Notes for a MSc Course in Group Theory

Andrea Caranti

Dipartimento di Matematica, Università degli Studi di Trento,
via Sommarive 14, 38123 Trento

Email address: `andrea.caranti@unitn.it`

URL: `https://caranti.maths.unitn.it/`

Introduction

In the Spring semester of 2018–19 I am giving for the first time a course (6 ECTS) in "Advanced Group Theory" for the MSc in Mathematics in Trento (the MSc is in English, in case you were wondering). This is meant as a second course in group theory, after a 6 ECTS course in Group Theory in the BSc (see the notes [**Car19**], in Italian), which followed a 12 ECTS basic course in Algebra (see the notes [**CM19**], in Italian), and a 6 ECTS course in Galois Theory.

Before and during the course, I plan to write up some notes, which are what you are reading right now. The most recent version can be downloaded from

https://caranti.maths.unitn.it/

Useful Texts

Among the many good texts in group theory, I recommend

- [**Mac12**] (I graduated with the author);
- [**Hup67**], an excellent text in German;
- [**Gor80**], another great classic;
- [**Rob96**], complete and crystal clear;
- [**Rot95**], a nice selection of arguments;
- [**Ser78**]/[**Ser77**], a compact classic;
- [**Ser16**], a bit intense, but magnificent;
- [**Isa06**], a great, a bit demanding classic.

For the course, in the end I am mainly using [**Ser16**], with some arguments taken from [**Isa06**], and sometimes [**Ser78**].

An excellent text for general algebra (covering for instance tensor products) is [**Lan02**].

To be written up next

Normal subgroups and Clifford?

Contents

Introduction 3

Useful Texts 3

To be written up next 3

Part 1. Preliminaries 7

Chapter 1. Preliminaries 9

1.1. The general linear group, and projections 9

1.2. The trace 9

1.3. Algebras over a field 10

1.4. The centre of the matrix algebra 10

1.5. Simultaneous diagonalization 11

1.6. Inner products 11

1.7. The class equation and the centre of finite p-groups 14

1.8. Endomorphisms, automorphisms and inner automorphisms 14

1.9. Elementary abelian groups 15

1.10. Modules 17

1.11. Integrality and algebraic integers 18

1.12. Another approach to algebraic integers 23

1.13. Finitely generated $\mathbf{Z}$-modules 24

1.14. Tensor Products 25

Part 2. Soluble and nilpotent groups 31

Chapter 2. Series 33

2.1. Series 33

2.2. Ω-groups and Ω-series 33

2.3. Ω-composition series 34

2.4. Uniqueness of the factors of an Ω-composition series 36

Chapter 3. Soluble groups 37

3.1. Commutators 37

3.2. The derived series and soluble groups 39

Chapter 4. Nilpotent groups 41

4.1. Central series and nilpotent groups 41

4.2. Finite nilpotent groups 43

4.3. Nilpotent groups are soluble 43

Part 3. Representations and Characters 45

Chapter 5. Representations 47
5.1. Permutation representations 47
5.2. Linear representations 47
5.3. Inner products 47
5.4. From permutation representations to linear representations 48
5.5. Subrepresentations 49
5.6. Examples 49
5.7. The group algebra 50
5.8. Maschke's Theorem 51
5.9. Irreducible representations 53
5.10. Morphisms and modules over the group algebra 53
5.11. Schur's Lemma 54
5.12. Orthogonality Relations 54

Chapter 6. Characters 57
6.1. Characters 57
6.2. Decomposing the regular representation 58
6.3. Number of irreducible characters 59
6.4. Representation and characters from quotients 60
6.5. Products of representations and characters 60
6.6. Kernels and centres 61
6.7. Character tables 62
6.8. Characters of abelian groups 62
6.9. Induced Characters 63
6.10. Characters of permutation representations 67
6.11. Character tables of small groups 68
6.12. A non-linear character vanishes somewhere 72
6.13. Integrality 73

Chapter 7. Applications 79
7.1. Burnside $p^a q^b$ 79
7.2. Frobenius groups and the Frobenius kernel 80
7.3. Groups with an abelian Sylow p-subgroup 83

Bibliography 85

Part 1

Preliminaries

CHAPTER 1

Preliminaries

Recalling some basic stuff, and establishing notation.

1.1. The general linear group, and projections

If V is a vector space over some field F, then $\mathrm{GL}(V)$ denotes the *general linear group* of invertible linear maps on V.

If V is of finite dimension n, and we fix a basis, then the elements of $\mathrm{GL}(V)$ can be represented via $n \times n$ matrices. We denote by $\mathrm{GL}(n, F)$ the corresponding group of matrices, which acts on the space F^n of row vectors. When F is finite of order q (this is not going to happen in these notes), one writes also $\mathrm{GL}(n, q)$.

Recall that if $V = U \oplus W$, the projection on U along W is the linear map

$$\begin{aligned} \pi : V &\to V \\ u + w &\mapsto u, \end{aligned}$$

where $u \in U$ and $w \in W$. Clearly $(u+w)\pi^2 = u\pi = (u+0)\pi = u$, so that $\pi^2 = \pi$. Conversely we have

1.1.1. Lemma. *Let V be a vector space.*
Let π be a linear map on V such that $\pi^2 = \pi$.
Then

$$V = V\pi \oplus \ker(\pi),$$

and π is the projection on $V\pi$ along $\ker(\pi)$.

Proof. Let $v \in V$. Then $(v - v\pi)\pi = v\pi - v\pi^2 = 0$, so that $v - v\pi \in \ker(\pi)$. If $v \in \ker(\pi) \cap V\pi$, then $v = u\pi$ for some $u \in V$, and thus $v = u\pi = u\pi^2 = v\pi = 0$. □

1.2. The trace

It is a well-known and elementary fact that the trace (of a square matrix) satisfies $\mathrm{trace}(AB) = \mathrm{trace}(BA)$ for all matrices A, B. In particular, if C is invertible we have

$$\mathrm{trace}(C^{-1}AC) = \mathrm{trace}(C^{-1}(AC)) = \mathrm{trace}((AC)C^{-1}) = \mathrm{trace}(A).$$

Note actually that this reminds us that the trace is defined for a linear map on a vector space, and it does not depend on the base with respect to which one writes it down as a matrix.

1.3. Algebras over a field

1.3.1. Definition. An *algebra* over the field F is a ring $A \neq \{ 0 \}$ with unity, which is also a vector space over F (with respect to the same "+" operation), such that for $a, b \in A$, $\lambda \in F$,

$$\lambda(ab) = (\lambda a)b = a(\lambda b). \tag{1.3.1}$$

It follows from the definitions that the ring product in an algebra is bilinear, that is, for $a, b \in A$ and $\lambda \in F$ we have

$$(\lambda a + \lambda b)c = (\lambda(a+b))c = \lambda((a+b)c) = \lambda(ac+bc) = \lambda(ac) + \lambda(bc),$$

and similarly on the right.

Note that if we take $a = 1$ in (1.3.1), we obtain first of all for all $b \in A$

$$(\lambda \cdot 1)b = \lambda(1 \cdot b) = \lambda b = \lambda(b \cdot 1) = b(\lambda \cdot 1).$$

This says first of all that $F \cdot 1 = \{ \lambda \cdot 1 : \lambda \in F \}$ is in the centre of A. And then in particular we have

$$\lambda b = (\lambda \cdot 1)b. \tag{1.3.2}$$

Since $A \neq \{ 0 \}$ by definition, the axioms of vector spaces then imply that

$$\begin{aligned} F &\to A \\ \lambda &\mapsto \lambda \cdot 1 \end{aligned}$$

is an injective morphism of rings with unity. (This is because a field F has only the ideals $\{ 0 \}$ and F, so a ring morphism from F to another ring either maps F to zero, or is injective. But in a vector space, multiplication by the scalar 1 is the identity.) Therefore $F \cdot 1 = \{ \lambda \cdot 1 : \lambda \in F \}$ is a subring of A isomorphic to F, which is often identified with F, because (1.3.2) shows that scalar multiplication by λ, or multiplication by $\lambda \cdot 1$ in A, are the very same thing.

