2] &= \sum_{i=1}^n \mathbb{P}\{\sigma(i) = i\} + \sum_{i \neq j} \mathbb{P}\{\sigma(i) = i, \sigma(j) = j\} \\ &= n\frac{(n-1)!}{n!} + n(n-1)\frac{(n-2)!}{n!} = 2. \end{aligned}$$

It turns out to be quite instructive to play the same sort of game with the regular representation of an arbitrary finite group G.

Recall from Example 2.3 that this is the representation λ defined by $\lambda(g)\mathbf{e}_h = \mathbf{e}_{gh}$ for $\{\mathbf{e}_s\}_{s \in G}$ a basis of the representation space. It has degree $d_\lambda = |G|$ and matrix form $\ell_{g,h}(s) = 1\{sh = g\}$.

Since $sg = g$ iff $s = id$, the character of the regular representation is

$$\chi_\lambda(s) = \sum_{g \in G} \ell_{g,g}(s) = \begin{cases} |G|, & s = id \\ 0, & \text{otherwise} \end{cases}.$$

Proposition 3.3. *Every irreducible representation is contained in the regular representation with multiplicity equal to its degree.*

Proof. Let λ be the regular representation of G and let ρ be an irreducible representation. Then

$$(\chi_\lambda \,|\, \chi_\rho) = \frac{1}{|G|}\sum_{s \in G} \chi_\lambda(s)\chi_\rho(s^{-1}) = \frac{1}{|G|}\chi_\lambda(id)\chi_\rho(id) = \frac{1}{|G|}\,|G|\,d_\rho = d_\rho. \qquad \square$$

Corollary 3.5. *Let $\rho_1, \ldots, \rho_m$ be a complete list of the irreducible representations of G with ρ_k having character χ_k and degree d_k.*

(1) $\sum_{k=1}^m d_k^2 = |G|$

(2) For $s \neq id$, $\sum_{k=1}^m d_k \chi_k(s) = 0$

Proof. Proposition 3.3 shows that the regular representation has character $\chi(s) = \sum_{k=1}^m d_k\chi_k(s)$.

Taking $s = id$ gives (1), and taking $s \neq id$ gives (2). $\square$

The preceding results give upper bounds on the degrees and number of irreducible representations of G, as well as a criterion for checking that one has an exhaustive list of the irreps.

Also, since we know that the entries of the irreducible representations in unitary matrix form are orthogonal with the entries of ρ_k having norm d_k^{-1}, the fact that there are $\sum_{k=1}^m d_k^2 = |G|$ of them gives

Proposition 3.4. *Let $[r_{i,j}^k]$ be the matrix form of the irreducible representation ρ_k with respect to a basis that makes it unitary. Then an orthonormal basis for $\mathbb{C}^G$ is given by $\{\sqrt{d_k}r_{i,j}^k : k \in [m],\, i,j \in [d_k]\}$.*

Example 3.5. For the dihedral group D_n with $n = 2m$, there are 4 one-dimensional representations,

$$\begin{aligned} \psi_1(r^k) &= 1 = \psi_1(r^k s), \\ \psi_2(r^k) &= 1 = -\psi_2(r^k s), \\ \psi_3(r^k) &= (-1)^k = \psi_3(r^k s), \\ \psi_4(r^k) &= (-1)^k = -\psi_4(r^k s). \end{aligned}$$

The remaining are two-dimensional: Writing $\omega = e^{\frac{2\pi i}{n}}$, we have for each $1 \leq \ell \leq m-1$,

$$\rho_\ell(r^k) = \begin{bmatrix} \omega^{\ell k} & 0 \\ 0 & \omega^{-\ell k} \end{bmatrix}, \qquad \rho_\ell(r^k s) = \begin{bmatrix} 0 & \omega^{-\ell k} \\ \omega^{\ell k} & 0 \end{bmatrix}.$$

Indeed, the maps $g \mapsto \psi_j(g)$ and $g \mapsto \rho_\ell(g)$ are clearly homomorphisms, and the ρ_ℓ are irreducible because $\rho_\ell(r^k)$ and $\rho_\ell(r^k s)$ have no eigenvectors in common, so there is no stable one-dimensional subspace.
As $4 \cdot 1^2 + (m-1) \cdot 2^2 = 2n$, this accounts for all of them.

Observe that the characters of the two-dimensional irreps satisfy

$$\chi_\ell(r^k) = \chi_\ell(r^{-k}) = \omega^{\ell k} + \omega^{-\ell k} = 2\cos\left(\frac{\pi \ell k}{m}\right),$$
$$\chi_\ell(sr^k) = 0.$$

Given a finite group G, we define the *convolution* of $f, g : G \to \mathbb{C}$ by

$$(f * g)(s) = \sum_{t \in G} f(st^{-1})g(t).$$

(Like much in this section, the above definition generalizes from finite to locally compact groups by replacing summation with integration against Haar measure. For instance, when G is the real numbers under addition, we get the familiar convolution operation $(f * g)(x) = \int_{-\infty}^{\infty} f(x-y)g(y)\,dy$.)

Repeated convolution is expressed in the exponential notation $f^{*1} = f$ and $f^{*k} = f * f^{*(k-1)}$.

Example 4.1. Convolution plays nicely with delta functions in the sense that

$$(\delta_g * \delta_h)(s) = \sum_{t \in G} \delta_g(st^{-1})\delta_h(t) = \delta_g(sh^{-1}) = \delta_{gh}(s).$$

Note that if $gh \neq hg$, then $\delta_g * \delta_h \neq \delta_h * \delta_g$, so the convolution product is not commutative in general.

However, if G is abelian, then the change of variables $u = t^{-1}s$ gives

$$(f * g)(s) = \sum_{t \in G} f(st^{-1})g(t) = \sum_{t \in G} f(t^{-1}s)g(t) = \sum_{u \in G} f(u)g(su^{-1}) = (g * f)(s).$$

Example 4.2. For any group G, the space $\mathbb{C}^G$ of functions from G to $\mathbb{C}$ forms a ring under the operations of pointwise addition and convolution.

The abelian group structure induced by the sum $(f + g)(s) = f(s) + g(s)$ is clear—$f \equiv 0$ is the identity and $-g$ is the additive inverse of g—as is the distributivity of convolution over this pointwise addition, so it remains only to check associativity of convolution and the existence of a multiplicative identity.

For the former, observe that for any $f, g, h : G \to \mathbb{C}$, the change of variables $r = tu$ gives

$$\begin{aligned}\big((f * g) * h\big)(s) &= \sum_{u \in G} (f * g)(su^{-1})h(u) = \sum_{u \in G}\sum_{t \in G} f(su^{-1}t^{-1})g(t)h(u) \\ &= \sum_{u \in G}\sum_{r \in G} f(sr^{-1})g(ru^{-1})h(u) = \sum_{r \in G} f(sr^{-1})(g * h)(r) = \big(f * (g * h)\big)(s).\end{aligned}$$

For the latter, we compute

$$\begin{aligned}(f * \delta_{id})(s) &= \sum_{t \in G} f(st^{-1})\delta_{id}(t) = f(s), \\ (\delta_{id} * f)(s) &= \sum_{t \in G} \delta_{id}(st^{-1})f(t) = f(s).\end{aligned}$$

(Alternatively, $f = \sum_{g \in G} f(g)\delta_g$ and, by Example 4.1, $\delta_{id} * \delta_g = \delta_g * \delta_{id}$...)

Now define the *Fourier transform* of $f : G \to \mathbb{C}$ at the representation ρ as the $d_\rho \times d_\rho$ matrix

$$\widehat{f}(\rho) = \sum_{s \in G} f(s)\rho(s).$$

This is a generalization of the discrete Fourier transform ($G = \mathbb{Z}/n\mathbb{Z}$) and, suitably interpreted, the standard Fourier transform for functions on $\mathbb{R}$. As such, it possesses many familiar properties.

Proposition 4.1. *The convolution of functions $f, g : G \to \mathbb{C}$ has Fourier transform*

$$\widehat{f * g}(\rho) = \widehat{f}(\rho)\widehat{g}(\rho).$$

Proof. Multiplying by $\rho(t^{-1})\rho(t)$ and making the change of variables $u = st^{-1}$ gives

$$\begin{aligned}\widehat{f * g}(\rho) &= \sum_{s\in G}(f * g)(s)\rho(s) = \sum_{s\in G}\sum_{t\in G} f(st^{-1})g(t)\rho(s) \\ &= \sum_{s\in G}\sum_{t\in G} f(st^{-1})g(t)\rho(st^{-1})\rho(t) \\ &= \sum_{u\in G} f(u)\rho(u)\sum_{t\in G} g(t)\rho(t) = \widehat{f}(\rho)\widehat{g}(\rho).\end{aligned}$$

□

The fact that Fourier transforms take convolutions to products is very useful in applications like probability and signal processing. Our next result provides a sort of dictionary that enables one to more explicitly capitalize on such observations.

The first part shows that a function is completely determined by its Fourier transforms at irreps and gives a rule for recovering the function from its transforms.
The second can be thought of as relating 'inner products' in the time and frequency domains.

Theorem 4.1. *Let $\rho_1, \dots, \rho_m$ be the irreducible representations of G with $d_1, \dots, d_m$ the corresponding degrees. Then for any $f, g : G \to \mathbb{C}$, we have*

Fourier Inversion Formula:

$$f(s) = \frac{1}{|G|}\sum_{i=1}^{m} d_i \mathrm{Tr}\big(\rho_i(s^{-1})\widehat{f}(\rho_i)\big)$$

Plancherel Formula:

$$\sum_{s\in G} f(s)g(s^{-1}) = \frac{1}{|G|}\sum_{i=1}^{m} d_i \mathrm{Tr}\big(\widehat{f}(\rho_i)\widehat{g}(\rho_i)\big)$$

Proof. Since both sides of the above equations are linear in f, it suffices to prove the result for $f(s) = \delta_{st}$, in which case $\widehat{f}(\rho_i) = \rho_i(t)$.

For the inversion formula, the right-hand side is then

$$\frac{1}{|G|}\sum_{i=1}^{m} d_i \mathrm{Tr}\big(\rho_i(s^{-1})\rho_i(t)\big) = \frac{1}{|G|}\sum_{i=1}^{m} d_i \chi_i(s^{-1}t) = \delta_{st}$$

by Corollary 3.5.

For the Plancherel formula, we must show that

$$g(t^{-1}) = \frac{1}{|G|}\sum_{i=1}^{m} d_i \mathrm{Tr}\big(\rho_i(t)\widehat{g}(\rho_i)\big),$$

and this follows immediately from the inversion formula. □

Next, we call $f \in \mathbb{C}^G$ a *class function* if it's constant on conjugacy classes—that is, $f(sts^{-1}) = f(t)$ for all $s, t \in G$—and we write $\mathcal{C}(G)$ for the set of class functions on G.

Proposition 4.2. *$\mathcal{C}(G)$ is the center of $\mathbb{C}^G$.*

Proof. Since $\mathrm{cl}(id) = \{id\}$, $\delta_{id}(sts^{-1}) = \delta_{id}(t)$ for all $s, t \in G$, hence $\mathcal{C}(G)$ contains the multiplicative identity. Also, for any $f, g \in \mathcal{C}(G)$, $s, t \in G$,

$$\begin{aligned} f * g(sts^{-1}) &= \sum_u f(sts^{-1}u^{-1})g(u) = \sum_u f(sts^{-1}(sus^{-1})^{-1})g(sus^{-1}) \\ &= \sum_u f(stu^{-1}s^{-1})g(sus^{-1}) = \sum_u f(tu^{-1})g(u) = f * g(t). \end{aligned}$$

As $f, g \in \mathcal{C}(G)$ clearly implies $f - g \in \mathcal{C}(G)$, we conclude that $\mathcal{C}(G)$ is a subring of $\mathbb{C}^G$.

Now suppose that $f \in \mathcal{C}(G)$, $h \in \mathbb{C}^G$, $s \in G$. Then the change of variables $r = st^{-1}$ gives

$$h * f(s) = \sum_t h(st^{-1})f(t) = \sum_t h(st^{-1})f(sts^{-1}) = \sum_r h(r)f(sr^{-1}) = f * h(s).$$

Conversely, if $h \in \mathbb{C}^G$ satisfies $f * h = h * f$ for all $f \in \mathbb{C}^G$, then for any $s, t \in G$,

$$h(ts) = \sum_u h(tu^{-1})\delta_{s^{-1}}(u) = h * \delta_{s^{-1}}(t) = \delta_{s^{-1}} * h(t) = \sum_u \delta_{s^{-1}}(tu^{-1})h(u) = h(st),$$

hence $h(sts^{-1}) = h\big((st)s^{-1}\big) = h\big(s^{-1}(st)\big) = h(t)$. □

Proposition 4.3. *If f is a class function on G, then its Fourier transform at an irreducible representation ρ is given by $\widehat{f}(\rho) = \lambda I$ with*

$$\lambda = \frac{1}{d_\rho}\sum_{s\in G} f(s)\chi_\rho(s) = \frac{|G|}{d_\rho}\left(f \,|\, \overline{\chi_\rho}\right).$$

Proof. Observe that

$$\begin{aligned} \rho(s)\widehat{f}(\rho)\rho(s)^{-1} &= \sum_t f(t)\rho(s)\rho(t)\rho(s)^{-1} = \sum_t f(t)\rho(sts^{-1}) \\ &= \sum_u f(s^{-1}us)\rho(u) = \sum_u f(u)\rho(u) = \widehat{f}(\rho), \end{aligned}$$

so Schur's lemma shows that $\widehat{f}(\rho) = \lambda I$.

Taking traces of both sides gives $d_\rho\lambda = \mathrm{Tr}\Big(\sum_t f(t)\rho(t)\Big) = \sum_t f(t)\chi_\rho(t)$. □

Example 4.3. A probability μ on a group G defines a Markov chain $\{X_k\}_{k=0}^\infty$ that proceeds by sampling independently from μ and left-multiplying, so that the *transition function* is

$$P(g, h) = \mathbb{P}\big\{X_{k+1} = h \,\big|\, X_k = g\big\} = \mu(hg^{-1}).$$

The distribution after k steps of the chain started at the identity is given by the k-fold convolution μ^{*k}, which has Fourier transform $\widehat{\mu^{*k}}(\rho_i) = \widehat{\mu}(\rho_i)^k$ by Proposition 4.1.

If μ is constant on the conjugacy classes of G (which happens in many natural examples), then Proposition 4.3 tells us that $\widehat{\mu}(\rho_i) = \lambda_i I$ and thus $\widehat{\mu^{*k}}(\rho_i) = \lambda_i^k I$ with $\lambda_i = \sum_{s\in G}\mu(s)\frac{\chi_i(s)}{d_i}$, the expectation of the associated *character ratio* under μ.

Applying the inversion formula then yields

$$\mathbb{P}\big\{X_k = s \,\big|\, X_0 = id\big\} = \mu^{*k}(s) = \frac{1}{|G|}\sum_{i=1}^m d_i \mathrm{Tr}\Big(\rho_i(s^{-1})\widehat{\mu^{*k}}(\rho_i)\Big) = \frac{1}{|G|}\sum_{i=1}^m d_i\lambda_i^k\overline{\chi_i(s)}.$$

Theorem 4.2. *The irreducible characters form an orthonormal basis for the space of class functions.*

Proof. The first orthogonality relation tells us that the irreducible characters are orthonormal with respect to $(\cdot \,|\, \cdot)$, so it remains only to show that there are enough of them. In particular, the result will follow upon demonstrating that any class function which is orthogonal to the conjugates of each irreducible character must be identically 0.

Suppose that f is such a function. Then for any irrep ρ_i, Proposition 4.3 shows that $\widehat{f}(\rho_i) = \lambda_i I$ with $\lambda_i = |G|\, d_i^{-1}\, (f \,|\, \overline{\chi_i}) = 0$. Fourier inversion then implies that $f \equiv 0$. □

Corollary 4.1. *The number of irreducible representations is equal to the number of conjugacy classes.*

Proof. We know that the irreducible characters of G form a basis for $\mathcal{C}(G)$. Another basis is $\{1_{C_k}\}_{k=1}^{r}$ where $C_1, \ldots, C_r$ are the distinct conjugacy classes of G. □

Corollary 4.2. *The irreducible representations of a finite abelian group are all one-dimensional.*

Proof. The conjugacy classes of an abelian group G all have size one, so Corollary 4.1 implies that the number of irreps is $|G|$. Since the sum of their squared degrees is also $|G|$, they must all have degree one. □

Another consequence of Theorem 4.2 is the *second orthogonality relation*.

Theorem 4.3. *If $\chi_1, \ldots, \chi_m$ are the irreducible characters of G, then for any $s, t \in G$,*

$$\frac{1}{|G|} \sum_{i=1}^{m} \chi_i(s)\overline{\chi_i(t)} = \frac{1}{|\mathrm{cl}(s)|} 1\{t \in \mathrm{cl}(s)\}.$$

Proof. Set $f_s(t) = 1\{t \in \mathrm{cl}(s)\}$. Then f_s is a class function, so Theorem 4.2 implies $f_s(t) = \sum_{i=1}^{m} \alpha_i \chi_i(t)$ where

$$\alpha_i = (f_s \,|\, \chi_i) = \frac{1}{|G|} \sum_t f_s(t)\overline{\chi_i(t)} = \frac{|\mathrm{cl}(s)|}{|G|}\overline{\chi_i(s)}.$$

Taking conjugates of

$$1\{t \in \mathrm{cl}(s)\} = \sum_{i=1}^{m} \frac{|\mathrm{cl}(s)|}{|G|}\overline{\chi_i(s)}\chi_i(t)$$

and dividing by $|\mathrm{cl}(s)|$ yields the assertion. □

5 Abelian Groups and Dimension

We have shown that the irreducible representations of a finite abelian group all have degree one, but we can actually be a little more specific.

Example 5.1. The irreps of $G = \mathbb{Z}/n\mathbb{Z}$ are all of the form $\omega(k) = \omega^k$ as ω ranges over the n^{th} roots of 1.

Indeed, for any $m \in \{0, 1, \ldots, n-1\}$, define $m : G \to \mathbb{C}^*$ by $m(k) = e^{\frac{2\pi i m}{n}k}$. Then

$$m(j+k) = e^{\frac{2\pi i m}{n}(j+k)} = e^{\frac{2\pi i m}{n}j} e^{\frac{2\pi i m}{n}k} = m(j)m(k),$$

so $(m, \mathbb{C})$ is a representation. It's irreducible because $d_m = 1$.

Since we have produced one irrep for each of the n conjugacy classes of G, the list is exhaustive.

Observe that in this case, the Fourier transform and inversion formula are given by

$$\widehat{f}(m) = \sum_{k=0}^{n-1} f(k) e^{\frac{2\pi i m}{n}k}, \qquad f(k) = \frac{1}{n} \sum_{m=0}^{n-1} \widehat{f}(m) e^{-\frac{2\pi i m}{n}k}.$$

As finite abelian groups are products of cyclic groups, knowing how to compute representations of products will tell us (in principle) all about their representation theory.

Recall that if G_1 and G_2 are groups, their direct product is the group $G_1 \times G_2$ with multiplication $(s_1, t_1)(s_2, t_2) = (s_1 s_2, t_1 t_2)$.

Given representations (ρ^1, V_1) of G_1 and (ρ^2, V_2) of G_2, we can define the representation $(\rho^1 \otimes \rho^2, V_1 \otimes V_2)$ of $G_1 \times G_2$ by

$$(\rho^1 \otimes \rho^2)_{(s,t)}(\mathbf{v}_1 \otimes \mathbf{v}_2) = \rho^1_s(\mathbf{v}_1) \otimes \rho^2_t(\mathbf{v}_2).$$

The associated character is $\chi_{\rho^1 \otimes \rho^2}\big((s,t)\big) = \chi_{\rho^1}(s)\chi_{\rho^2}(t)$ and the degree is thus $d_{\rho^1 \otimes \rho^2} = d_{\rho^1} d_{\rho^2}$.

This follows by thinking about $(\rho^1 \otimes \rho^2)_{(s,t)}$ as a block diagonal matrix as we did when discussing the tensor product of two representations of a single group.

(When $G_1 = G_2 = G$, the restriction of the representation $\rho^1 \otimes \rho^2$ of $G \times G$ to the diagonal gives the representation $\rho^1 \otimes \rho^2$ of G.)

Proposition 5.1. *If G_1 and G_2 are finite groups, then every irreducible representation of $G_1 \times G_2$ is equivalent to some $\rho^1 \otimes \rho^2$ with $\rho^i \in Irr(G_i)$.*

Proof. Let ρ^1 and ρ^2 be irreducible representations of G_1 and G_2, respectively. Then

$$\begin{aligned}
\big(\chi_{\rho^1 \otimes \rho^2} \,|\, \chi_{\rho^1 \otimes \rho^2}\big) &= \frac{1}{|G_1 \times G_2|} \sum_{(s,t) \in G_1 \times G_2} \chi_{\rho^1 \otimes \rho^2}(s,t) \overline{\chi_{\rho^1 \otimes \rho^2}(s,t)} \\
&= \frac{1}{|G_1|\,|G_2|} \sum_{s \in G_1} \sum_{t \in G_2} \chi_{\rho^1}(s)\chi_{\rho^2}(t) \overline{\chi_{\rho^1}(s)\chi_{\rho^2}(t)} \\
&= \frac{1}{|G_1|} \sum_{s \in G_1} \chi_{\rho^1}(s) \overline{\chi_{\rho^1}(s)} \cdot \frac{1}{|G_2|} \sum_{t \in G_2} \chi_{\rho^2}(t) \overline{\chi_{\rho^2}(t)} \\
&= \big(\chi_{\rho^1} \,|\, \chi_{\rho^1}\big)\big(\chi_{\rho^2} \,|\, \chi_{\rho^2}\big) = 1 \cdot 1 = 1,
\end{aligned}$$

hence $\rho^1 \otimes \rho^2$ is irreducible by Corollary 3.4.

If $\eta^1 \in Irr(G_1)$ is not equivalent to ρ^1 or $\eta^2 \in Irr(G_2)$ is not equivalent to ρ^2, an analogous computation gives $\left(\chi_{\rho^1 \otimes \rho^2} \,|\, \chi_{\eta^1 \otimes \eta^2}\right) = \left(\chi_{\rho^1} \,|\, \chi_{\eta^1}\right)\left(\chi_{\rho^2} \,|\, \chi_{\eta^2}\right) = 0$, so this construction produces distinct irreps.

To see that all have been accounted for, observe that

$$\sum_{i=1}^{m_1}\sum_{j=1}^{m_2} d^2_{\rho_i \otimes \rho_j} = \sum_{i=1}^{m_1}\sum_{j=1}^{m_2} d_i^2 d_j^2 = \sum_{i=1}^{m_1} d_i^2 \sum_{j=1}^{m_2} d_j^2 = |G_1|\,|G_2| = |G_1 \times G_2|\,. \qquad \square$$

Taken together, Example 5.1 and Proposition 5.1 also imply Corollary 4.2, but it is still quite remarkable that partial knowledge of the number and dimensions of the irreducible representations can lead so easily to such a sweeping result. Thus inspired, we now set out to establish a few more facts of this nature.

We have defined the commutator of $g, h \in G$ as $[g,h] = ghg^{-1}h^{-1}$, which equals the identity if and only if g and h commute.

The *commutator subgroup* (or *derived subgroup*) $G' = [G,G]$ is defined to be the subgroup generated by the commutators, $G' = \langle [g,h] : g, h \in G \rangle$.

For any $K \leq G$, if $G' \subseteq K$, then for all $x \in G$, $k \in K$, $xkx^{-1} = xkx^{-1}k^{-1}k = [x,k]k \in K$, so $K \lhd G$.

Moreover, for any $x, y \in G$, $(xK)(yK) = xyK = yxx^{-1}y^{-1}xyK = yx[x^{-1}, y^{-1}]K = yxK = (yK)(xK)$, so G/K is abelian.

In particular, $G' \lhd G$ and G/G' is abelian.

In fact, if $N \lhd G$ is such that G/N is abelian, then for all $x, y \in G$, $xyN = (xN)(yN) = (yN)(xN) = yxN$, so $xyx^{-1}y^{-1} \in N$. As x and y were arbitrary, we conclude that $G' \subseteq N$. The commutator subgroup is thus the smallest normal subgroup whose quotient is abelian, so we can think of G' as a measure of abelianity: the larger the commutator, the less abelian the group.

Theorem 5.1. *The number of one-dimensional representations of a finite group G is $[G : G'] = |G| \,/\, |G'|$.*

Proof. Let $\pi : G \to G/G'$ be the natural map $\pi(x) = xG'$. We will show that $\psi \mapsto \psi \circ \pi$ defines a bijection from the irreducible representations of G/G' to the one-dimensional representations of G. (Since G/G' is abelian, its irreps are all one-dimensional, and of course, one-dimensional representations are always irreducible.)

On one hand, if $\psi : G/G' \to \mathbb{C}^*$ is a homomorphism, then $\psi \circ \pi : G \to \mathbb{C}^*$ is a composition of homomorphisms and thus is a homomorphism, so it defines a one-dimensional representation of G.

On the other, if $\rho : G \to \mathbb{C}^*$ is a homomorphism, then $\ker(\rho) \lhd G$ and $\mathrm{Im}(\rho) \cong G/\ker(\rho)$ is a subgroup of $\mathbb{C}^*$ and thus abelian. It follows that $G' \subseteq \ker(\rho)$. Define $\psi : G/G' \to \mathbb{C}^*$ by $\psi(xG') = \rho(x)$. This is coherent because if $xG' = yG'$, then $x^{-1}y \in G' \subseteq \ker(\rho)$, hence $\rho(y) = \rho(x)\rho(x^{-1}y) = \rho(x)$. Also, $\psi(xG'zG') = \psi(xzG') = \rho(xz) = \rho(x)\rho(z) = \rho(xG')\rho(zG')$, so ψ is a homomorphism. By construction $\rho = \psi \circ \pi$. $\square$

Our next order of business is to prove that the degrees of the irreps divide the order of the group.

We begin by recalling that $z \in \mathbb{C}$ is said to be an *algebraic integer* if it is the root of a monic polynomial with integer coefficients, and we write $\mathbb{A}$ for the set of algebraic integers. By the rational roots test, $\mathbb{A} \cap \mathbb{Q} = \mathbb{Z}$. This gives us a sneaky means of establishing divisibility: If $d, n \in \mathbb{Z}$ with $n/d \in \mathbb{A}$, then $d \,|\, n$.

Observe that characters are traces of unitary matrices and thus sums of roots of unity. Since an n^{th} root of one solves $x^n - 1 = 0$ and Proposition 7.2 ensures that $\mathbb{A}$ is closed under addition, we see that if χ is a character of a finite group G, then $\chi(g)$ is an algebraic integer for each $g \in G$. (Alternatively, if χ is the character of a degree d representation and $o(g) = m$, then $\chi(g)^m - d = \chi(g^m) - \chi(id) = 0$.)

Lemma 5.1. *Let ρ be an irreducible representation of a finite group G having degree d and character χ. Then for any $g \in G$, $\dfrac{|\mathrm{cl}(g)|\,\chi(g)}{d} \in \mathbb{A}$.*

Proof. Let $C_1, \ldots, C_r$ be the distinct conjugacy classes of G, write $h_i = |C_i|$, and let χ_i denote the value of χ on C_i. We wish to show that $h_i\chi_i/d$ is an algebraic integer for each i.