1.3.2. Example. The $n \times n$ matrices over a field A are an algebra. The identification just mentioned means that we consider $\lambda \in F$ and the scalar matrix $\lambda \cdot I$ with λ on the diagonal as the same thing.

1.4. The centre of the matrix algebra

Recall that if G is a group, then its centre is

$$Z(G) = \{ z \in G : gz = zg, \text{ for all } g \in G \}.$$

If A is a ring, its centre is

$$Z(A) = \{ z \in A : az = za, \text{ for all } a \in A \}.$$

1.4.1. Proposition. *Show that the centre of the algebra of $n \times n$ matrices over a field F consists of the scalar matrices.*

Proof. Let A be the algebra of $n \times n$ matrices, and $z \in Z(A)$. Proceeding by way of contradiction, assume there is a vector $v \neq 0$ such that v and vz are independent. Complete v, vz to a basis, and consider the linear map T that is zero on all basis vectors, except that it takes vz to v. Then $v(zT) = (vz)T = v$, but $v(Tz) = (vT)z = 0z = 0 \neq v$, a contradiction.

Thus for all $0 \neq v \in V$ there is a uniquely defined $a(v) \in F$ such that $vz = a(v)z$. (We are saying that all non-zero vectors are eigenvectors for z.) If v, w are independent vectors then

$$\begin{aligned}a(v+w)v + a(v+w)w = a(v+w)(v+w) = (v+w)z = \\ = vz + wz = a(v)v + a(w)w,\end{aligned}$$

whence

$$a(v) = a(v+w) = a(w),$$

and z is a scalar matrix. □

1.4.2. Exercise. *Find other proofs for this elementary, but basic, fact. (Books are OK, but Internet search is also allowed.)*

1.4.1. Centre of the general linear group. $GL(V)$, the general linear group, is the group of invertible linear maps on the finite-dimensional vector space V. Its centre consists of the non-zero scalar matrix. A proof is a variant of the previous one, in which one takes at T the linear map that is the identity on all basis elements, except that $(vz)T = vz + v$. One checks easily that this is indeed invertible, and then $vTz = vz \neq vz + v = vzT$.

An alternative proof over an infinite field F consists in noting that any matrix A can be written as the difference of two invertible matrices. If $\lambda \in F$ is *not* an eigenvalue of A, we have in fact $A = \lambda\mathbf{1} + (A - \lambda\mathbf{1})$. Thus if a matrix commutes with all the invertible matrices, it commutes with all the matrices.

1.5. Simultaneous diagonalization

Let $A_o, A_1, \dots, A_k$ be $n \times n$ matrices over a field F, such that each of them is diagonalizable, and the matrices commute pairwise, that is $A_i A_j = A_j A_i$. then there is a basis with respect to which all A_i are diagonal.

Let $V = F^n$. For each eigenvalue λ of $B = A_0$, consider the eigenspace $V(\lambda) = \{ v \in V : vB = \lambda v \}$. By hypothesis, V is a direct sum of the $V(\lambda)$.

We claim that for each λ and $i \geq 1$ we have $V(\lambda)A_i \subseteq V(\lambda)$. In fact if $v \in V(\lambda)$ one has

$$(vA_i)B = v(A_iB) = v(BA_i) = (vB)A_i = \lambda(vA_i),$$

so that $vA_i \in V(\lambda)$. By induction on k, the restrictions of the A_i, for $i \geq 1$ to $V(\lambda)$ can be simultaneously diagonalized. As the restriction of B to $V(\lambda)$ is scalar with respect to any base (see Section 1.4), all the A_i are then diagonalized at once.

1.6. Inner products

Let V be a finite-dimensional vector space over $\mathbf{C}$. An *inner product* on V is a map

$$\langle \cdot, \cdot \rangle : V \times V \to \mathbf{C}$$

which satisfies the following properties.

(1) For $\lambda \in \mathbf{C}$, $x, y, z \in V$

$$\langle x, \lambda y \rangle = \lambda \langle x, y \rangle, \qquad \langle x, y+z \rangle = \langle x, y \rangle + \langle x, z \rangle.$$

This states that $\langle \cdot, \cdot \rangle$ is linear in the second component.

(2) For $x, y \in V$

$$\langle x, y \rangle = \overline{\langle y, x \rangle}.$$

Here the bar denotes complex conjugation. This implies that

(3) for $x \in V$

$$\langle x, x \rangle = \overline{\langle x, x \rangle}$$

is real, and

(4) $\langle \cdot, \cdot \rangle$ is semilinear (antilinear, conjugate-linear) in the first component, that is, for $\lambda \in \mathbf{C}$, $x, y, z \in V$

$$\langle \lambda x, y \rangle = \overline{\lambda} \langle x, y \rangle, \qquad \langle x+y, z \rangle = \langle x, z \rangle + \langle y, z \rangle.$$

The first part of (4) follows from

$$\langle \lambda x, y \rangle = \overline{\langle y, \lambda x \rangle} = \overline{\lambda \langle y, x \rangle} = \overline{\lambda}\, \overline{\langle y, x \rangle} = \overline{\lambda} \langle x, y \rangle.$$

(5) For $x \in V$

$$\langle x, x \rangle \geq 0,$$

and

(6) for $x \in V$

$$\langle x, x \rangle = 0 \quad \text{if and only if} \quad x = 0. \tag{1.6.1}$$

Note that (1.6.1) implies

(7) for $x \in V$

$$\langle x, y \rangle = 0 \text{ for all } y \in V \quad \text{if and only if} \quad x = 0. \tag{1.6.2}$$

If $V = \mathbf{C}^n$, the standard inner product is given by

$$\langle x, y \rangle = \overline{x} \cdot y^t = \sum_{i=1}^{n} \overline{x_i} y_i. \tag{1.6.3}$$

If V has an inner product $\langle \cdot, \cdot \rangle$, we can build an *orthonormal basis* with the *Gram-Schmidt* process. An orthonormal basis is a basis $e_1, \dots, e_n$ such that $\langle e_i, e_j \rangle = \delta_{ij}$ for all i, j, where δ_{ij} is the Kronecker delta.

Start with an arbitrary basis $v_1, \dots, v_n$, and define

$$e_1 = \frac{1}{\langle v_1, v_1 \rangle^{1/2}} v_1,$$

so that $\langle e_1, e_1 \rangle = 1$.

Assume you have defined $e_1, \dots, e_{k-1}$ as linear combinations of $v_1, \dots, v_{k-1}$, for $k \leq n$. Define first

$$v_k' = v_k - \langle v_k, e_1 \rangle e_1 - \dots - \langle v_k, e_{k-1} \rangle e_{k-1},$$

a vector which is independent of $e_1, \dots, e_{k-1}$, so that for $1 \leq i < k$ we have

$$\langle v_k', e_i \rangle = \langle v_k, e_i \rangle - \langle v_k, e_i \rangle \langle e_i, e_i \rangle = 0.$$

Then normalise

$$e_k = \frac{1}{\langle\, v_k', v_k' \,\rangle^{1/2}}\, v_k'.$$

If V has an inner product $\langle\, \cdot, \cdot \,\rangle$, and U is a subspace with orthonormal basis $f_1, \dots, f_k$, then the map

$$\begin{aligned} p : V &\to U \\ v &\mapsto \sum_{i=1}^{k} \langle\, v, f_i \,\rangle\, f_i \end{aligned}$$

is a projection onto U along the subspace

$$U^\perp = \langle\, v \in V : \langle\, v, u \,\rangle = 0 \text{ for all } u \in U \,\rangle\,.$$

In fact $Vp \subseteq U$, and if

$$u = \sum_{i=1}^{k} a_i f_i \in U,$$

then $\langle\, u, f_i \,\rangle = a_i$, so that $up = u$. Moreover $v \in \ker(p)$ if and only if $\langle\, v, f_i \,\rangle = 0$ for all i, that is, $v \in U^\perp$.