Setting $T_i = \sum_{x \in C_i} \rho(x)$, we see that for any $g \in G$, $\rho(g)T_i\rho(g)^{-1} = \sum_{x \in C_i} \rho(gxg^{-1}) = \sum_{y \in C_i} \rho(y) = T_i$.

Schur's lemma thus implies that $T_i = \lambda_i I_d$, and taking traces shows that $d\lambda_i = \sum_{x \in C_i} \mathrm{Tr}\big(\rho(x)\big) = h_i\chi_i$.

Also, $T_iT_j = \sum_{x \in C_i} \sum_{y \in C_j} \rho(xy) = \sum_{g \in G} |S_{ij}(g)|\,\rho(g)$ where $S_{ij}(g) := \big\{(x, y) \in C_i \times C_j : xy = g\big\}$.

Now if $g' = aga^{-1}$ for some $a \in G$, then the map $\varphi : S_{ij}(g) \to S_{ij}(g')$ defined by $\varphi\big((x, y)\big) = (axa^{-1}, aya^{-1})$ is clearly a bijection, so we can set $a_{ijk} := |S_{ij}(g)|$ where g is any element of C_k to get

$$T_iT_j = \sum_{g \in G} |S_{ij}(g)|\,\rho(g) = \sum_{k=1}^{r} \sum_{g \in C_k} a_{ijk}\rho(g) = \sum_{k=1}^{r} a_{ijk}T_k.$$

Substituting $T_k = (h_k\chi_k/d)I_d$ into this expression and examining the $(1,1)$-entry gives $(h_i\chi_i/d)(h_j\chi_j/d) = \sum_{k=1}^{r} a_{ijk}(h_k\chi_k/d)$, so Lemma 7.8 tells us that $h_i\chi_i/d \in \mathbb{A}$ as desired. □

With the preceding in hand, we can now prove our main result.

Theorem 5.2. *If ρ is a d-dimensional irreducible representation of G, then $d \,\big|\, |G|$.*

Proof. Let χ denote the character of ρ and keep the notation of Lemma 5.1.

The first orthogonality relation gives $1 = (\chi \,|\, \chi) = \frac{1}{|G|} \sum_{s \in G} \chi(s)\overline{\chi(s)}$, so

$$\frac{|G|}{d} = \sum_{s \in G} \frac{\chi(s)}{d}\overline{\chi(s)} = \sum_{i=1}^{r} \sum_{s \in C_i} \frac{\chi(s)}{d}\overline{\chi(s)} = \sum_{i=1}^{r} \frac{h_i\chi_i}{d} \cdot \overline{\chi_i}.$$

Since $h_i\chi_i/d$ and χ_i are algebraic integers and $\mathbb{A}$ is closed under conjugates, products, and sums, we have that $|G|/d$ is a rational algebraic integer and the claim follows. □

Remark 5.1. Theorem 5.2 can actually be strengthened to show that the dimension of any irrep divides the index of the center, $[G : Z(G)]$.

Example 5.2. We have seen that if p is prime and $|G| = p^2$, then G is abelian.

Another proof proceeds by noting that if $\rho \in Irr(G)$ has dimension d, then $d \,\big|\, p^2$, so $d \in \{1, p, p^2\}$. Since $p^2 = \sum_{\rho \in Irr(G)} d_\rho^2 = 1 + \sum_{\rho \in Irr(G) \setminus \{\rho_0\}} d_\rho^2$, we cannot have any representation of degree p or p^2, so all irreps are one-dimensional and G is abelian.

Similarly, if p and q are primes with $p < q$ and $q \not\equiv 1 (\mathrm{mod}\, p)$, then any group G of order pq must be abelian: If $d_1, \ldots, d_s$ are the degrees of the irreps, then we must have $pq = d_1^2 + \cdots + d_s^2$ and $d_k \,\big|\, pq$ for each k. As $p < q$, this means that $d_k \in \{1, p\}$ for all k.

Writing m and n for the number of degree 1 and p irreps, respectively, we have $pq = m + np^2$. Since $m = pq - np^2$ is divisible by p, $m \,|\, |G|$ by Theorem 5.1, $m \geq 1$ because of the trivial representation, we must have $m = p$ or $m = pq$. But $m = p$ gives $pq = p(1 + np)$ implies $q \equiv 1 (\mathrm{mod}\, p)$. It follows that the number of one-dimensional irreps is pq, hence G is abelian.

Yet another famous example of this general line of reasoning is Burnside's Theorem that if p and q are primes and $a, b \in \mathbb{N}_0$, then any group of order $p^a q^b$ is solvable, but this requires a bit more preparatory work than we have time for right now.

6 Restriction and Induction

If $\psi : G \to K$ is a homomorphism and $\rho : K \to GL(V)$ is a representation, then $\rho \circ \psi : G \to GL(V)$ is a composition of homomorphisms and thus a homomorphism. When ψ is surjective and ρ is irreducible, this representation of G is irreducible: If $W \leq V$ is stable under $\rho \circ \psi$, then for any $y \in K$, there is an $x \in G$ with $y = \psi(x)$ and thus $\rho(y)W = \rho\big(\psi(x)\big)W \subseteq W$, so W is stable under ρ as well.

An important example is $K = G/N$ for some $N \lhd G$ with $\psi : G \to G/N$ the natural map. In fact, this is how we got the one-dimensional irreps of G from those of its *abelianization* G/G'.

In addition to this *lifting* of representations, we can sometimes *descend* from the group to a suitable quotient. If $\rho : G \to GL(V)$ is a representation of a finite group G with kernel containing $N \lhd G$, then $\rho'(gN) = \rho(g)$ defines a representation of G/N. ($N \subseteq \ker(\rho)$ ensures that $gN = hN$ implies $\rho(g) = \rho(h)$, so the map $\rho' : G/N \to GL(V)$ is well-defined.) Its character satisfies

$$\begin{aligned}\langle \chi', \chi' \rangle_{G/N} &= \frac{1}{|G/N|} \sum_{xN \in G/N} \chi'(xN)\overline{\chi'(xN)} = \frac{|N|}{|G|} \sum_{xN \in G/N} \chi(x)\overline{\chi(x)} \\ &= \frac{1}{|G|} \sum_{xN \in G/N} \sum_{y \in xN} \chi(y)\overline{\chi(y)} = \frac{1}{|G|} \sum_{y \in G} \chi(y)\overline{\chi(y)} = \langle \chi, \chi \rangle_G ,\end{aligned}$$

hence ρ' is irreducible precisely when ρ is.

(In this section, we write $\langle \cdot, \cdot \rangle_G$ in place of $(\cdot \,|\, \cdot)$ to enhance readability and emphasize the underlying group.)

A kindred question is the how the representations of a group G relate to those of a subgroup $H \leq G$.

Certainly, if $\rho : G \to GL(V)$ is a representation of G, then the *restriction* $\rho_H : H \to GL(V)$, defined by $\rho_H(x) = \rho(x)$ for all $x \in H$, is a representation of H. However, irreducibility of ρ does not necessarily entail irreducibility of ρ_H. For instance, if H abelian and G is not, then G has a representation of degree greater than one whose restriction to H cannot be irreducible. (Of course, if ρ is not irreducible, then ρ_H is not either since a subspace $W \leq V$ that is stable under ρ will also be stable under ρ_H.)

Example 6.1. Let G be a finite group and $A \leq G$ an abelian subgroup. Then every irreducible representation of G has degree at most $|G| \,/\, |A|$.

Indeed, suppose that $\rho : G \to GL(V)$ is irreducible and let ρ_A be its restriction to A. Let $W \leq V$ be an irreducible subrepresentation so that $\dim(W) = 1$ by Corollary 4.2. Define $V' \leq V$ to be the vector subspace generated by $\rho(s)W$ as s ranges over G. By construction, V' is stable under ρ, so irreducibility implies $V' = V$. Also, for any $g \in G$, $a \in A$, we have $\rho(ga)W = \rho(g)\rho(a)W = \rho(g)W$. It follows that the number of distinct $\rho(g)W$, and thus the dimension of V, is at most $[G : A]$.

To better understand these issues, it helps to start by thinking about class functions:

Given a finite group G and a subgroup $H \leq G$, define the *restriction map* $\mathrm{Res}^G_H : \mathcal{C}(G) \to \mathcal{C}(H)$ by $\mathrm{Res}^G_H \varphi(h) = \varphi(h)$ for $\varphi \in \mathcal{C}(G)$, $h \in H$.

For $\psi \in \mathcal{C}(H)$, let $\widetilde{\psi} : G \to \mathbb{C}$ be its 'extension by zero,' $\widetilde{\psi}(g) = \psi(g)1\{g \in H\}$, and define the *induction map* $\mathrm{Ind}^G_H : \mathcal{C}(H) \to \mathcal{C}(G)$ by $\mathrm{Ind}^G_H \psi(g) = \dfrac{1}{|H|} \displaystyle\sum_{x \in G} \widetilde{\psi}(x^{-1}gx)$.

Proposition 7.3 in the appendix shows that these maps are well-defined and linear. Our next result, known as *Frobenius reciprocity*, shows they are adjoint.

Theorem 6.1. *Suppose $H \leq G$, $\psi \in \mathcal{C}(H)$, and $\varphi \in \mathcal{C}(G)$. Then*

$$\left\langle \psi, \mathrm{Res}_H^G \varphi \right\rangle_H = \left\langle \mathrm{Ind}_H^G \psi, \varphi \right\rangle_G .$$

Proof. We compute

$$\begin{aligned}
\left\langle \mathrm{Ind}_H^G \psi, \varphi \right\rangle_G &= \frac{1}{|G|} \sum_{g \in G} \mathrm{Ind}_H^G \psi(g) \overline{\varphi(g)} \\
&= \frac{1}{|G|} \sum_{g \in G} \frac{1}{|H|} \sum_{x \in G} \widetilde{\psi}(x^{-1} g x) \overline{\varphi(g)} \\
&= \frac{1}{|G|} \frac{1}{|H|} \sum_{x \in G} \sum_{g \in G} \widetilde{\psi}(x^{-1} g x) \overline{\varphi(g)} \\
&= \frac{1}{|G|} \frac{1}{|H|} \sum_{x \in G} \sum_{y \in G} \widetilde{\psi}(y) \overline{\varphi(xyx^{-1})} \\
&= \frac{1}{|G|} \frac{1}{|H|} \sum_{x \in G} \sum_{h \in H} \psi(h) \overline{\varphi(xhx^{-1})} \\
&= \frac{1}{|H|} \sum_{h \in H} \psi(h) \frac{1}{|G|} \sum_{x \in G} \overline{\varphi(h)} \\
&= \frac{1}{|H|} \sum_{h \in H} \psi(h) \overline{\varphi(h)} = \left\langle \psi, \mathrm{Res}_H^G \varphi \right\rangle_H .
\end{aligned}$$

□

Proposition 6.1. *If $x_1, \ldots, x_r$ is a transversal of H in G, then $\mathrm{Ind}_H^G \psi(g) = \sum_{i=1}^{r} \widetilde{\psi}(x_i^{-1} g x_i)$.*

Proof. If $h \in H$, then $h^{-1} g h \in H$ iff $g \in H$, so if $\psi \in \mathcal{C}(H)$ then $\widetilde{\psi}(h^{-1} g h) = \widetilde{\psi}(g)$. Accordingly, we have

$$\begin{aligned}
\mathrm{Ind}_H^G \psi(g) &= \frac{1}{|H|} \sum_{x \in G} \widetilde{\psi}(x^{-1} g x) = \frac{1}{|H|} \sum_{i=1}^{r} \sum_{h \in H} \widetilde{\psi}\big((x_i h)^{-1} g (x_i h)\big) \\
&= \sum_{i=1}^{r} \frac{1}{|H|} \sum_{h \in H} \widetilde{\psi}\big(x_i^{-1} g x_i\big) = \sum_{i=1}^{r} \widetilde{\psi}\big(x_i^{-1} g x_i\big).
\end{aligned}$$

□

Example 6.2. Suppose $K \leq H \leq G$ and let $\varphi \in \mathcal{C}(G)$. It is easy to see that $\mathrm{Res}_K^H \mathrm{Res}_H^G \varphi = \mathrm{Res}_K^G \varphi$. More interestingly, induction is also transitive. Indeed, if $\psi \in \mathcal{C}(K)$, then

$$\begin{aligned}
\mathrm{Ind}_H^G \mathrm{Ind}_K^H \psi(g) &= \frac{1}{|H|} \sum_{y \in G} \mathrm{Ind}_K^H \psi(y^{-1} g y) 1\{y^{-1} g y \in H\} \\
&= \frac{1}{|H|} \sum_{y \in G} \frac{1}{|K|} \sum_{x \in H} \psi(x^{-1} y^{-1} g y x) 1\{x^{-1} y^{-1} g y x \in K\} 1\{y^{-1} g y \in H\} \\
&= \frac{1}{|K|} \sum_{x \in H} \frac{1}{|H|} \sum_{y \in G} \psi(x^{-1} y^{-1} g y x) 1\{x^{-1} y^{-1} g y x \in K,\, y^{-1} g y \in H\} \\
&= \frac{1}{|K|} \sum_{x \in H} \frac{1}{|H|} \sum_{z \in G} \psi(z^{-1} g z) 1\{z^{-1} g z \in K,\, x z^{-1} g z x^{-1} \in H\} \\
&= \frac{1}{|K|} \sum_{x \in H} \frac{1}{|H|} \sum_{z \in G} \psi(z^{-1} g z) 1\{z^{-1} g z \in K\} \\
&= \frac{1}{|K|} \sum_{z \in G} \psi(z^{-1} g z) 1\{z^{-1} g z \in K\} = \mathrm{Ind}_K^G \psi(g).
\end{aligned}$$

Now if ρ is a representation of G with character χ_ρ, then the restriction of ρ to $H \leq G$ has character $\chi_{\rho_H}(h) = \mathrm{Tr}\big(\rho_H(h)\big) = \mathrm{Tr}\big(\rho(h)\big) = \chi_\rho(h)$.

In other words, if we denote the restriction of ρ to H by $\mathrm{Res}^G_H\rho$, then $\chi_{\mathrm{Res}^G_H\rho} = \mathrm{Res}^G_H\chi_\rho$.

Our next goal is to show that the induction map likewise sends characters to characters, so given a representation η of H, we can define $\mathrm{Ind}^G_H\eta$ to be the representation of G having character $\mathrm{Ind}^G_H\chi_\eta$.

Frobenius reciprocity then tells us that if ρ is an irreducible representation of G and η is an irreducible representation of H, the multiplicity of η in $\mathrm{Res}^G_H\rho$ is equal to the multiplicity of ρ in $\mathrm{Ind}^G_H\eta$.

Example 6.3. Let χ_0 be the trivial character on the trivial subgroup $\{id\} \leq G$. Since $x^{-1}gx = id$ iff $g = id$, we see that

$$\mathrm{Ind}^G_{\{id\}}\chi_0(g) = \sum_{x\in G}\widetilde{\chi_0}(x^{-1}gx) = \begin{cases} |G|\,, & g = id \\ 0, & g \neq id \end{cases}$$

is the character of the regular representation.

More generally, let $H \leq G$ and consider the coset action of G on G/H, $g(xH) = gxH$.
The set of points fixed by $g \in G$ is $\mathrm{Fix}(g) = \big\{xH : x^{-1}gx \in H\big\}$. Since each coset has $|H|$ elements, $|\mathrm{Fix}(g)| = \frac{1}{|H|}\big|\big\{x \in G : x^{-1}gx \in H\big\}\big|$.

Thus if χ_0 is the trivial character on H, then $\widetilde{\chi_0}(x^{-1}gx) = 1\big\{x^{-1}gx \in H\big\}$, so

$$\mathrm{Ind}^G_H\chi_0(g) = \frac{1}{|H|}\sum_{x\in G}\widetilde{\chi_0}(x^{-1}gx) = |\mathrm{Fix}(g)|\,,$$

the character of the permutation representation associated with the coset action.

(Recall that if G acts on X, then the permutation representation is (ρ, V) where V has basis $\{\mathbf{e}_x\}_{x\in X}$ and ρ is defined by $\rho(g)\mathbf{e}_x = \mathbf{e}_{gx}$. Its character is thus $\chi(g) = \sum_{x\in X}\langle\rho(g)\mathbf{e}_x, \mathbf{e}_x\rangle = \sum_{x\in X} 1\big\{gx = x\big\} = |\mathrm{Fix}(g)|$.)

The definition of the induction map as an average over conjugates is quite natural (especially in light of Theorem 6.1) and suggests something about the general form an induced representation should take. We will elaborate on the intuition soon, but first we provide a 'reasonable' construction and check that it works:

Given $\eta : H \to GL_d(\mathbb{C})$, let $\widetilde{\eta}(g) = \eta(g)1\{g \in H\}$, let $t_1, \ldots, t_r$ be a transversal of H in G, and define $\mathrm{Ind}^G_H\eta(g)$ to be the $rd \times rd$ block matrix with (i,j)-block $\widetilde{\eta}(t_i^{-1}gt_j)$ for $i, j \in [r]$.

Note that if $s_k \in t_kh_k$ with $h_k \in H$, then $\widetilde{\eta}(s_i^{-1}gs_j) = \widetilde{\eta}(h_i^{-1}t_i^{-1}gt_jh_j) = \eta(h_i)^{-1}\widetilde{\eta}(t_i^{-1}gxt_j)\eta(h_j)$, so changing the coset representatives just amounts to conjugating by a block-diagonal matrix of the form $\mathrm{diag}\big(\eta(h_1), \ldots, \eta(h_r)\big)$.

Theorem 6.2. *Suppose H is a subgroup of G of index r and $\eta : H \to GL_d(\mathbb{C})$ is a representation of H. Then $\mathrm{Ind}^G_H\eta : G \to GL_{rd}(\mathbb{C})$ is a representation of G with character $\chi_{\mathrm{Ind}^G_H\eta} = \mathrm{Ind}^G_H\chi_\eta$.*

Proof. Let $t_1, \ldots, t_r$ be a transversal of H in G, and for ease of notation, write $\eta^G_x = \mathrm{Ind}^G_H\eta(x)$ for the $rd\times rd$ block diagonal matrix with (i,j)-block $[\eta^G_x]_{ij} = \widetilde{\eta}(t_i^{-1}xt_j)$.
For any $x, y \in G$, $i, j \in [r]$, $\eta^G_x\eta^G_y$ has (i,j)-block

$$[\eta^G_x\eta^G_y]_{ij} = \sum_{k=1}^r [\eta^G_x]_{ik}[\eta^G_y]_{kj} = \sum_{k=1}^r \widetilde{\eta}(t_i^{-1}xt_k)\widetilde{\eta}(t_k^{-1}yt_j).$$

In order for $\widetilde{\eta}(t_k^{-1}yt_j)$ to not be the zero matrix, we must have $t_k^{-1}yt_j \in H$ or $yt_j \in t_kH$. Let t_ℓ be the unique representative of the coset containing yt_j so that $[\eta_x^G\eta_y^G]_{ij} = \widetilde{\eta}(t_i^{-1}xt_\ell)\eta(t_\ell^{-1}yt_j)$. This is nonzero precisely when $t_i^{-1}xt_\ell \in H$ or $t_iH = xt_\ell H = xyt_jH$, which in turn is equivalent to $t_i^{-1}xyt_j \in H$. In this case, $\widetilde{\eta}(t_i^{-1}xt_\ell)\eta(t_\ell^{-1}yt_j) = \eta(t_i^{-1}xt_\ell)\eta(t_\ell^{-1}yt_j) = \eta(t_i^{-1}xyt_j)$.

We have thus shown that $[\eta_x^G\eta_y^G]_{ij} = [\eta_{xy}^G]_{ij}$, so $\eta^G : G \to GL_{rd}(\mathbb{C})$ is indeed a homomorphism. Appealing to Proposition 6.1, its character is

$$\chi_{\eta^G}(g) = \mathrm{Tr}\big(\eta_g^G\big) = \sum_{i=1}^{r} \mathrm{Tr}\big(\widetilde{\eta}(t_i^{-1}gt_i)\big) = \sum_{i=1}^{r} \widetilde{\chi_\eta}(t_i^{-1}gt_i) = \mathrm{Ind}_H^G\chi_\eta(g). \qquad \square$$

The idea is that we start with a representation (η, W) of H, let V be a direct sum of $[G:H]$ copies of W, and let G act on V by (1) permuting the summands according to the coset action and (2) acting within each summand according to η.

Specifically, let $t_1, \ldots, t_r$ be a transversal of H in G and set $V = \bigoplus_{k=1}^r t_kW$ with $t_kW \cong W$ for $k = 1, \ldots, r$. For each $g \in G$, $k \in [r]$, there are unique $k(g) \in [r]$, $h_{g,k} \in H$ with $gt_k = t_{k(g)}h_{g,k}$. Given $\mathbf{v} = \sum_{k=1}^r t_k\mathbf{w}_k$, we define $\rho(g)\mathbf{v} = \sum_{k=1}^r t_{k(g)}\eta(h_{g,k})\mathbf{w}_k$. (Be aware that the t_k are not acting as scalars, they are keeping track of the 'coordinates.')

By taking natural bases for $W \cong \mathbb{C}^d$, $V \cong \mathbb{C}^{rd}$, we see that our matrix representation must satisfy $\rho(g)(\mathbf{w}_1, \ldots, \mathbf{w}_r)^\mathsf{T} = (\mathbf{v}_1, \ldots, \mathbf{v}_r)^\mathsf{T}$ where $\mathbf{v}_i = \eta(h)\mathbf{w}_j$ with $gt_j = t_ih$. That is, $\rho(g)$ is the block matrix having (i,j)-block $\widetilde{\eta}(t_i^{-1}gt_j)$.

Example 6.4. The *quaternion group* is defined as $Q_8 = \{\pm 1, \pm\hat{\imath}, \pm\hat{\jmath}, \pm\hat{k}\}$ with $\hat{\imath}^2 = \hat{\jmath}^2 = \hat{k}^2 = \hat{\imath}\hat{\jmath}\hat{k} = -1$. These relations imply -1 is central and $\hat{\imath}, \hat{\jmath}, \hat{k}$ multiply cyclically like cross-products of unit vectors in $\mathbb{R}^3$.

One can check by hand that Q_8 has commutator subgroup $Q_8' = \{\pm 1\}$, so there are $4 = [Q_8 : Q_8']$ one-dimensional irreps by Theorem 5.1. (Specifically, there's the trivial representation ρ_0, and the representations ρ_x, $x = \hat{\imath}, \hat{\jmath}, \hat{k}$, that map elements of $\langle x\rangle$ to 1 and the rest to -1.)

Since the sum of the squared degrees is 8, the remaining irrep is two-dimensional, and it turns out that we can compute it by induction: Set $H = \langle\hat{\imath}\rangle$ and consider the representation $\rho(\hat{\imath}^k) = i^k$. (The dotted i is the imaginary unit and the hatted $\hat{\imath}$ is the group element.) A transversal of H in Q_8 is $\{1, \hat{\jmath}\}$, so our formula gives $\mathrm{Ind}_H^G\rho(x) = \begin{bmatrix} \widetilde{\rho}(x) & \widetilde{\rho}(x\hat{\jmath}) \\ \widetilde{\rho}(-\hat{\jmath}x) & \widetilde{\rho}(-\hat{\jmath}x\hat{\jmath}) \end{bmatrix}$, which works out to

$$\mathrm{Ind}_H^G\rho(\pm 1) = \pm\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad \mathrm{Ind}_H^G\rho(\pm\hat{\imath}) = \pm\begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}$$
$$\mathrm{Ind}_H^G\rho(\pm\hat{\jmath}) = \pm\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \qquad \mathrm{Ind}_H^G\rho(\pm\hat{k}) = \pm\begin{bmatrix} 0 & -i \\ -i & 0 \end{bmatrix}.$$

Since $\mathrm{Ind}_H^G\rho(\hat{\imath})$ and $\mathrm{Ind}_H^G\rho(\hat{k})$ have no eigenvectors in common, this representation is indeed irreducible.

Example 6.4 shows that irreps sometimes induce to irreps, while Example 6.3 shows they sometimes do not. Given a representation η of $H \le G$, Corollary 3.4 and Theorem 6.1 imply $\mathrm{Ind}_H^G\eta \in Irr(G)$ if and only if

$$1 = \left\langle \chi_{\mathrm{Ind}_H^G\eta}, \chi_{\mathrm{Ind}_H^G\eta} \right\rangle_G = \left\langle \mathrm{Ind}_H^G\chi_\eta, \mathrm{Ind}_H^G\chi_\eta \right\rangle_G = \left\langle \chi_\eta, \mathrm{Res}_H^G\mathrm{Ind}_H^G\chi_\eta \right\rangle_H.$$

If $\eta = \eta_1 \oplus \eta_2$, then Frobenius reciprocity or linearity of the induction map gives $\langle \chi_{\mathrm{Ind}_H^G \eta}, \chi_{\mathrm{Ind}_H^G \eta} \rangle_G \geq \langle \chi_{\mathrm{Ind}_H^G \eta_1}, \chi_{\mathrm{Ind}_H^G \eta_1} \rangle_G + \langle \chi_{\mathrm{Ind}_H^G \eta_2}, \chi_{\mathrm{Ind}_H^G \eta_2} \rangle_G \geq 2$, so $\mathrm{Ind}_H^G \eta$ is reducible as well.

As such, we just need to understand $\mathrm{Res}_H^G \mathrm{Ind}_H^G \chi_\eta$ for $\eta \in Irr(H)$.

It turns out that it's about as easy to consider the slightly more general case of $\mathrm{Res}_H^G \mathrm{Ind}_K^G \chi_\eta$ for $H, K \leq G$ and η any representation of H. A key notion in this analysis is that of a double coset:

Given $H, K \leq G$, let $H \times K$ act on G by $(h,k)g = hgk^{-1}$. The orbit of g under this action is the *double coset* $HgK = \{hgk : h \in H, k \in K\}$, and we write $H\backslash G/K$ for the set of double cosets of H and K in G.

Theorem 6.3 (Mackey Decomposition)**.** *Let $H, K \leq G$ and let S be a complete set of double coset representatives for $H\backslash G/K$. Then for any $\psi \in \mathcal{C}(K)$,*

$$\mathrm{Res}_H^G \mathrm{Ind}_K^G \psi = \sum_{s \in S} \mathrm{Ind}_{H \cap K^s}^H \mathrm{Res}_{H \cap K^s}^{K^s} \psi^s$$

where $K^s = sKs^{-1}$ and $\psi^s \in \mathcal{C}(K^s)$ is given by $\psi^s(x) = \psi(s^{-1}xs)$.

Proof. We begin by determining a suitable transversal of K in G.

First, for each $s \in S$, choose a transversal V_s of $H \cap K^s$ in H, so that $H = \bigsqcup_{v \in V_s} v(H \cap K^s)$. One readily checks that $(H \cap K^s)sK = sK$, so we have

$$HsK = \bigcup_{v \in V_s} v(H \cap K^s)sK = \bigcup_{v \in V_s} vsK.$$

This union is disjoint since $vsK = v'sK$ for some $v, v' \in V_s$ implies $s^{-1}v^{-1}v's \in K$ and thus $v^{-1}v' \in K^s$. Because $v, v' \in H$ as well, we have $v^{-1}v' \in H \cap K^s$ or $v' \in v(H \cap K^s)$, hence $v = v'$ by definition of V_s.