Incidentally, this shows that $V = U \oplus U^\perp$. This also follows from the two facts

(1) (1.6.1) implies $U \cap U^\perp = \{\, 0 \,\}$, and
(2) the condition $\langle\, x, u \,\rangle = 0$ for all $u \in U$ translates into a homogeneous linear system in the coordinates of x. The rank of the matrix of the system equals to the dimension of U. Therefore $U^\perp$, which is the set of solutions, has dimension $\dim(V) - \dim(U)$.

With respect to an orthonormal basis, an inner product takes the standard form (1.6.3).

Given an inner product on V, and $A \in \mathrm{End}_F(V)$, one can define the *adjoint* A^* of A as the unique $B \in \mathrm{End}_F(V)$ such that $\langle\, xB, y \,\rangle = \langle\, x, yA \,\rangle$ for all $x, y \in V$. In fact, every linear map $V \to \mathbf{C}$, that is, every element of the dual space V^* can be represented in the form

$$y \mapsto \langle\, x, y \,\rangle$$

for some $x \in V$. This is because (1.6.2) shows that the dimension of the latter maps is exactly $\dim(V) = \dim(V^*)$. Since for each $x \in V$ the map

$$y \mapsto \langle\, x, yA \,\rangle$$

is linear, we will have, for all $x, y \in V$

$$\langle\, x, yA \,\rangle = \langle\, xB, y \,\rangle$$

for a unique map $B : V \to V$. But then B is linear, as for all $x, y, z \in V$ we have

$$\begin{aligned} \langle\, (\lambda x + \mu y)B, z \,\rangle &= \langle\, \lambda x + \mu y, zA \,\rangle \\ &= \overline{\lambda}\, \langle\, x, zA \,\rangle + \overline{\mu}\, \langle\, y, zA \,\rangle \\ &= \overline{\lambda}\, \langle\, xB, z \,\rangle + \overline{\mu}\, \langle\, yB, z \,\rangle \\ &= \langle\, \lambda(xB) + \mu(yB), z \,\rangle\,. \end{aligned}$$

Given the standard inner product, and $A \in \mathrm{GL}(V)$, we have

$$\langle\, x, yA \,\rangle = \overline{x} \cdot (yA)^t = \overline{x} \cdot A^t y^t = \overline{x\overline{A}^t} y^t = \left\langle\, x\overline{A}^t, y \,\right\rangle,$$

so that the adjoint of A is

$$A^* = \overline{A}^t.$$

$A \in \mathrm{GL}(V)$ is *unitary* if

$$\langle\, xA, yA \,\rangle = \langle\, x, y \,\rangle \qquad \text{for all } x, y \in V.$$

In other words, A is unitary if and only if for all $x, y \in V$ one has

$$\langle\, xA, yA \,\rangle = \langle\, xAA^*, y \,\rangle = \langle\, x, y \,\rangle,$$

that is $A^* = A^{-1}$. If $0 \neq \lambda \in \mathbf{C}$ is an eigenvalue of a unitary A, with eigenvector v, we have

$$\langle\, v, v \,\rangle = \langle\, vA, vA \,\rangle = \langle\, v\lambda, v\lambda \,\rangle = \lambda\overline{\lambda} \langle\, v, v \,\rangle,$$

so that λ has absolute value 1.

1.7. The class equation and the centre of finite p-groups

Let G be a finite group acting on itself by conjugation. Since the orbits (that is, the conjugacy classes) form a partition, we have the *class equation*

$$|\, G \,| = |\, Z(G) \,| + \sum |\, G : C_G(a) \,|,$$

where the sum is over a set of representatives of the conjugacy classes $a^G \neq \{\, a \,\}$, so that each $|\, G : C_G(a) \,| = \left| a^G \right| > 1$.

If $G \neq \{\, 1 \,\}$ is a finite p-group, then $p \mid |\, G \,|$ and p divides each $|\, G : C_G(a) \,| = \left| a^G \right|$. Thus $p \mid |\, Z(G) \,|$, so that $|\, Z(G) \,| > 1$.

1.8. Endomorphisms, automorphisms and inner automorphisms

Let G be a group. A (homo)morphism $G \to G$ is called an *endomorphism* of G. The set of all endomorphisms of a group G is denoted by $\mathrm{End}(G)$, and is a monoid under the composition $\circ$ of map.

1.8.1. Lemma. *If $(G, +, 0)$ is an* abelian *group, then $\mathrm{End}(G)$ becomes a ring under pointwise addition*

$$g^{s+t} = g^s + g^t, \quad \textit{for } g \in G \textit{ and } s, t \in \mathrm{End}(G),$$

and composition of maps

$$g^{s \circ t} = (g^s)^t, \quad \textit{for } g \in G \textit{ and } s, t \in \mathrm{End}(G).$$

1.8.2. Exercise. *Prove the Lemma. The point where the hypothesis that G is abelian plays a crucial role is when proving that if $s, t \in \mathrm{End}(G)$, then $s + t \in \mathrm{End}(G)$.*

1.8.3. Exercise. *If $(G, \cdot, 1)$ is any group, and s, t are* maps *on G, one can define a map $s + t$ on G by $g^{s+t} = g^s \cdot g^t$.*

Let I be the identity map on G, so that $I \in \mathrm{End}(G)$. Show that $I + I \in \mathrm{End}(G)$ if and only if G is abelian.

(Hint: Note that for $g \in G$ one has $g^{I+I} = g^I \cdot g^I = g^2$. So one has to prove that the map $g \mapsto g^2$ is a homomorphism if and only if G is abelian.)

An invertible element of $\mathrm{End}(G)$, that is, an endomorphism which is a bijective map, is called an *automorphism* of G. The automorphisms of G form thus a group $\mathrm{Aut}(G)$.

For $g \in G$, the map

$$\begin{aligned} \iota(g) : G &\to G \\ x &\mapsto g^{-1}xg \end{aligned}$$

is an automorphism of G, the *inner automorphism* induced by g. The map

$$\begin{aligned} \iota : G &\to \mathrm{Aut}(G) \\ g &\mapsto \iota(g) \end{aligned}$$

is a group morphism.

The image of ι is the group $\mathrm{Inn}(G)$ of the *inner automorphisms* of G

1.8.4. Exercise. *Prove these statements.*

(Hint: It is possibly better to start by proving that each $\iota(g)$ is an endomorphism of G. Then show that $\iota(gh) = \iota(g)\iota(h)$ and $\iota(1) = 1$, and then show that $\iota(g^{-1}) = \iota(g)^{-1}$.)

1.8.5. Exercise. *Prove that* $\mathrm{Inn}(G) \trianglelefteq \mathrm{Aut}(G)$.

(Hint: Let $g \in G$ and $\varphi \in \mathrm{Aut}(G)$. For $x \in G$ we have

$$x^{\iota(g)^\varphi} = x^{\varphi^{-1}\iota(g)\varphi} = (g^{-1}x^{\varphi^{-1}}g)^\varphi = (g^\varphi)^{-1}xg^\varphi = x^{\iota(g^\varphi)},$$

so that $\iota(g)^\varphi = \iota(g^\varphi) \in \mathrm{Inn}(G)$.)

Now

$$\ker(\iota) = \{\, g \in G : \iota(g) = 1 \,\} = \left\{ g \in G : g^{-1}xg = x \text{ for all } x \in G \right\} = Z(G),$$

the centre of G. The first isomorphism theorem implies

$$\mathrm{Inn}(G) \cong G/Z(G).$$

1.9. Elementary abelian groups

Let p be a prime. A finite abelian group G such that $x^p = 1$ for all $x \in G$ is said to be *elementary abelian.* So if G has order p^n, it is isomorphic to the direct product of n (cyclic) groups of order p.

1.9.1. Lemma. *Let $(V, +, 0)$ be an abelian group, and F a field. The following are equivalent*

(1) a structure of F-vector space on $(V, +, 0)$, and

(2) a homomorphism of rings with unity $F \to \mathrm{End}(V)$.

Proof. Suppose $\varphi : F \to \mathrm{End}(V)$ is a homomorphism of rings with unity. For $a \in F$ and $v \in V$, denote by $va = v^{\varphi(a)}$, the image of v under the endomorphism $\varphi(a)$. Let us see what this means.

Every $\varphi(a)$ is an endomorphism of V, that is

(1) for $u, v \in V$ and $a \in F$, one has

$$(u+v)a = (u+v)^{\varphi(a)} = u^{\varphi(a)} + v^{\varphi(a)} = ua + va.$$

Then $\varphi(a+b) = \varphi(a) + \varphi(b)$, that is

(2) for $a, b \in F$ and $v \in V$ one has

$$v(a+b) = v^{\varphi(a+b)} = v^{\varphi(a)+\varphi(b)} = v^{\varphi(a)} + v^{\varphi(b)} = va + vb.$$

We have used the fact that addition of $f, g \in \mathrm{End}(V)$ is defined pointwise, that is, for $v \in V$

$$v^{f+g} = v^f + v^g.$$

Now multiplication in $\mathrm{End}(V)$ is defined as composition of maps, so that $\varphi(ab) = \varphi(a) \circ \varphi(b)$, that is

(3) for $a, b \in F$ and $v \in V$ one has

$$v(ab) = v^{\varphi(ab)} = v^{\varphi(a)\circ\varphi(b)} = (v^{\varphi(a)})^{\varphi(b)} = (va)^{\varphi(b)} = (va)b.$$

Finally,

(4) $v1 = v^{\varphi(1)} = v^I = v$, where 1 denotes the unity of F and I the identity map on V.

So we see that (1)–(4) are just the axioms of an F-vector space.

The converse is immediate. □

1.9.2. Theorem. *Let $(V, +, 0)$ be an elementary abelian group of order $p^n > 1$, where p is a prime. For $a \in \mathbf{Z}$ and $v \in V$, the notation va stands for the a-th multiple of v.*

Let $F = \mathbf{Z}/p\mathbf{Z}$ be the field with p elements.

Then for $a \in \mathbf{Z}$ and $v \in V$ the operation

$$v(a + p\mathbf{Z}) = va \tag{1.9.1}$$

defines a multiplication by scalars that makes V into an F-vector space.

Proof. The map

$$\begin{aligned} \psi : \mathbf{Z} &\to \mathrm{End}(V) \\ a &\mapsto (v \mapsto va) \end{aligned}$$

is a morphism of rings with unity, by the properties of multiples.

Since $V \neq \{0\}$, and $vp = 0$ for all $v \in V$, its kernel is $p\mathbf{Z}$. The first isomorphism theorem for rings yields a morphism of rings with unity as in (1.9.1), so that we may appeal to lemma 1.9.1. □

Recall that if V is a finite-dimensional vector space over a field F, then $\mathrm{GL}(V)$ denotes the group of invertible linear maps on the V. If F is the field with p elements, then, once a basis is chosen, $\mathrm{GL}(V)$ is isomorphic to the group $\mathrm{GL}(n,p)$ of $n \times n$ matrices with non-zero determinant in F.

1.9.3. Proposition. *Let $(V,+,0)$ be an elementary abelian group of order p^n, where p is a prime, so that V can be regarded as a vector space on the field with p elements.*

Then the group $\mathrm{Aut}(V)$ *of automorphisms of V coincides with the group* $\mathrm{GL}(V) \cong \mathrm{GL}(n,p)$.

Proof. If $\beta \in \mathrm{Aut}(V)$, then for $a \in \mathbf{Z}$ and $v \in V$ one has

$$(v(a+p\mathbf{Z}))^\beta = (va)^\beta = v^\beta a = v^\beta(a+p\mathbf{Z}),$$

so that β is also a linear map on the vector space V. □

1.10. Modules

On the model of Lemma 1.9.1, we give the following

1.10.1. Definition. Let $(M,+,0)$ be an abelian group, R a unital ring.
An *R-module structure on M* is a morphism of unital rings $A \to \mathrm{End}(M)$.

1.10.2. Exercise. *Show that this definition is equivalent to requiring that there is a map*

$$\begin{aligned} M \times R &\to R \\ (m,r) &\mapsto mr \end{aligned}$$

which satisfies the axioms, for $m, m_1, m_2 \in M$ and $r, r_1, r_2 \in R$:

(1) $(m_1+m_2)r = m_1r + m_2r$
(2) $m(r_1+r_2) = mr_1 + mr_2$
(3) $m(r_1r_2) = (mr_1)r_2$
(4) $m1 = m$

1.10.3. Exercise. *Let R be a unital ring, $M = (R,+,0)$ its abelian group. Show that M becomes a right R-module by taking*

$$\begin{aligned} M \times R &\to R \\ (m,r) &\mapsto mr \end{aligned}$$

and a left R-module by taking

$$\begin{aligned} M \times R &\to R \\ (m,r) &\mapsto mr \end{aligned}$$

1.10.4. Remark. If R is an algebra over the field F, then an R-module is also an F-module, and thus a vector space over F.

1.10.5. Remark. These are *right* modules, as we compose maps left-to-right, so that (3) holds. If we compose maps right-to-left, then we write rm instead of mr, and replace (3) with

(3') $(r_1 r_2)m = r_1(r_2 m)$.

When R is a commutative ring, there is no difference between left and right modules.

If M is any abelian group, $m \in M$ and $r \in \mathbf{Z}$, the map

$$r \mapsto (m \mapsto mr),$$

which takes the integer r to the the map that takes $m \in M$ to its r-th multiple, satisfies the axioms, and thus turns M into a Z-module. Conversely, a $\mathbf{Z}$-module structure on an abelian group M is completely determined by the group structure, as (4) and (2) yield that mr is indeed the r-th multiple of m.

The definition of a module is an analogue of that of a vector space, and specialises to that one when R is a field. The main (and very important!) difference is that a module need not have a basis. (The definition of a basis for modules is exactly the same as for vector space.) For instance, the $\mathbf{Z}$-module $M = \mathbf{Z}/2\mathbf{Z} = \{ [0], [1] \}$. Does not have a basis. The only candidate would be $[1]$, which generates M, but $1[1] = 3[1] = 5[1] = \dots$ shows that uniqueness of representation ("linear independence") does not hold.

1.10.6. Definition. An R-module is said to be *free* if it has a basis.

So for instance $R^n = R \times \cdots \times R$ is free, with the usual basis as in the case of vector spaces.

The following definition is an analogue of that of subspace

1.10.7. Definition. Let M be an R-module. A subgroup N of M is said to be an R-submodule (or simply a submodule) if for $n \in N$ and $r \in R$ we have $nr \in N$. It follows that an R-submodule is an R-module in its own right.

Quotients M/N of a module with respect to a submodule N are defined as in vector spaces.

The following definition is an analogue of that of linear maps between vector spaces.

1.10.8. Definition. Let R be a unital ring, and M_1, M_2 be two R-modules.

A map $f : M_1 \to M_2$ is a *morphism of modules* if it is a group morphism, and then

$$(mr)f = (mf)r$$

for $m \in M_1$, $r \in R$.

As in the case of vector spaces, it follows that if $f : M_1 \to M_2$ is a morphism of modules, then the kernel $\ker(f)$ is a submodule of M_1 and the image $M_1 f$ is a submodule of M_2.

1.11. Integrality and algebraic integers

Let R be a commutative ring of characteristic zero, so that $\mathbf{Z} \subseteq R$.

1.11.1. Proposition. *For $x \in R$, the following are equivalent:*

(1) there exists $n \geq 1$ and $a_1, \dots, a_n \in \mathbf{Z}$ such that

$$x^n + a_1 x^{n-1} + \cdots + a_n = 0.$$

(2) The subring

$$\mathbf{Z}[x] = \left\{ a_0 + a_1 x + \cdots + a_k x^k : k \in \mathbf{N}, a_i \in \mathbf{Z} \right\}$$

of R is finitely generated as an abelian group, that is, as a $\mathbf{Z}$-module.

(3) The subring $\mathbf{Z}[x]$ of R is contained *in a finitely generated $\mathbf{Z}$-submodule of R.*

Proof. If (1) holds, then $\mathbf{Z}[x]$ is generated as a $\mathbf{Z}$-module by $1, x, \dots, x^{n-1}$, so that (2) holds.