Accordingly, writing $T_s = \{vs : v \in V_s\}$ for each $s \in S$ and setting $T = \bigcup_{s \in S} T_s$, we see that

$$G = \bigsqcup_{s \in S} HsK = \bigsqcup_{s \in S} \bigsqcup_{v \in V_s} vsK = \bigsqcup_{s \in S} \bigsqcup_{t \in T_s} tK = \bigsqcup_{t \in T} tK,$$

hence T is a transversal of K in G.

Two applications of Proposition 6.1 now show that for each $h \in H$,

$$\begin{aligned}
\mathrm{Ind}_K^G \psi(h) &= \sum_{t \in T} \widetilde{\psi}(t^{-1}ht) \\
&= \sum_{s \in S} \sum_{t \in T_s} \widetilde{\psi}(t^{-1}ht) \\
&= \sum_{s \in S} \sum_{v \in V_s} \widetilde{\psi}(s^{-1}v^{-1}hvs) \\
&= \sum_{s \in S} \sum_{v \in V_s} \psi^s(v^{-1}hv) 1\{v^{-1}hv \in K^s\} \\
&= \sum_{s \in S} \sum_{v \in V_s} \mathrm{Res}_{H \cap K_s}^{K_s} \psi^s(v^{-1}hv) 1\{v^{-1}hv \in H \cap K^s\} \\
&= \sum_{s \in S} \mathrm{Ind}_{H \cap K^s}^H \mathrm{Res}_{H \cap K^s}^{K^s} \psi^s(h).
\end{aligned}$$

□

We say that representations η and λ of G are *disjoint* if they have no equivalent (nonzero) subrepresentations. If we denote the irreps of G by $\rho_1, \ldots, \rho_r$, then we have the direct sum decompositions $\eta = m_1\rho_1 \oplus \cdots \oplus m_r\rho_r$ and $\lambda = n_1\rho_1 \oplus \cdots \oplus n_r\rho_r$ where $m_k, n_k \in \mathbb{N}_0$ and the *isotypic component* $m_k\rho_k$ is the direct sum of m_k copies of ρ_k. Disjointness means that there is no $k \in [r]$ with $m_k, n_k > 0$ and thus

$$\langle \chi_\eta, \chi_\lambda \rangle = \left\langle \sum_{i=1}^r m_i \chi_{\rho_i}, \sum_{j=1}^r n_j \chi_{\rho_j} \right\rangle = \sum_{i=1}^r \sum_{j=1}^r m_i n_j \left\langle \chi_{\rho_i}, \chi_{\rho_j} \right\rangle = \sum_{i=1}^r m_i n_i = 0.$$

That is, representations are disjoint precisely when their characters are orthogonal.

Theorem 6.4 (Mackey's Irreducibility Criterion)**.** *Let H be a subgroup of G and η a representation of H. For $s \in G$, set $H^s = sHs^{-1}$ and let η^s be the representation of H^s defined by $\eta^s(x) = \eta(s^{-1}xs)$.*

Then $\mathrm{Ind}_H^G \eta$ is an irreducible representation of G if and only if η is irreducible and the representations $\mathrm{Res}^H_{H\cap H^s}\eta$ and $\mathrm{Res}^{H^s}_{H\cap H^s}\eta^s$ are disjoint for every $s \notin H$.

Proof. Let S be a set of representatives for $H\backslash G/H$. Since any $s \notin H$ may be chosen as the representative of its double coset, it will suffice to show that $\mathrm{Res}^H_{H\cap H^s}\eta$ and $\mathrm{Res}^{H^s}_{H\cap H^s}\eta^s$ are disjoint for each $s \in S \setminus H$.

Write χ and χ^s for the characters of η and η^s, respectively. For the sole element h of $S \cap H$, we have $H^h \cap H = H$ and $\eta^h = \eta$, so Theorem 6.3 gives

$$\mathrm{Res}^G_H \mathrm{Ind}^G_H \chi = \sum_{s\in S} \mathrm{Ind}^H_{H\cap H^s} \mathrm{Res}^{H^s}_{H\cap H^s} \chi^s = \chi + \sum_{s\in S\setminus H} \mathrm{Ind}^H_{H\cap H^s} \mathrm{Res}^{H^s}_{H\cap H^s} \chi^s.$$

Applying Frobenius reciprocity twice, we find that

$$\begin{aligned}
\left\langle \mathrm{Ind}^G_H\chi, \mathrm{Ind}^G_H\chi \right\rangle_G &= \left\langle \chi, \mathrm{Res}^G_H \mathrm{Ind}^G_H \chi \right\rangle_H \\
&= \langle \chi, \chi \rangle_H + \sum_{s\in S\setminus H} \left\langle \chi, \mathrm{Ind}^H_{H\cap H^s} \mathrm{Res}^{H^s}_{H\cap H^s} \chi^s \right\rangle_H \\
&= \langle \chi, \chi \rangle_H + \sum_{s\in S\setminus H} \left\langle \mathrm{Res}^H_{H\cap H^s}\chi, \mathrm{Res}^{H^s}_{H\cap H^s} \chi^s \right\rangle_{H\cap H^s}.
\end{aligned}$$

Note that the final equality made use of the fact that inner products of characters take values in $\mathbb{N}_0 \subseteq \mathbb{R}$. Since $\langle \chi, \chi \rangle_H \geq 1$ with equality iff $\eta \in Irr(G)$ and $\left\langle \mathrm{Res}^H_{H\cap H^s}\chi, \mathrm{Res}^{H^s}_{H\cap H^s} \chi^s \right\rangle_{H\cap H^s} \geq 0$ with equality iff $\mathrm{Res}^H_{H\cap H^s}\chi$ and $\mathrm{Res}^{H^s}_{H\cap H^s}\chi^s$ are disjoint, the theorem has been proved. □

Corollary 6.1. *Let η be a representation of $H \lhd G$ and S a transversal of H in G. Then $\mathrm{Ind}_H^G \eta \in Irr(G)$ if and only if $\eta \in Irr(H)$ and for each $s \in S \setminus H$, η and η^s are inequivalent.*

Proof. As before, we only need to check disjointness for a set of double coset representatives, and $H \lhd G$ implies $HxH = xHH = xH$, hence $H\backslash G/H = G/H$. Moreover, $H^s = sHs^{-1} = H$ and

$$\langle \chi, \chi \rangle_H = \sum_{h\in H} \chi(h)\overline{\chi(h)} = \sum_{h\in H} \chi(s^{-1}hs)\overline{\chi(s^{-1}hs)} = \langle \chi^s, \chi^s \rangle_H,$$

hence η is irreducible if and only if η^s is. (The third equality is a reindexing of H, not a statement about characters being invariant under conjugation; $\chi \in \mathcal{C}(H)$ and $s \notin H$.) The assertion now follows from Theorem 6.4 since irreps are disjoint precisely when they're inequivalent. □

We conclude with a description of the *Mackey machine*, which allows one to construct the irreps of the semidirect product of a subgroup by an abelian normal subgroup from those of the constituent subgroups.

Specifically, suppose $A \lhd G$ is abelian and H is a complement of A in G.

The irreps of A are all one-dimensional and form a group, $\widehat{A}$, under pointwise multiplication. Also, G acts on $\widehat{A}$ by $g\chi = \chi^g$ with $\chi^g(x) = \chi(g^{-1}xg)$ for $x \in A$; this is well-defined since A is normal.

(Note also that we can uniquely write $g = ah$ so that $\chi^g(x) = \chi(h^{-1}a^{-1}xah) = \chi(h^{-1}xh) = \chi^h(x)$.)

Let $\{\chi_i\}_{i\in I}$ consist of one element from each orbit in $\widehat{A}/G$. For $i \in I$, let $H_i = \{h \in H : \chi_i^h = \chi_i\}$ be the stabilizer in H of χ_i, and let $G_i = AH_i$ be the semidirect product of A and H_i. Observe that χ_i extends to a one-dimensional representation of G_i by $\widetilde{\chi}_i(ah) = \chi_i(a)$ since

$$\begin{aligned}\widetilde{\chi}_i(a_1h_1a_2h_2) &= \widetilde{\chi}_i(a_1h_1a_2h_1^{-1}h_1h_2) = \chi_i(a_1h_1a_2h_1^{-1}) \\ &= \chi_i(a_1)\chi_i(h_1a_2h_1^{-1}) = \chi_i(a_1)\chi_i(a_2) = \widetilde{\chi}_i(a_1h_1)\widetilde{\chi}_i(a_2h_2).\end{aligned}$$

Similarly, let ρ be an irreducible representation of H_i. Since the projection $\pi : G_i \to H_i$ defined by $\pi(ah) = h$ is clearly a surjective homomorphism, ρ lifts to the irreducible representation $\widetilde{\rho} = \rho \circ \pi$ of G_i.

Example 3.3 shows that $\widetilde{\chi}_i \otimes \widetilde{\rho}$ is an irreducible representation of G_i, and we can form the representation $\theta_{i,\rho} = \mathrm{Ind}_{G_i}^G(\widetilde{\chi}_i \otimes \widetilde{\rho})$ of G.

The following theorem asserts that these are precisely the irreps of G.

Theorem 6.5.

(1) Each $\theta_{i,\rho}$ is an irreducible representation of G.
(2) If $\theta_{i,\rho}$ is equivalent to $\theta_{j,\eta}$ then $i = j$ and ρ is equivalent to η.
(3) Every irreducible representation of G is equivalent to some $\theta_{i,\rho}$.

Proof.

(1) We appeal to Theorem 6.4. Since $\widetilde{\chi}_i \otimes \widetilde{\rho} \in Irr(G_i)$, we just need to check that $\mathrm{Res}^{G_i}_{G_i\cap G_i^s}(\widetilde{\chi}_i \otimes \widetilde{\rho})$ and $\mathrm{Res}^{G_i^s}_{G_i\cap G_i^s}(\widetilde{\chi}_i \otimes \widetilde{\rho})^s$ are disjoint for $s \notin G_i$.

Now $G_i^s = sAH_is^{-1} = AH_i^s$, so we can uniquely write each $g \in G_i \cap G_i^s$ as $g = at$ with $a \in A$ and $t \in S_i$ for some transversal S_i of A in $G_i \cap G_i^s$; without loss, $A \cap S_i = \{id\}$, so $S_i \subseteq H_i^s$.

In this case, $\chi_{\widetilde{\chi}_i\otimes\widetilde{\rho}}(g) = \widetilde{\chi}_i(g)\chi_{\widetilde{\rho}}(g) = \chi_i(a)\chi_\rho(t)$ and $\chi_{(\widetilde{\chi}_i\otimes\widetilde{\rho})^s}(g) = \widetilde{\chi}_i(s^{-1}gs)\chi_{\widetilde{\rho}}(s^{-1}gs) = \chi_i^s(a)\chi_\rho^s(t)$, hence

$$\begin{aligned}\left\langle \mathrm{Res}^{G_i}_{G_i\cap G_i^s}\chi_{(\widetilde{\chi}_i\otimes\widetilde{\rho})}, \mathrm{Res}^{G_i^s}_{G_i\cap G_i^s}\chi_{(\widetilde{\chi}_i\otimes\widetilde{\rho})^s}\right\rangle_{G_i\cap G_i^s} &= \frac{1}{|G_i \cap G_i^s|}\sum_{g\in G_i\cap G_i^s}\chi_{\widetilde{\chi}_i\otimes\widetilde{\rho}}(g)\overline{\chi_{(\widetilde{\chi}_i\otimes\widetilde{\rho})^s}(g)} \\ &= \frac{1}{|G_i \cap G_i^s|}\sum_{t\in S_i}\sum_{a\in A}\widetilde{\chi}_i(a)\chi_{\widetilde{\rho}}(t)\overline{\chi_i^s(a)\chi_\rho^s(t)} \\ &= \frac{1}{|G_i \cap G_i^s|}\sum_{t\in S_i}\chi_{\widetilde{\rho}}(t)\overline{\chi_\rho^s(t)}\sum_{a\in A}\chi_i(a)\overline{\chi_i^s(a)}.\end{aligned}$$

Since $s \notin G_i$ ensures that χ_i and χ_i^s are inequivalent irreps of A, $\sum_{a\in A}\chi_i(a)\overline{\chi_i^s(a)} = 0$. This establishes disjointness and thereby the claim.

(2) Let T_i be a transversal of H_i in H, which we may suppose contains id. Since A is abelian and H a complement of A in G, T_i is also a transversal for $G_i = AH_i$ in G. Moreover, for any $x \in G$, $AxG_i = xAG_i = xG_i$, so T_i serves as a set of (A, G_i)-double coset representatives.

Writing $\chi_{i,\rho}$ for the character of $\theta_{i,\rho}$, Theorem 6.3 yields

$$\begin{aligned}\mathrm{Res}_A^G \chi_{i,\rho} = \mathrm{Res}_A^G \mathrm{Ind}_{G_i}^G(\widetilde{\chi}_i \otimes \chi_{\widetilde{\rho}}) &= \sum_{s\in T_i} \mathrm{Ind}_{A\cap G_i^s}^A \mathrm{Res}_{A\cap G_i^s}^{G_i^s}(\widetilde{\chi}_i \otimes \chi_{\widetilde{\rho}})^s \\ &= \sum_{s\in T_i} \mathrm{Ind}_{A\cap G_i^s}^A \chi_i^s \chi_{\widetilde{\rho}}^s = \sum_{s\in T_i} d_\rho \chi_i^s,\end{aligned}$$

where the final equality used $A \leq G_i^s$ and $\widetilde{\rho}^s(a) = \widetilde{\rho}(a) = \rho(id)$ for all $s \in G$, $a \in A$. As the restriction $\theta_{i,\rho}$ to A only involves characters in the orbit of χ_i, $\theta_{i,\rho}$ uniquely determines i.