Conversely, assume (2) holds. Since $\mathbf{Z}[x]$ is a finitely generated module over the Noetherian ring $\mathbf{Z}$, we have that $\mathbf{Z}[x]$ is Noetherian as a $\mathbf{Z}$-module. (See Lemma 1.11.3 below.) Considering the ascending chain of $\mathbf{Z}$-submodules of $\mathbf{Z}[x]$ generated by $1, x, \dots, x^i$, we see that there must be an m such that $\mathbf{Z}[x]$ is generated by $1, x, \dots, x^{n-1}$ as a $\mathbf{Z}$-module. Hence x^n is a $\mathbf{Z}$-linear combination of $1, x, \dots, x^{n-1}$, so that (1) holds.

Clearly (2) implies (3). We are thus left with proving that (3) implies (2).

This follows from the general fact that a submodule of a finitely generated $\mathbf{Z}$-module (more generally, a module over a PID) is finitely generated itself, see Section 1.13 below, but also [**Jac85**, Chapter 3] for a more general picture. □

1.11.2. Remark. See also [**Mar18**, Theorem 2, p. 11] for other slick proofs.

1.11.3. Lemma. *Let A be a commutative, unital Noetherian ring.*
Then each finitely generated A-module is Noetherian.

Proof. By the correspondence theorem, it is enough to prove this for the free A-module $M = A^n = A \oplus \cdots \oplus A$. We proceed by induction on n. The case $n = 1$ being clear, let $n \geq 2$.

Let $L_1 \subseteq L_2 \subseteq \dots$ be an ascending chain of A-submodules of M. Write $M = A \oplus K$, where $K = A^{n-1}$. Since $M/K \cong A$ is Noetherian, there is m such that $L_i + K = L_m + K$ for all $i \geq m$. Thus for $i \geq m$ we have, using Dedekind's identity,

$$L_i = L_i \cap (L_m + K) = L_m + (L_i \cap K).$$

Now the $L_i \cap K$ are submodules of $K = A^{n-1}$. Thus there is $n \geq m$ such that for $i \geq n$ we have $L_i \cap K = L_n \cap K$, so that for $i \geq n$

$$L_n \subseteq L_i = L_m + (L_n \cap K) \subseteq L_n.$$

□

The following result is a weak form of distributivity of intersection over sum, and has a group version as well

1.11.4. Lemma (Dedekind's Identity). *Let A, B, C be submodules of a module M, with $A \supseteq B$. Then*

$$A \cap (B + C) = B + (A \cap C).$$

Proof. If $x \in A \cap (B + C)$, then $A \ni x = b + c$ for some $b \in B$ and $c \in C$. Then $c = x - b \in A + B = A$, as $A \supseteq B$, so that $c \in A \cap C$.

Conversely, it is clear that $B + (A \cap C) \subseteq A$ and $B + (A \cap C) \subseteq B + C$. □

1.11.5. Exercise. *Show that if A, B, C are submodules of a module M, the identity*

$$A \cap (B + C) = (A \cap B) + (A \cap C)$$

does not hold in general. (Hint: Take M to be a vector space of dimension 2 over your favourite field, and A, B, C be three distinct subspaces of dimension 1.)

1.11.6. Definition. An element x which satisfies the properties of Proposition 1.11.1 is said to be *integral.* If $R = \mathbf{C}$, one speaks of an *algebraic integer*

1.11.7. Proposition. *The set of the integral elements in R is a subring of R.*

Proof. Let x, y be integral, with

$$x^n + a_1 x^{n-1} + \dots + a_n = 0 = y^m + b_1 y^{m-1} + \dots + b_m,$$

for some $a_i, b_i \in \mathbf{Z}$.

Consider the subring $\mathbf{Z}[x, y]$ of R. Each element of $\mathbf{Z}[x, y]$ can be written as

$$\alpha = \sum_{i=0}^{k} f_i(x) y^i,$$

for some k, and some polynomials $f_i \in \mathbf{Z}[y]$.

Consider the polynomial ring $\mathbf{Z}[x][z]$ with coefficients in $\mathbf{Z}[x]$. Dividing the polynomial

$$\sum_{i=0}^{k} f_i(x) z^i,$$

by the (monic) polynomial

$$z^m + b_1 z^{m-1} + \dots + b_m,$$

and evaluating at y, we obtain that

$$\alpha = \sum_{i=0}^{m-1} f_i'(x) y^i$$

for some

$$f_i'(x) = \sum_{j=0}^{n-1} c_{ij} x^j$$

for some $c_{ij} \in \mathbf{Z}$. It follows that $\mathbf{Z}[x, y]$ is generated as a $\mathbf{Z}$-module by the nm elements $x^j y^i$. Since $x + y, xy \in \mathbf{Z}[x, y]$, we are done. □

1.11.8. Remark. The proof could be made slicker using tensor products, which I may have to use anyway for the product of characters.

1.11.9. Proposition. *A rational number is an algebraic integer if and only if it is an integer.*

This result will turn out to be very useful to prove that $b \mid a$, for given integers a and $b \neq 0$. “Just” prove that the rational number a/b is an algebraic integer!

PROOF. If p/q, with p, q coprime integers, is a root of

$$x^n + a_1 x^{n-1} + \dots + a_n = 0,$$

for $a_i \in \mathbf{Z}$, then

$$p^n + a_1 p^{n-1} q + \dots + a_n q^n = 0,$$

so that $q \mid p$, and thus $q = \pm 1$. □

Clearly roots of unity (i.e., roots of $x^k - 1$ for some k) are algebraic integers.

1.11.10. LEMMA. *Let E be a subfield of $\mathbf{C}$, such that $E/\mathbf{Q}$ is a Galois extension. Suppose that $\operatorname{Gal}(E/\mathbf{Q})$ is abelian.*

Then for $\alpha \in E$ and $g \in \operatorname{Gal}(E/\mathbf{Q})$ we have $\lvert\, \alpha^g \,\rvert = \lvert\, \alpha \,\rvert^g$.

PROOF. Let c be the restriction of the complex conjugate to E. Since $\operatorname{Gal}(E/\mathbf{Q})$ is abelian, we have $\alpha^{gc} = \alpha^{cg}$, so that

$$\lvert\, \alpha^g \,\rvert^2 = \alpha^g \overline{\alpha^g} = \alpha^g \alpha^{gc} = \alpha^g \alpha^{cg} = \alpha^g \overline{\alpha}^g = (\alpha\overline{\alpha})^g = (\lvert\, \alpha \,\rvert^2)^g.$$

□

1.11.11. REMARK. The Lemma does not hold true anymore if $\operatorname{Gal}(E/Q)$ is non-abelian.

Let $E/\mathbf{Q}$ be the splitting field of $f = x^3 - 2$. Since f is irreducible over $\mathbf{Q}$ (by Eisenstein, say), we have that $3 \mid \lvert\, E : \mathbf{Q} \,\rvert$. Now the roots of f are $\alpha, \alpha\omega, \alpha\omega^2$, with $\omega \in \mathbf{C} \setminus \mathbf{R}$ a primitive third root of unity, thus $Q(\alpha) \subsetneq E$, and thus $\lvert\, E : \mathbf{Q} \,\rvert = 6$. Since an element of $G = \operatorname{Gal}(E/\mathbf{Q})$ is determined by the permutation of the roots of f it induces, G is isomorphic to S_3.

Now conjugacy induces the 2-cycle $c = (\alpha\omega, \alpha\omega^2) \in G$. Consider the 3-cycle $g = (\alpha, \alpha\omega, \alpha\omega^2) \in G$. Then

$$\alpha^{gc} = (\alpha\omega)^c = \alpha\omega^2,$$

while

$$\alpha^{cg} = \alpha^g = \alpha\omega.$$

It follows that

$$\lvert\, \alpha^g \,\rvert^2 = \alpha^g \overline{\alpha^g} = \alpha\omega\alpha\omega^2 = \alpha^2 \neq \alpha^2\omega^2 = (\alpha^2)^g = (\alpha\overline{\alpha})^g = (\lvert\, \alpha \,\rvert^2)^g.$$

1.11.12. LEMMA ([**Ser16**, Lemma 8.6]). *Let $z_1, \dots, z_n \in \mathbf{C}$ have all absolute value 1.*

If $\lvert\, z_1 + \dots + z_n \,\rvert = n$, then $z_1 = \dots = z_n$.

PROOF. Let $z_j = e^{i\varphi_j}$, for $\varphi_j \in \mathbf{R}/2\pi\mathbf{Z}$, and z be the sum of the z_j. Then

$$z\overline{z} = \sum_{j,k} e^{i(\varphi_j - \varphi_k)} = n + 2\sum_{j<k} \cos(\varphi_j - \varphi_k),$$

as for $j < k$, setting $\varphi = \varphi_j - \varphi_k$, we have

$$\begin{aligned} e^{i(\varphi_j - \varphi_k)} + e^{-i(\varphi_k - \varphi_j)} &= e^{i\varphi} + e^{-i\varphi} \\ &= \cos(\varphi) + i\sin(\varphi) + \cos(-\varphi) + i\sin(-\varphi) \\ &= \cos(\varphi) + i\sin(\varphi) + \cos(\varphi) - i\sin(\varphi) \\ &= 2\cos(\varphi_j - \varphi_k). \end{aligned}$$

Note first that we cannot have $\cos(\varphi_j - \varphi_k) = -1$ for some j, k, as this would mean $\varphi_j = \varphi_k + \pi$, so that $z_j + z_k = e^{i\varphi_j} + e^{i\varphi_j} e^{i\pi} = e^{i\varphi_j} - e^{i\varphi_j} = 0$. Clearly this implies $|\, z \,| \le n - 2$.

If the φ_j are not all equal, then one of the cosines is different from 1, and then less then 1 in absolute value, so that

$$n^2 = z\overline{z} = \left| n + 2 \sum_{j<k} \cos(\varphi_j - \varphi_k) \right| \le n + 2 \sum_{j<k} |\cos(\varphi_j - \varphi_k)\,| < n + 2\binom{n}{2} = n^2,$$

a contradiction. □

Alternative Proof, same notation. By the triangle inequality, if the absolute value of the sum of all the z_j is n, then the absolute value of the sum of any m distinct z_j must be m.

In particular, for every $j \ne k$ we must have that

$$z_j + z_k = e^{i\varphi_j} + e^{i\varphi_k} = e^{i\varphi_k}(1 + e^{i(\varphi_j - \varphi_k)}) = e^{i\varphi_k}((1 + \cos(\varphi)) + i\sin(\varphi)),$$

where $\varphi = \varphi_j - \varphi_k$, has absolute value 2, that is,

$$4 = (1 + \cos(\varphi))^2 + \sin(\varphi)^2 = 1 + \cos(\varphi)^2 + \sin(\varphi)^2 + 2\cos(\varphi) = 2(1 + \cos(\varphi)),$$

which yields $\cos(\varphi) = 1$, that is, $\varphi = \varphi_j - \varphi_k = 0$, as required. □

1.11.13. Lemma. *Let $\omega_1, \dots, \omega_n \in \mathbf{C}$ be roots of unity.*
If

$$\alpha = \frac{\omega_1 + \cdots + \omega_n}{n} \tag{1.11.1}$$

is an algebraic integer, then either $\alpha = 0$, or $\omega_1 = \cdots = \omega_n = \alpha$.

Proof. Let the ω_i be all k-th roots of unity, say. Let E be the splitting field over $\mathbf{Q}$ of $x^k - 1$. We have that $E/\mathbf{Q}$ is a Galois extension, with abelian Galois group $\mathrm{Gal}(E/\mathbf{Q})$, and $\alpha \in E$. An element $g \in \mathrm{Gal}(E/\mathbf{Q})$ maps roots of unity to roots of unity, so α^g has the same form.

Note that

$$|\, \alpha \,| = \frac{|\, \omega_1 + \cdots + \omega_n \,|}{n} \le \frac{|\, \omega_1 \,| + \cdots + |\, \omega_n \,|}{n} = 1,$$

and thus the same holds for the conjugates α^g.

Now the norm (in the sense of field theory)

$$\prod_{g \in \mathrm{Gal}(E/\mathbf{Q})} \alpha^g \tag{1.11.2}$$

of α is fixed by $\mathrm{Gal}(E/\mathbf{Q})$, and thus is in $\mathbf{Q}$. It is an algebraic integer, as a product of algebraic integers, and thus it is an integer. Each term of (1.11.2) is of the same form as α, and thus has absolute value at most 1. It follows that the absolute value of (1.11.2) is at most 1.

Now there are two possibilities.

(1) (1.11.2) might be zero, so that one factor is zero, and thus they are all zero, so that $\alpha = 0$.

(2) (1.11.2) is ± 1, and thus each factor, which has absolute value at most 1, must have absolute value 1. In particular

$$|\,\omega_1 + \dots + \omega_n\,| = n.$$

But this can hold only if all ω_i are equal, as per the previous Lemma.

□

1.12. Another approach to algebraic integers

This is in development, and disconnected from the above at the moment

A complex number is said to be an *algebraic integer* if it is the root of a monic polynomial with integer coefficients. Examples: integers, i, $\sqrt{2}$, $\sqrt{2}+\sqrt{3}$.

Therefore every algebraic integer is algebraic (over the rationals), but not vice versa (Proposition 1.11.9).

The following treatment is taken from [**Isa06**].

Let us start by noting that if S is a subring of $\mathbf{C}$ which contains $\mathbf{Z}$, given $z_i \in \mathbf{Z}$ e y_i in S, for $i = 1, \dots, n$, we may consider the linear combinations (with integer coefficients)

$$z_1 y_1 + \dots z_i y_i \in S.$$

This does not yield a vector space, as $\mathbf{Z}$ is not a field, but what is called a $\mathbf{Z}$-module.

1.12.1. Lemma. *Let $X = \{\,\alpha_1, \dots, \alpha_n\,\}$ be a finite set of algebraic integers. Then there are*

(1) a subring S of $\mathbf{C}$, containing $\mathbf{Z}$ and X, and
(2) a finite subset $Y \subseteq S$

such that every element of S can be written as a linear combination with integer coefficients of the elements of Y.

Proof. Each α_i will be a root of a monic polynomial in $\mathbf{Z}[x]$ degree n_i. As in the case of algebraic elements, $\alpha_i^{n_i}$ can be written as a linear combination with integer coefficients of $1, \alpha_i, \dots, \alpha_i^{n_i - 1}$.

Let

$$Y = \left\{\, \alpha_1^{k_1} \cdots \alpha_n^{k_n} : 0 \le k_i < n_i \,\right\},$$

and let S be the set of all linear combinations, with integer coefficients, of the elements of Y. For each j, we have two possibilities

(1) either $k_j < n_i - 1$, so that

$$\alpha_j \cdot (\alpha_1^{k_1} \cdots \alpha_n^{k_n}) \in Y,$$

(2) or $k_j = n_i - 1$, and then

$$\alpha_j \cdot (\alpha_1^{k_1} \cdots \alpha_n^{k_n}) = \alpha_1^{k_1} \cdots \alpha_j^{k_j} \cdots \alpha_n^{k_n}$$

can be written as a linear combination with integer coefficients of the elements of Y

$$\alpha_1^{k_1} \cdots \alpha_j^{t} \cdots \alpha_n^{k_n},$$

per $0 \leq t < n_j$.

Therefore S is closed under multiplication by all α_j, and thus under multiplication by all elements of Y. The distributive property implies that S is a subring of $\mathbf{C}$. □

For the converse, we have

1.12.2. Theorem. *Let S be a subring of $\mathbf{C}$ containing $\mathbf{Z}$. Let $Y \subseteq S$ be a finite set, such that each element of S can be written as a linear combination with integer coefficients of the elements of Y.*

Then every element of S is an algebraic integer.