Now let V be the representation space of $\theta_{i,\rho}$ and W the representation space of $\widetilde{\chi}_i \otimes \widetilde{\rho}$. Then we can write $V = \bigoplus_{s\in T_i} sW$ with $sW \cong W$ and $\theta_{i,\rho}(g)\sum_{s\in T_i} s\mathbf{w}_s = \sum_{s\in T_i} s_g\big(\widetilde{\chi}_i \otimes \widetilde{\rho}\big)(h_{g,s})\mathbf{w}_s$ where $s_g \in T_i$, $h_{g,s} \in G_i$ satisfy $gs = s_g h_{g,s}$. When $g \in G_i$ and $s = id$, we have $s_g = id$,$h_{g,s} = g$, thus if $\mathbf{v} \in V' = \big\{\sum_{s\in T_i} s\mathbf{w}_s : \mathbf{w}_s = \mathbf{0} \text{ for } s \neq id\big\}$, then $\theta_{i,\rho}(a)\mathbf{v} = \chi_i(a)\mathbf{v}$ for all $a \in A$ and $\theta_{i,\rho}(h)\mathbf{v} = \rho(h)\mathbf{v}$ for all $h \in H_i$. In particular, $(\mathrm{Res}_{H_i}^G \theta_{i,\rho}, V')$ is equivalent to ρ, hence $\theta_{i,\rho}$ determines ρ as well.

(3) $\theta_{i,\rho} = \mathrm{Ind}_{G_i}^G(\widetilde{\chi}_i \otimes \widetilde{\rho})$ has degree $\frac{|G|}{|G_i|} d_{\widetilde{\chi}_i \otimes \widetilde{\rho}} = \frac{|G|}{|G_i|} d_\rho$, so for each fixed $i \in I$,

$$\sum_{\rho\in Irr(G_i)} d_{\theta_{i,\rho}}^2 = \frac{|G|^2}{|G_i|^2} \sum_{\rho\in Irr(G_i)} d_\rho^2 = \frac{|G|^2}{|G_i|^2}|H_i| = \frac{|A|^2|H|^2}{|A|^2|H_i|^2}|H_i| = \frac{|H|^2}{|H_i|}.$$

Now the orbit-stabilizer theorem tells us that $\frac{|H|}{|H_i|} = |\mathcal{O}(\chi_i)|$, so

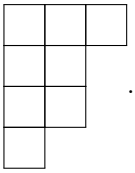
$$\sum_{i\in I}\sum_{\rho\in Irr(H_i)} d_{\theta_{i,\rho}}^2 = \sum_{i\in I} \frac{|H|^2}{|H_i|} = |H| \sum_{i\in I} |\mathcal{O}(\chi_i)| = |H|\left|\widehat{A}\right| = |H|\,|A| = |G|\,. \qquad \square$$

7 The Symmetric Group

In this section, we will explore the representation theory of S_n, which is particularly beautiful and replete with interesting combinatorics and connections to other areas of mathematics. Of course, we will only be able to scratch the surface here.

To begin, a *partition* λ of n (denoted $\lambda \vdash n$) is an ordered collection of positive integers $\lambda = (\lambda_1, \ldots, \lambda_k)$ where $\lambda_1 \geq \cdots \geq \lambda_k$ and $\lambda_1 + \cdots + \lambda_k = n$. The partitions of n index the conjugacy classes of S_n by specifying cycle type, and they can be represented by *Young diagrams*, which are left-justified arrays of boxes having λ_i in the i^{th} row.

For instance the partition $(3, 2, 2, 1) \vdash 8$ has Young diagram



.

If the boxes are populated by distinct elements of $[n]$, then the resulting object is called a *Young tableau* of *shape* λ, or λ-*tableau* for short. The following will serve as our running example:

$$t' = \begin{array}{|c|c|c|}\hline 3 & 5 & 1 \\ \hline 4 & 2 \\ \cline{1-2} 6 & 8 \\ \cline{1-2} 7 \\ \cline{1-1} \end{array} \quad .$$

Define a partial order on partitions by $(\lambda_1, \ldots, \lambda_k) \trianglerighteq (\mu_1, \ldots, \mu_\ell)$ if $\lambda_1 + \cdots + \lambda_i \geq \mu_1 + \cdots + \mu_i$ for all $i \geq 1$ (with $\lambda_i = 0$ when $i > k$, say). Thus, $(4, 2) \trianglerighteq (3, 3), (4, 1, 1)$, but $(3, 3)$ and $(4, 1, 1)$ are not comparable.

Lemma 7.1. *Suppose that $\lambda, \mu \vdash n$, let t be a λ-tableau, and let s be a μ-tableau. If for each i, the numbers in the i^{th} row of s belong to different columns of t, then $\lambda \trianglerighteq \mu$.*

Proof. We can construct a λ-tableau u that has the same entries as t in each column and contains the entries from the first i rows of s in its first i rows. This implies the claim since it guarantees that the number of entries in the first i rows of u is at least the number of entries in the first i rows of s.

To get u from t, start by moving each entry in the first row of s to the top of the column that contains it in t. Then move each entry from the second row of s to the topmost position in its column of t that has not yet been used. Continuing thusly we arrive at u, and each step is legitimate since the entries of any row of s belong to different columns of t. □

Now the symmetric group acts on λ-tableaux by permuting their entries. Thus if $\pi = (17)(245)(368)$, then

$$\pi t' = \begin{array}{|c|c|c|}\hline 6 & 2 & 7 \\ \hline 5 & 4 \\ \cline{1-2} 8 & 3 \\ \cline{1-2} 1 \\ \cline{1-1} \end{array} \quad .$$

Given a Young tableau t, define its *column stabilizer* C_t to be the group of all permutations π such that t and πt have the same elements in each column. For example, $C_{t'} = S_{\{3,4,6,7\}} S_{\{2,5,8\}} S_{\{1\}} \cong S_4 \times S_3$, where S_X denotes the subgroup of S_n that fixes all elements in $[n] \setminus X$.

Define an equivalence relation on the set of λ-tableaux by declaring that $t_1 \sim t_2$ if they have the same entries in each row, e.g.

$$\begin{array}{|c|c|c|}\hline 3&5&1\\\hline 4&2\\\cline{1-2} 6&8\\\cline{1-2} 7\\\cline{1-1}\end{array} \sim \begin{array}{|c|c|c|}\hline 1&3&5\\\hline 2&4\\\cline{1-2} 6&8\\\cline{1-2} 7\\\cline{1-1}\end{array} .$$

(The tableau on the right is called *standard* since its entries increase along each row and column.)

Equivalence classes of tableaux are called *tabloids*. The tabloid associated with tableau t is denoted $\{t\}$ and the set of λ-tabloids is denoted T^λ. If $\lambda = (\lambda_1, \dots \lambda_k)$, then $\left|T^\lambda\right| = n!/\lambda_1! \cdots \lambda_k!$.

Also, if $t_1 \sim t_2$, then $\sigma t_1 \sim \sigma t_2$ for all $\sigma \in S_n$ because if a is in row i of σt_1, then $\sigma^{-1}(a)$ is in row i of t_1, so $\sigma^{-1}(a)$ is in row i of t_2, so a is in row i of σt_2, and conversely. In particular, $\pi\{t\} = \{\pi t\}$ defines an action of S_n on T^λ.

For a fixed partition λ, define the *permutation module* M^λ to be the complex vector space with basis vectors $\left\{\mathbf{e}_{\{t\}} : \{t\} \in T^\lambda\right\}$, and consider the representation $(\rho_\lambda, M^\lambda)$ defined by $\rho_\lambda(\pi)\mathbf{e}_{\{t\}} = \mathbf{e}_{\{\pi t\}}$.
We will construct a unique irreducible subrepresentation of each M^λ. As the number of partitions equals the number of conjugacy classes, this will account for all of them.

Example 7.1. For the trivial partition (n), there is a single tabloid, so $\rho_{(n)}$ is the trivial representation.

At the other extreme, each tabloid $\{t\}$ of shape $(1^n) := (1, \dots, 1)$ contains only the tableau t, which we can represent as the permutation σ with $\sigma(i)$ the element in row i of t. $\rho_{(1^n)}$ is thus the regular representation $\rho_{(1^n)}(\pi)\mathbf{e}_\sigma = \mathbf{e}_{\pi\sigma}$.
The tabloids of shape $(n-1,1)$ are determined by the content of the single box in the second row, so $M^{(n-1,1)}$ has basis $\mathbf{e}_1, \dots, \mathbf{e}_n$ with $\rho_{(n-1,1)}(\pi)\mathbf{e}_k = \mathbf{e}_{\pi(k)}$. This is the n-dimensional permutation representation, which splits as the direct sum of the trivial representation and the standard representation.

Given partitions $\lambda, \mu \vdash n$ and a λ-tableau t, define the operator $A_t : M^\mu \to M^\mu$ by $A_t = \sum_{\pi \in C_t} \operatorname{sgn}(\pi)\rho_\mu(\pi)$. When $\mu = \lambda$, we call $\mathbf{f}_t := A_t\mathbf{e}_{\{t\}} = \sum_{\pi \in C_t} \operatorname{sgn}(\pi)\mathbf{e}_{\{\pi t\}}$ the *polytabloid* associated with t.

Proposition 7.1. *$\rho_\lambda(\sigma)\mathbf{f}_t = \mathbf{f}_{\sigma t}$ for all tableaux t of shape λ and all $\sigma \in S_n$.*

Proof. For $\sigma \in S_n$, $S \subseteq S_n$, $s \subseteq [n]$, write $\sigma S = \left\{\sigma\pi : \pi \in S\right\}$, $S\sigma = \left\{\pi\sigma : \pi \in S\right\}$, and $\sigma s = \left\{\sigma(i) : i \in s\right\}$. If c_j is the set of entries in the j^{th} column of t, then $c'_j = \sigma c_j$ is the set of entries in the j^{th} column of σt and we have

$$\begin{aligned}\pi \in C_{\sigma t} &\iff \pi c'_j = c'_j \ \forall j \iff (\pi\sigma)c_j = \sigma c_j \ \forall j\\ &\iff (\sigma^{-1}\pi\sigma)c_j = c_j \ \forall j \iff \sigma^{-1}\pi\sigma \in C_t\\ &\iff \pi \in \sigma C_t \sigma^{-1}.\end{aligned}$$

As $\operatorname{sgn}(\cdot)$ is invariant under conjugation and $C_{\sigma t} = \sigma C_t \sigma^{-1}$, a change of variables gives

$$\begin{aligned}\mathbf{f}_{\sigma t} &= \sum_{\pi \in C_{\sigma t}} \operatorname{sgn}(\pi)\mathbf{e}_{\{\pi\sigma t\}} = \sum_{\pi \in \sigma C_t \sigma^{-1}} \operatorname{sgn}(\sigma^{-1}\pi\sigma)\mathbf{e}_{\{\pi\sigma t\}}\\ &= \sum_{\eta \in C_t} \operatorname{sgn}(\eta)\mathbf{e}_{\{\sigma\eta t\}} = \sum_{\eta \in C_t} \operatorname{sgn}(\eta)\rho_\lambda(\sigma)\mathbf{e}_{\{\eta t\}} = \rho_\lambda(\sigma)\mathbf{f}_t.\end{aligned}$$

□

Thus if we define the *Specht module* S^λ to be the span of the polytabloids for tableaux of shape λ, then S^λ is a stable subspace of M^λ. (S_n acts transitively on the set of λ-tableaux, so for any λ-tableau t, we can write $S^\lambda = \operatorname{span}\{\mathbf{f}_{\sigma t} : \sigma \in S_n\}$.) Our goal is to show that the S^λ are irreducible and inequivalent.

Example 7.2. We saw in Example 7.1 that each tabloid t of shape (1^n) corresponds to some $\sigma \in S_n$. Clearly its column stabilizer is $C_t = S_n$, so the associated polytabloid is $\mathbf{f}_t = \sum_{\pi\in S_n} \operatorname{sgn}(\pi)\mathbf{e}_{\{\pi\sigma\}}$. Since $\operatorname{sgn} : S_n \to \{\pm 1\}$ is a homomorphism,

$$\begin{aligned}\rho_{(1^n)}(\alpha)\mathbf{f}_t = \mathbf{f}_{\alpha t} &= \sum_{\pi\in S_n} \operatorname{sgn}(\pi)\mathbf{e}_{\{\pi\alpha\sigma\}} \\ &= \operatorname{sgn}(\alpha^{-1}) \sum_{\pi\in S_n} \operatorname{sgn}(\pi\alpha)\mathbf{e}_{\{\pi\alpha\sigma\}} = \operatorname{sgn}(\alpha)\mathbf{f}_t,\end{aligned}$$

hence $S^{(1^n)}$ is the sign representation.

Lemma 7.2. *Let $\lambda, \mu \vdash n$. Suppose t is a λ-tableau and s is a μ-tableau with $A_t\mathbf{e}_{\{s\}} \neq \mathbf{0}$. Then $\lambda \trianglerighteq \mu$, and if $\lambda = \mu$, then $A_t\mathbf{e}_{\{s\}} = \pm\mathbf{f}_t$.*

Proof. Suppose there exist $a, b \in [n]$ such that a and b are in the same row of s and the same column of t. Then $\rho_\mu\big((ab)\big)\mathbf{e}_{\{s\}} = \mathbf{e}_{\{s\}}$ and $\langle (ab)\rangle \leq C_t$. Choosing a transversal $\sigma_1, \ldots, \sigma_r$ gives

$$\begin{aligned}A_t\mathbf{e}_{\{s\}} &= \sum_{\pi\in C_t} \operatorname{sgn}(\pi)\rho_\mu(\pi)\mathbf{e}_{\{s\}} = \sum_{k=1}^{r} \big[\operatorname{sgn}(\sigma_k)\rho_\mu(\sigma_k)\mathbf{e}_{\{s\}} + \operatorname{sgn}\big(\sigma_k(ab)\big)\rho_\mu\big(\sigma_k(ab)\big)\mathbf{e}_{\{s\}}\big] \\ &= \sum_{k=1}^{r} \operatorname{sgn}(\sigma_k)\rho_\mu(\sigma_k)\left[\mathbf{e}_{\{s\}} - \rho_\mu\big((ab)\big)\mathbf{e}_{\{s\}}\right] = \mathbf{0}.\end{aligned}$$

But this contradicts our assumption, so it must be the case that any elements belonging to the same row of s are in different columns of t, hence $\lambda \trianglerighteq \mu$ by Lemma 7.1.