Proof. Let $Y = \{ y_1, \dots, y_n \}$. For each i, and each $s \in S$, we have

$$y_i s = \sum_{j=1}^{n} a_{ij} y_j,$$

for suitable integers a_{ij}. But then s is an eigenvalue of the matrix $A = [a_{ij}]$, and this a root of its characteristic polynomial

$$\det(x\mathbf{1} - A),$$

which is a monic polynomial in $\mathbf{Z}[x]$. □

1.12.3. Corollary. *Sums and products of algebraic integers are algebraic integers.*

Proof. Let α, β be algebraic integers. By Lemma 1.12.1, there is a subring S of $\mathbf{C}$, containing $\mathbf{Z}$ and α, β, such that every element of S can be written as a linear combination with integer coefficients of the elements of a suitable finite subset Y of S.

As $\alpha+\beta, \alpha\cdot\beta \in S$, by Theorem 1.12.2 these elements are algebraic integers. □

1.13. Finitely generated Z-modules

Of course $\mathbf{Z}$-modules are the same thing as abelian groups.

We want to prove

1.13.1. Theorem. *Let A be a $\mathbf{Z}$-module, which is generated by n elements. Let B be a submodule of A.*

Then B can be generated by $\leq n$ elements.

If A is free, then so is B.

Proof. We first reduce to the free case. A free $\mathbf{Z}$-module of rank n is a $\mathbf{Z}$-module that has a basis, so that, very much as in the case of vector spaces, it is isomorphic to $\mathbf{Z}^n$.

Since A is n-generated, there is a surjective morphism of $\mathbf{Z}$-modules $f : \mathbf{Z}^n \to A$. Then $f^{-1}(B)$ is a submodule of $\mathbf{Z}^n$, and $f\mid_{f^{-1}(B)}: f^{-1}(B) \to B$ is a surjective morphism.

It is therefore enough to show that a submodule C of $\mathbf{Z}^n$ is free on $k \leq n$ generators.

Proceeding by induction on n, the basis is provided by the fact that every subgroup (i.e. $\mathbf{Z}$-submodule) of $\mathbf{Z}$ is of the form $n\mathbf{Z}$, which is either zero, or isomorphic to $\mathbf{Z}$. (And it turns out that this proof holds more generally for modules over any PID.)

So suppose $n \geq 2$, and consider the first projection

$$\begin{aligned}\pi_1 : \mathbf{Z}^n &\to \mathbf{Z}\\ (z_1, z_2, \dots, z_n) &\mapsto z_1,\end{aligned}$$

with kernel

$$K = \{\, 0 \,\} \times \mathbf{Z} \times \cdots \times \mathbf{Z} \cong \mathbf{Z}^{n-1}.$$

The induction hypothesis yields that $C \cap K$ is free of some rank $h \leq n - 1$. If $\pi_1(C) = \{\, 0 \,\}$, that is, $C \leq K$, we have $C = C \cap K$, and we are finished. If $\pi_1(C) \neq \{\, 0 \,\}$, then $\pi_1(C) = \langle\, z \,\rangle$, for some $z \neq 0$. Let $c \in C$ be such that $\pi_1(c) = z$. We claim that

$$C = \langle\, c \,\rangle \oplus (C \cap K),$$

which will show that C is free of rank $h \leq n$.

If $x \in C$, then $\pi_1(x) = \lambda z$ for some $\lambda \in \mathbf{Z}$, so that $\pi_1(x - \lambda c) = 0$, and $x - \lambda c \in C \cap K$. This shows that $C = \langle\, c \,\rangle + (C \cap K)$.

If $x \in \langle\, c \,\rangle \cap (C \cap K)$, then $x = \lambda c$ for some $\lambda \in \mathbf{Z}$, and $0 = \pi_1(x) = \lambda z$. Since $z \neq 0$ we get $\lambda = 0$, so that $c = 0$. □

1.14. Tensor Products

This is basically taken from [**Lan02**].

1.14.1. The commutative case. Let us start with a *commutative* ring R. Let $M_1, \dots, M_n, N$ be R-modules. (Since R is commutative, left and right modules are the same.)

A map

$$f : M_1 \times \cdots \times M_n \to N$$

is said to be *multilinear* if for each i and $m_1, \dots m_{i-1}, m_{i+1}, \dots, m_m$ one has that the map

$$\begin{aligned}f_i : M_i &\to N\\ x &\mapsto f(m_1, \dots, m_{i-1}, x, m_{i+1}, \dots, m_m)\end{aligned}$$

is a morphism of R-modules. A tensor product of the M_i is a construction that allows one to replace the whole of multilinear maps with just one of them, and then linear maps, that is, plain morphisms of modules.

1.14.1. Definition. Let $M_1, \dots, M_n, N$ be R-modules. A tensor product of the M_i is an R-module

$$M_1 \otimes \cdots \otimes M_n.$$

together with a multilinear map

$$\iota : M_1 \times \cdots \times M_n \to M_1 \otimes \cdots \otimes M_n,$$

such that for every R-module N and every multilinear map

$$f : M_1 \times \cdots \times M_n \to N$$

there is a unique morphism of R-modules

$$g : M_1 \otimes \cdots \otimes M_n \to N$$

that makes the diagram (1.14.1) commute.

$$\begin{tikzcd} M_1 \times \cdots \times M_n \arrow[r, "f"] \arrow[d, "\iota"] & N \\ M_1 \otimes \cdots \otimes M_n \arrow[ur, dotted, "g"'] & \end{tikzcd} \tag{1.14.1}$$

One writes

$$\iota(m_1, \dots, m_n) = m_1 \otimes \cdots \otimes m_n.$$

Familiar arguments show that a tensor product, if it exists, is unique up to an isomorphism of modules.

Existence can be proved as follows. Start with the free module M which has as basis the elements of

$$M_1 \times \cdots \times M_n.$$

Consider the submodule K of M generated by the following elements

$$\begin{aligned}(m_1, \dots, m_{i-1}, x + y, m_{i+1}, \dots, m_m)\\ - (m_1, \dots, m_{i-1}, x, m_{i+1}, \dots, m_m) - (m_1, \dots, m_{i-1}, y, m_{i+1}, \dots, m_m)\end{aligned}$$

and

$$(m_1, \dots, m_{i-1}, rx, m_{i+1}, \dots, m_m) - r(m_1, \dots, m_{i-1}, x, m_{i+1}, \dots, m_m)$$

for all i, $m_j \in M_j$ and $x, y \in M_i$, and $r \in R$. We claim the quotient module M/K is a tensor product of the M_i. In fact, as $M_1 \times \cdots \times M_n$ is a subset of M, we can compose the inclusion map with the projection $M \to M/K$ to get

$$\iota : M_1 \times \cdots \times M_n \to M/K.$$

This is a multilinear map, by the very definition of K. It is also a tensor product. In fact, if

$$f : M_1 \times \cdots \times M_n \to N$$

is a multilinear map, this can be extended linearly to a morphism of modules $f' : M \to N$. The generators of K are visibly in $\ker(f')$, so that we obtain a morphism of modules $g : M/K \to N$, which satisfies by construction

$$g(\iota(m_1, \dots, m_n) = g((m_1, \dots, m_n) + K) = f'(m_1, \dots, m_n) = f(m_1, \dots, m_n),$$

so that the diagram commutes.