If $\lambda = \mu$, then the fact that the elements in each row of s are in different columns of t implies that $s = \widetilde{\pi}t$ for some $\widetilde{\pi} \in C_t$. (Since the first rows of s and t have the same number of elements and those in the first row of s are in different columns of t, there is a permutation $\pi_1 \in C_t$ so that the sets of elements in the first rows of s and $\pi_1 t$ agree. Likewise, there is a permutation $\pi_2 \in C_t$ that fixes the first row of $\pi_1 t$ and ensures that the second rows of s and $\pi_2\pi_1 t$ contain the same elements...) It follows that

$$\begin{aligned}A_t\mathbf{e}_{\{s\}} &= \sum_{\pi\in C_t} \operatorname{sgn}(\pi)\mathbf{e}_{\{\pi s\}} = \sum_{\pi\in C_t} \operatorname{sgn}(\pi)\mathbf{e}_{\{\pi\widetilde{\pi}t\}} \\ &= \sum_{\sigma\in C_t} \operatorname{sgn}(\sigma)\operatorname{sgn}(\widetilde{\pi}^{-1})\mathbf{e}_{\{\sigma t\}} = \operatorname{sgn}(\widetilde{\pi})\mathbf{f}_t.\end{aligned}$$

□

Lemma 7.3. *Let $\mathbf{v} \in M^\lambda$ and let t be a λ-tableau. Then $A_t\mathbf{v} = c\mathbf{f}_t$ for some scalar $c \in \mathbb{C}$.*

Proof. Lemma 7.2 gives $A_t\mathbf{e}_{\{s\}} = c_{\{s\}}\mathbf{f}_t$ where $c_{\{s\}} \in \{-1, 0, 1\}$. Since $\mathbf{v} = \sum_{\{s\}\in T^\lambda} \alpha_{\{s\}}\mathbf{e}_{\{s\}}$ for some scalars $\{\alpha_{\{s\}}\} \subseteq \mathbb{C}$, we have

$$\begin{aligned}A_t\mathbf{v} &= \sum_{\pi\in C_t} \operatorname{sgn}(\pi)\rho_\lambda(\pi) \sum_{\{s\}\in T^\lambda} \alpha_{\{s\}}\mathbf{e}_{\{s\}} \\ &= \sum_{\{s\}\in T^\lambda} \alpha_{\{s\}} \sum_{\pi\in C_t} \operatorname{sgn}(\pi)\rho_\lambda(\pi)\mathbf{e}_{\{s\}} = \left(\sum_{\{s\}\in T^\lambda} \alpha_{\{s\}}c_{\{s\}}\right)\mathbf{f}_t.\end{aligned}$$

□

Now put an inner product $\langle \cdot, \cdot \rangle$ on M^λ that makes $\{\mathbf{e}_{\{s\}}\}_{\{s\}\in T^\lambda}$ an orthonormal set and for which ρ_λ is unitary. Then for any λ-tableau t and any $\mathbf{u}, \mathbf{v} \in M^\lambda$, we have

$$\begin{aligned}\langle A_t\mathbf{u}, \mathbf{v}\rangle &= \sum_{\pi\in C_t} \operatorname{sgn}(\pi)\,\langle \rho_\lambda(\pi)\mathbf{u}, \mathbf{v}\rangle = \sum_{\pi\in C_t} \operatorname{sgn}(\pi^{-1})\,\langle \mathbf{u}, \rho_\lambda(\pi^{-1})\mathbf{v}\rangle \\ &= \sum_{\pi\in C_t} \operatorname{sgn}(\pi)\,\langle \mathbf{u}, \rho_\lambda(\pi)\mathbf{v}\rangle = \langle \mathbf{u}, A_t\mathbf{v}\rangle\,,\end{aligned}$$

thus A_t self-adjoint with respect to $\langle \cdot, \cdot \rangle$.

Theorem 7.1 (Submodule Theorem)**.** *Let V be a stable subspace of M^λ. Then either $S^\lambda \subseteq V$ or $V \subseteq \left(S^\lambda\right)^\perp$.*

Proof. Suppose there is a λ-tableau t and a vector $\mathbf{v} \in V$ such that $A_t\mathbf{v} \neq \mathbf{0}$. Lemma 7.3 and the invariance of V imply that there is a nonzero $c \in \mathbb{C}$ with $c\mathbf{f}_t = A_t\mathbf{v} \in V$. It follows that $\mathbf{f}_{\sigma t} = \rho_\lambda(\sigma)\mathbf{f}_t \in V$ for all $\sigma \in S_n$ and thus $S^\lambda = \operatorname{span}\{\mathbf{f}_{\sigma t} : \sigma \in S_n\} \subseteq V$.

The only other possibility is that for every λ-tableau t and vector $\mathbf{v} \in V$, $A_t\mathbf{v} = \mathbf{0}$, so $\langle \mathbf{f}_t, \mathbf{v}\rangle = \langle A_t\mathbf{e}_{\{t\}}, \mathbf{v}\rangle = \langle \mathbf{e}_{\{t\}}, A_t\mathbf{v}\rangle = 0$ and thus $V \subseteq \left(S^\lambda\right)^\perp$. □

Corollary 7.1. *For each $\lambda \vdash n$, $(\rho_\lambda, S^\lambda)$ is irreducible.*

Proof. Suppose that V is a proper stable subspace of S^λ. Then Theorem 7.1 implies $V \subseteq \left(S^\lambda\right)^\perp$, and the assertion follows since $S^\lambda \cap \left(S^\lambda\right)^\perp = \{\mathbf{0}\}$. □

It remains only to understand how the Specht modules corresponding to different partitions relate to one another. Specifically, to account for all irreps, we must show that S^λ and S^μ are not equivalent when $\lambda \neq \mu$.

Lemma 7.4. *Let $\lambda, \mu \vdash n$ and suppose that T is a linear map from M^λ to M^μ that commutes with the action of S_n. If $S^\lambda \not\subseteq \ker(T)$, then $\lambda \trianglerighteq \mu$. Moreover, if $\lambda = \mu$, then $T|_{S^\lambda}$ is a scalar multiple of the identity.*

Proof. Suppose $S^\lambda \not\subseteq \ker(T)$. Since T is an intertwining map, $\ker(T)$ is a stable subspace of M^λ, so Theorem 7.1 implies $\ker(T) \subseteq \left(S^\lambda\right)^\perp$. Thus for any λ-tableau t, $\mathbf{0} \neq T\mathbf{f}_t = TA_t\mathbf{e}_{\{t\}} = A_tT\mathbf{e}_{\{t\}}$, where the final equality used $TA_t = \sum_{\pi\in C_t} \operatorname{sgn}(\pi)T\rho_\lambda(\pi) = \sum_{\pi\in C_t} \operatorname{sgn}(\pi)\rho_\mu(\pi)T = A_tT$.

Since $T\mathbf{e}_{\{t\}} = \sum_{\{s\}\in M^\mu} \alpha_{\{s\}}\mathbf{e}_{\{s\}}$ for some scalars $\{\alpha_{\{s\}}\}_{\{s\}\in M^\mu}$, it must be the case that $A_t\mathbf{e}_{\{s\}} \neq \mathbf{0}$ for some $\{s\} \in M^\mu$, hence $\lambda \trianglerighteq \mu$ by Lemma 7.2.

If $\lambda = \mu$, then Lemma 7.3 ensures the existence of some $c \in \mathbb{C}$ such that $T\mathbf{f}_t = A_tT\mathbf{e}_{\{t\}} = c\mathbf{f}_t \in S^\lambda$, hence T maps S^λ to itself. As S^λ is irreducible, Schur's lemma implies that $T|_{S^\lambda} = cI$. □

Lemma 7.5. *Let $T : S^\lambda \to S^\mu$ be a linear map that commutes with the action of S_n. If $T \neq 0$, then $\lambda \trianglerighteq \mu$.*

Proof. Any such T can be extended to a linear map T' from M^λ to M^μ by declaring $T'\mathbf{v} = \mathbf{0}$ for all $\mathbf{v} \in \left(S^\lambda\right)^\perp$, and the invariance of S^λ and $\left(S^\lambda\right)^\perp$ implies that for each $\mathbf{u} \in S^\lambda$, $\mathbf{v} \in \left(S^\lambda\right)^\perp$, $\sigma \in S_n$,

$$\begin{aligned}T'\rho_\lambda(\sigma)(\mathbf{u}+\mathbf{v}) &= T'\rho_\lambda(\sigma)\mathbf{u} + T'\rho_\lambda(\sigma)\mathbf{v} = T\rho_\lambda(\sigma)\mathbf{u} \\ &= \rho_\mu(\sigma)T\mathbf{u} = \rho_\mu(\sigma)T'(\mathbf{u}+\mathbf{v}).\end{aligned}$$

Thus if $T \neq 0$, then Lemma 7.4 tells us that $\lambda \trianglerighteq \mu$. □

Theorem 7.2. *Specht modules corresponding to distinct partitions are inequivalent.*

Proof. If ρ_λ and ρ_μ are equivalent, then there is a nonzero linear map $T : S^\lambda \to S^\mu$ with $T \circ \rho_\lambda = \rho_\mu \circ T$, so Lemma 7.5 implies $\lambda \trianglerighteq \mu$. A symmetric argument shows that $\mu \trianglerighteq \lambda$, proving the contrapositive of the assertion. □

In fact, Lemma 7.4 tells us a bit more about the direct sum decomposition of the permutation modules:

Corollary 7.2. (ρ_μ, S^μ) *has multiplicity one in* (ρ_μ, M^μ)*, and any other irreducible constituent* $(\rho_\lambda, S^\lambda)$ *must satisfy* $\lambda \trianglerighteq \mu$*.*

Finally, we observe that while the Specht module S^λ is spanned by the λ-polytabloids, they are not linearly independent in general; see Example 7.2. However, one can show that the set of polytabloids corresponding to standard tableaux of shape λ actually forms a basis for S^λ. The degree of ρ_λ is thus the number of standard λ-tableau, which can be computed using the famous hook length formula.

Clearly there's plenty more representation theory to study, even apart from considering infinite groups or fields other than $\mathbb{C}$!

Lemma 7.6 (Bézout's lemma). *The greatest common divisor of positive integers $b_1, \ldots, b_k$ is the smallest positive integer which can be expressed as an integral linear combination of $b_1, \ldots, b_k$.*

Proof. Let $d = \alpha_1 b_1 + \ldots + \alpha_k b_k$ be the smallest positive integer which can be so expressed. (Such a number exists by the well-ordering of $\mathbb{N}$.) If any b_j is divided by d, then its remainder $0 \leq r_j < d$ is of the form $r_j = \alpha_1' b_1 + \ldots + \alpha_k' b_k$ since it is obtained by subtracting a multiple of d from b_j. As d is the smallest positive integer of this form, it must be the case that $r_j = 0$, hence d is a divisor of each b_j. If c is any other common divisor of $b_1, \ldots, b_k$, then c divides d as well, hence d is the greatest common divisor. □

Theorem 7.3 (Second Isomorphism Theorem). *If $H \lhd G$ and $K \leq G$, then $HK \leq G$, $H \cap K \lhd K$, and*

$$K/(H \cap K) \cong HK/H.$$

Proof. We have already seen that $HK \leq G$, and since $H \subseteq HK$ with $H \lhd G$, we have that $H \lhd HK$.

Now suppose $xH \in HK/H$. Then $x = hk = kk^{-1}hk = kh'$ for some $h, h' \in H$, $k \in K$, so $xH = kH$. It follows that the map $\varphi : K \to HK/H$ defined by $\varphi(k) = kH$ is a surjective homomorphism. Its kernel is clearly $H \cap K$, so Theorem 1.1 shows that $H \cap K \lhd K$ and $K/(H \cap K) = K/\ker(\varphi) \cong \mathrm{Im}(\varphi) = HK/H$. □

Theorem 7.4 (Third Isomorphism Theorem). *If $H, K \lhd G$ and $K \leq H$, then $H/K \lhd G/K$ and*

$$(G/K)/(H/K) \cong G/H.$$

Proof. Let $\varphi : G/K \to G/H$ be given by $\varphi(gK) = gH$. φ is well-defined because if $aK = bK$, then $a^{-1}b \in K \leq H$, so $aH = bH$, and one readily checks that it is a surjective homomorphism. Since $gH = H$ iff $g \in H$, $\ker(\varphi) = \{gK \in G/K : gH = H\} = H/K$, so Theorem 1.1 gives $(G/K)/(H/K) \cong G/H$. □

Theorem 7.5 (Correspondence Theorem). *Suppose that $N \lhd G$ and define $\mathcal{S}(G;N) = \{H \leq G : N \leq H\}$, $\mathcal{S}(G/N) = \{K : K \leq G/N\}$. Then the map $\psi : \mathcal{S}(G;N) \to \mathcal{S}(G/N)$ defined by $\psi(S) = S/N$ is a bijection. Also, for any $S, T \in \mathcal{S}(G;N)$, $T \leq S$ if and only if $\psi(T) \leq \psi(S)$, in which case $[S:T] = [\psi(S) : \psi(T)]$. Moreover, $T \lhd S$ if and only if $\psi(T) \lhd \psi(S)$, in which case $S/T \cong \psi(S)/\psi(T)$.*

Proof. If $N \leq H \leq G$, then H/N is a subgroup of G/N since it is a subset which is itself a group. To see that ψ is injective observe that if $S, T \in \mathcal{S}(G;N)$ with $S/N = T/N$, then for any $s \in S$, there is a $t \in T$ with $sN = tN$ so that $s = tn$ for some $n \in N \leq T$, hence $s \in T$. The reverse inclusion is proved similarly. To see that ψ is surjective, suppose that $A' \leq G/N$ and define $A = \{x \in G : xN \in A'\}$. If $x, y \in A$, then $xN, yN \in A'$, hence $xy^{-1}N = xNy^{-1}N = (xN)(yN)^{-1} \in A'$. This shows that A is a subgroup of G, and it contains N since $nN = N \in A'$ for all $n \in N$. Surjectivity follows since $A' = \psi(A)$ by construction.

Since bijections preserve set inclusion, it's clear that $T \leq S$ iff $\psi(T) \leq \psi(S)$, and one can check that the map $sT \mapsto (sN)\psi(T)$ is bijection from cosets of T in S to those of T/N in S/N. (For finite groups, $[S:T] = [\psi(S) : \psi(T)]$ is an arithmetic consequence of $[G:H] = |G| \,/\, |H|$.)

If $T \lhd S$, then the third isomorphism theorem gives $T/N \lhd S/N$ and $(S/N)/(T/N) \cong S/T$. Finally, if $T/N \lhd S/N$, $t \in T$, and $s \in S$, then $(sts^{-1})N = (sN)(tN)(sN)^{-1} \in T/N$, hence $sts^{-1} = t'n'$ for some $t' \in T$, $n' \in nN \leq T$, showing that $sts^{-1} \in T$. □

Lemma 7.7. *If G is a p-group acting on a finite set X and $\mathrm{Fix}_G(X) = \{x \in X : gx = x \text{ for all } g \in G\}$ is the set of fixed points, then $|X| \equiv |\mathrm{Fix}_G(X)| \pmod p$.*

Proof. Write $X = \bigsqcup_{i=1}^{r} \mathcal{O}(x_i)$ with $\{x_1, \ldots, x_s\} = \mathrm{Fix}_G(X)$ so that $|X| = |\mathrm{Fix}_G(X)| + \sum_{i=s+1}^{r} |\mathcal{O}(x_i)|$. Since $|\mathcal{O}(x_i)| = |G| / |G_{x_i}|$ is divisible by p for $i > s$, the result follows. □

Theorem 7.6 (Sylow II)**.** *If P and Q are p-Sylow subgroups of G, then $Q = gPg^{-1}$ for some $g \in G$.*

Proof. Let Q act on the family of cosets G/P by left multiplication and write $\mathrm{Fix}_Q(G/P)$ for the set of fixed points. Lemma 7.7 gives $|G/P| \equiv |\mathrm{Fix}_Q(G/P)| \pmod p$. Since $|G/P| = |G| / |P|$ is not divisible by p, $|\mathrm{Fix}_Q(G/P)| \neq 0$, so there is some $gP \in \mathrm{Fix}_Q(G/P)$. This means that $q(gP) = gP$, hence $qg \in gP$, for all $q \in Q$, so $Q \subseteq gPg^{-1}$. As conjugation by g is an automorphism of G, $|gPg^{-1}| = |P| = |Q|$ and we conclude that $Q = gPg^{-1}$. □

Theorem 7.7 (Sylow III)**.** *Suppose that $|G| = p^k m$ with $p \nmid m$ and write n_p for the number of p-Sylow subgroups of G. Then $n_p \mid m$ and $n_p \equiv 1 \pmod p$. If P is any p-Sylow subgroup of G and $N_G(P)$ is its normalizer in G, then $n_p = [G : N_G(P)]$.*

Proof. Write $\mathrm{Syl}_p(G)$ for the set of p-Sylow subgroups of G, and let P act on $\mathrm{Syl}_p(G)$ by conjugation. ($h \mapsto ghg^{-1}$ is an automorphism of G and thus maps subgroups to subgroups of the same size.) We will show that $\mathrm{Fix}_P\big(\mathrm{Syl}_p(G)\big) = \{P\}$, so that $n_p \equiv 1 \pmod p$ by Lemma 7.7.

To this end, note that $\mathrm{Fix}_P\big(\mathrm{Syl}_p(G)\big) = \big\{Q \in \mathrm{Syl}_p(G) : gQg^{-1} = Q \text{ for all } g \in P\big\}$, so we certainly have that $P \in \mathrm{Fix}_P\big(\mathrm{Syl}_p(G)\big)$. If $Q \in \mathrm{Fix}_P\big(\mathrm{Syl}_p(G)\big)$, then $P \leq N_G(Q)$. Since $Q \leq N_G(Q)$ as well, we see that P and Q are p-Sylow subgroups of $N_G(Q)$ and thus are conjugate. But $Q \lhd N_G(Q)$, so it must be the case that $Q = P$, and the assertion follows.

For the remaining claims, let G act on $\mathrm{Syl}_p(G)$ by conjugation. Theorem 7.6 implies that $\mathrm{Syl}_p(G) = \mathcal{O}(P)$ and the corresponding stabilizer is $N_G(P)$, hence $n_p = |\mathcal{O}(P)| = [G : N_G(P)]$. As this must divide $|G| = p^k m$ and we know that $n_p \equiv 1 \pmod p$, we conclude that $n_p \mid m$. □

Fact 7.1. *Given $M \in \mathbb{C}^{d \times d}$, define $\mathrm{Tr}(M) = \sum_{k=1}^{d} M_{k,k}$.*

(1) If $A \in \mathbb{C}^{m \times n}$ and $B \in \mathbb{C}^{n \times m}$, then $\mathrm{Tr}(AB) = \mathrm{Tr}(BA)$.

(2) If $T : V \to V$ is a linear transformation, B is the matrix for T with respect to a basis $\mathcal{B}$, and C is the matrix for T with respect to a basis $\mathcal{C}$, then $\mathrm{Tr}(B) = \mathrm{Tr}(C)$.

(3) If $A \in \mathbb{C}^{n \times n}$ has eigenvalues (counting multiplicity) $\lambda_1, \ldots, \lambda_n$, then $\mathrm{Tr}(A) = \sum_{k=1}^{n} \lambda_k$.

Proof.

(1)

$$\begin{aligned} \mathrm{Tr}(AB) &= \sum_{k=1}^{m} (AB)_{k,k} = \sum_{k=1}^{m} \sum_{\ell=1}^{n} A_{k,\ell} B_{\ell,k} \\ &= \sum_{\ell=1}^{n} \sum_{k=1}^{m} B_{\ell,k} A_{k,\ell} = \sum_{\ell=1}^{n} (BA)_{\ell,\ell} = \mathrm{Tr}(BA). \end{aligned}$$

(2) Let P be the change-of-basis matrix from $\mathcal{B}$-coordinates to $\mathcal{C}$-coordinates. Then $C = PBP^{-1}$, so

$$\begin{aligned}\operatorname{Tr}(C) &= \operatorname{Tr}(PBP^{-1}) = \operatorname{Tr}\big((PB)P^{-1}\big)\\ &= \operatorname{Tr}\big(P^{-1}(PB)\big) = \operatorname{Tr}\big((P^{-1}P)B\big) = \operatorname{Tr}(B).\end{aligned}$$

(3) Let A be given in Jordan normal form as $A = PJP^{-1}$. Then J is upper-triangular with the eigenvalues of A on its main diagonal, so

$$\operatorname{Tr}(A) = \operatorname{Tr}(PJP^{-1}) = \operatorname{Tr}(J) = \sum_{k=1}^{n} J_{k,k} = \sum_{k=1}^{n} \lambda_k.$$

Alternatively, the coefficient of λ^{n-1} in $\varphi(\lambda) = \det(\lambda I - A)$ is $-\operatorname{Tr}(A)$, and the coefficient of λ^{n-1} in $\varphi(\lambda) = \prod_{k=1}^{n}(\lambda - \lambda_k)$ is $-\sum_{k=1}^{n}\lambda_k$.

□

Fact 7.2 (Rational Roots Test)**.** *Let $p(z) = a_n z^n + \cdots + a_1 z + a_0$ with $a_0, \ldots, a_n \in \mathbb{Z}$. If $r, s \in \mathbb{Z}$ satisfy $(r, s) = 1$ and $p(r/s) = 0$, then $r \mid a_0$ and $s \mid a_n$.*

Proof. Multiplying $a_n(r/s)^n + \cdots + a_1(r/s) + a_0 = 0$ by s^n gives $a_n r^n + a_{n-1} r^{n-1} s \cdots + a_1 r s^{n-1} + a_0 s^n = 0$. It follows that $-a_0 s^n = r(a_n r^{n-1} + \cdots + a_1 s^{n-1})$, hence $r \mid a_0$, and $-a_n r^n = s(a_{n-1} r^{n-1} + \cdots + a_0 s^{n-1})$, hence $s \mid a_n$. □

Lemma 7.8. *$y \in \mathbb{C}$ is an algebraic integer if and only if there exist $w_1, \ldots, w_n \in \mathbb{C}$, not all zero, such that $w_i y = \sum_{j=1}^{n} b_{ij} w_j$ for some integers $\{b_{ij}\}_{i,j=1}^{n}$.*

Proof. If y is an algebraic integer, then there exist $a_0, \ldots, a_{n-1} \in \mathbb{Z}$ with $y^n + a_{n-1}y^{n-1} + \cdots + a_1 y + a_0 = 0$. Taking $w_i = y^{i-1}$ for $i \in [n]$ gives $w_i y = y^i = w_{i+1}$ for $1 \le i < n$ and $w_n y = y^n = -a_0 w_1 - \cdots - a_{n-1} w_n$.

Conversely, suppose $w_i y = \sum_{j=1}^{n} b_{ij} w_i$ and let B be the $n \times n$ matrix with (i, j)-entry b_{ij}. Setting $\mathbf{w} = \begin{bmatrix} w_1 \\ \vdots \\ w_n \end{bmatrix}$, we have $B\mathbf{w} = y\mathbf{w}$, so y is an eigenvalue of the integer matrix B and thus a root of the monic polynomial $p(z) = \det(zI - B)$. □

Proposition 7.2. *The algebraic integers form a subring of $\mathbb{C}$.*

Proof. $1 \in \mathbb{A}$ since it solves $z - 1 = 0$. Suppose $y, z \in \mathbb{A}$. By Lemma 7.8, there are $w_1, \ldots, w_m, x_1, \ldots, x_n \in \mathbb{C}$, $\{b_{ij}\}_{i,j=1}^{m}, \{c_{k\ell}\}_{k,\ell=1}^{n} \subseteq \mathbb{Z}$ satisfying $w_i y = \sum_{j=1}^{m} b_{ij} w_j$ and $x_k z = \sum_{\ell=1}^{n} c_{k\ell} x_\ell$.

It follows that $w_i(-y) = \sum_{j=1}^{m}(-b_{ij})w_j$, hence $-y \in \mathbb{A}$; $(w_i x_k)(y + z) = \sum_{j=1}^{m} b_{ij} w_j x_k + \sum_{\ell=1}^{n} c_{k\ell} w_i x_\ell$, hence $y + z \in \mathbb{A}$; and $(w_i x_k)(yz) = (w_i y)(x_k z) = \Big(\sum_{j=1}^{m} b_{ij} w_j\Big)\Big(\sum_{\ell=1}^{n} c_{k\ell} x_\ell\Big) = \sum_{j=1}^{m}\sum_{\ell=1}^{n}(b_{ij}c_{k\ell})(w_j x_\ell)$, hence $yz \in \mathbb{A}$. □

Proposition 7.3. *The maps* $\mathrm{Res}_H^G : \mathcal{C}(G) \to \mathcal{C}(H)$ *and* $\mathrm{Ind}_H^G : \mathcal{C}(H) \to \mathcal{C}(G)$ *are linear.*

Proof. If $\Phi \in \mathcal{C}(G)$ and $x, y \in H$, then $xyx^{-1} \in H$, hence

$$\mathrm{Res}_H^G \Phi(xyx^{-1}) = \Phi(xyx^{-1}) = \Phi(y) = \mathrm{Res}_H^G \Phi(y).$$

If $\Psi \in \mathcal{C}(G)$ and $\alpha, \beta \in \mathbb{C}$ as well, then

$$\mathrm{Res}_H^G \big[\alpha\Phi + \beta\Psi\big](x) = \alpha\Phi(x) + \beta\Psi(x) = \alpha \mathrm{Res}_H^G \Phi(x) + \beta \mathrm{Res}_H^G \Psi(x).$$

If $\phi \in \mathcal{C}(H)$ and $x, y \in G$, then

$$\mathrm{Ind}_H^G \phi(xyx^{-1}) = \frac{1}{|H|} \sum_{z \in G} \widetilde{\phi}(z^{-1}xyx^{-1}z) = \frac{1}{|H|} \sum_{z \in G} \widetilde{\phi}(w^{-1}yw) = \mathrm{Ind}_H^G \phi(y),$$

where we made use of the reindexing $w = x^{-1}z$.

If $\psi \in \mathcal{C}(H)$ and $\alpha, \beta \in \mathbb{C}$ as well, then

$$\begin{aligned} \mathrm{Ind}_H^G \big[\alpha\phi + \beta\psi\big](x) &= \frac{1}{|H|} \sum_{z \in G} \widetilde{\alpha\phi + \beta\psi}(z^{-1}xz) = \frac{1}{|H|} \sum_{z \in G} \big[\alpha\widetilde{\phi}(z^{-1}xz) + \beta\widetilde{\psi}(z^{-1}xz)\big] \\ &= \alpha \frac{1}{|H|} \sum_{z \in G} \widetilde{\phi}(z^{-1}xz) + \beta \frac{1}{|H|} \sum_{z \in G} \widetilde{\psi}(z^{-1}xz) = \alpha \mathrm{Ind}_H^G \phi(x) + \beta \mathrm{Ind}_H^G \psi(x). \end{aligned}$$

□