Write

$$\iota(m_1, \dots, m_m) = m_1 \otimes \cdots \otimes m_m \in M_1 \otimes \cdots \otimes M_n.$$

We will see that in general $M_1 \otimes \cdots \otimes M_n$ is not the set of the $m_1 \otimes \cdots \otimes m_m$, but by construction these elements do generate it,

Again by construction we have, for $a \in R$,

$$a(m_1 \otimes \cdots \otimes m_m) = (am_1) \otimes \cdots \otimes m_m = \cdots = m_1 \otimes \cdots \otimes (am_m).$$

The tensor product of two non-trivial modules can well be trivial. For instance

$$\mathbf{Z}/2\mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Z}/3\mathbf{Z} = \{ 0 \}$$

where $\otimes_{\mathbf{Z}}$ means that this is a tensor product of $\mathbf{Z}$-modules. In fact, for $a, b \in \mathbf{Z}$ we have

$$[a]_2 \otimes [b]_3 = (3[a]_2) \otimes [b]_3 = [a]_2 \otimes (3[b]_3) = [a]_2 \otimes [0]_3 = 0.$$

1.14.2. Exercise.

(1) Show that if $\gcd(m, n) = 1$*, then*

$$\mathbf{Z}/m\mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Z}/n\mathbf{Z} = \{ 0 \}.$$

(2) Show that

$$\mathbf{Z}/m\mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Z}/n\mathbf{Z} \cong \mathbf{Z}/\gcd(m, n)\mathbf{Z}.$$

(Hint: Show that $1 \otimes 1$ generates $\mathbf{Z}/m\mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Z}/n\mathbf{Z}$. Therefore the morphism $\mathbf{Z} \to \mathbf{Z}/m\mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Z}/n\mathbf{Z}$ which maps $x \mapsto k([1]_m \otimes 1_n)$ is surjective. What is its kernel? Certainly the kernel is contained in $\gcd(m, n)\mathbf{Z}$, because if $[k]_m \times [1]_n = 0 = [1]_m \otimes [k]_n$, then $m \mid k$ and $n \mid k$, so that $\gcd(m, n) \mid k$. Now the map $\mathbf{Z}/m\mathbf{Z} \times \mathbf{Z}/n\mathbf{Z} \to \mathbf{Z}/\gcd(m, n)\mathbf{Z}$ given by $([a]_m, [b]_n) \mapsto [an]_{\gcd(m,n)}$ is well-defined, bilinear and surjective, so it induces an epimorphism $\mathbf{Z}/m\mathbf{Z} \otimes_{\mathbf{Z}} \mathbf{Z}/n\mathbf{Z} \to \mathbf{Z}/\gcd(m, n)\mathbf{Z}$. Put the two things together.)

1.14.2. The case of vector spaces. The case of vector spaces is much more manageable. Let V, W be vector spaces of dimension n, m over a field F, and bases $v_1, \dots, v_n$ and $w_1, \dots, w_m$. We first give an alternative construction of the tensor product $V \otimes_F W$. Consider the set $\operatorname{Bil}(V, W; F)$ of bilinear maps $V \times W \to F$. This is a vector space over F, with the usual operations on the images. By definition of the tensor product, there is a bijection between $\operatorname{Bil}(V, W; F)$ and the dual $(V \otimes W)^*$. So let us define $V \otimes W = \operatorname{Bil}(V, W; F)^*$, and let us show that this has indeed the properties of a tensor product.

First note that a basis of $\operatorname{Bil}(V, W; F)$ is given by the elements E_{ij}

$$(v_s, w_t)E_{ij} = \begin{cases} 1 & \text{if } s = i,\ t = j, \\ 0 & \text{otherwise.} \end{cases}$$

Consider the dual basis $v_i \otimes w_j = E_{ij}^*$ of $\operatorname{Bil}(V, W; F)^*$. The map defined by

$$\begin{aligned} \iota : V \times W &\to \operatorname{Bil}(V, W; F)^* \\ (\sum a_i v_i, \sum b_j w_j) &\mapsto \sum a_i b_j (v_i \otimes w_j) \end{aligned}$$

is bilinear, and if $f : V \times W \to U$ is a bilinear map, for some vector space U, then the linear map $g : V \otimes W \to U$ defined by

$$(v_i \otimes w_j)g = (v_i, w_j)f$$

satisfies indeed

$$(\sum a_i v_i, \sum b_j w_j)\iota g = (\sum a_i b_j (v_i \otimes w_j))g = \sum a_i b_j (v_i, w_j)f = (\sum a_i v_i, \sum b_j w_j)f.$$

We have obtained

1.14.3. Proposition. *Let V, W be vector spaces of dimension n, m over a field F, and bases $v_1, \dots, v_n$ and $w_1, \dots, w_m$.*

Then $\dim(V \otimes_F W) = nm$, and a basis of $V \otimes W$ is given by the $\iota(v_i, w_j) = v_i \otimes u_j$.

Let now $X \in \mathrm{End}_F(V)$, $Y \in \mathrm{End}_F(W)$. Then the map

$$\begin{aligned} V \times W &\to V \otimes W \\ (v, w) &\mapsto (vX) \otimes (wY) \end{aligned}$$

is readily seen to be bilinear. It follows that it induces an element of $\mathrm{End}(V \otimes W)$ given by

$$X \otimes Y : v \otimes w \mapsto (vX) \otimes (wY). \tag{1.14.2}$$

This map is called the tensor, or Kronecker, product of X and Y. If X, Y are given in terms of matrices, then one checks that the elements of $X \times Y$, which is a $(nm) \times (nm)$ matrix, are the $x_{ij}y_{kl}$. By choosing an appropriate ordering of the basis of $V \otimes W$, $X \otimes Y$ can be written as a block matrix

$$\begin{bmatrix} x_{11}Y & x_{12}Y & \dots & x_{1n}Y \\ x_{21}Y & x_{22}Y & \dots & x_{2n}Y \\ & & \ddots & \\ x_{n1}Y & x_{n2}Y & \dots & x_{nn}Y \end{bmatrix},$$

or also (with respect to a different, appropriate ordering of the basis), as

$$\begin{bmatrix} y_{11}X & y_{12}X & \dots & y_{1m}X \\ y_{21}X & y_{22}X & \dots & y_{2m}X \\ & & \ddots & \\ y_{m1}X & y_{m2}X & \dots & y_{mn}X \end{bmatrix}.$$

From either form one gets that

$$\begin{aligned} \mathrm{trace}(X \otimes Y) &= \mathrm{trace}(x_{11}Y) + \dots + \mathrm{trace}(x_{nn}Y) = \\ &= x_{11}\,\mathrm{trace}(Y) + \dots + x_{nn}\,\mathrm{trace}(Y) = \mathrm{trace}(X)\,\mathrm{trace}(Y). \end{aligned} \tag{1.14.3}$$

From the definition (1.14.2) we obtain immediately the following property, which will be important in what follows.

$$(X_1 \otimes X_2)(Y_1 \otimes Y_2) = (X_1 Y_1) \otimes (X_2 Y_2). \tag{1.14.4}$$

1.14.3. The non-commutative case. When R is a non-commutative ring, one can do the tensor product $M_R \otimes_R {}_RN$ of a right module $M = M_R$ and a left module $N = {}_RN$ (the indices serve to remember which is which). This is defined so that its generators satisfy

$$xr \otimes y = x \otimes ry,$$

for $x \in M$, $y \in N$, $r \in R$. Note that the equalities, for $r, s \in R$,

$$x(rs) \otimes y = (xr)s \otimes y = xr \otimes sy = x \otimes r(sy) = x \otimes (rs)y$$

show why one has to take one right and one left module. Note also that $M_R \otimes_R {}_R N$ is only an abelian group (but more about this in a second), unless R is commutative, so that right and left modules are one and the same thing. In fact if one tries and define

$$r(x \otimes y) = xr \otimes y = x \otimes ry$$

one sees that this makes $M_R \otimes_R {}_R N$ into both a right and left module.

This will have an important application to induced representations, where we will have a situation like this. Let G be a finite group, $H \leq G$, and V a (right) $\mathbf{C}[H]$ module. Consider the tensor product

$$V \otimes_{\mathbf{C}[H]} \mathbf{C}[G]. \tag{1.14.5}$$

Here $\mathbf{C}[G]$ is a *left* $\mathbf{C}[H]$-module in a natural way. Since $\mathbf{C}[G]$ is also naturally a *right* $\mathbf{C}[G]$-module, we can make (1.14.5) into a *right* $\mathbf{C}[G]$*-module* via

$$(v \otimes \mathfrak{r})\mathfrak{s} = v \otimes (\mathfrak{r}\mathfrak{s}),$$

for $\mathfrak{r}, \mathfrak{s} \in \mathbf{C}[G]$.

Actually, if R, S, T are rings, ${}_R M_S$ is a left R-module and a right S-module (such a beast is called a *bimodule*) and ${}_S M_T$ is a left S-module and a right T-module, then

$${}_R M_S \otimes_S {}_S M_T$$

becomes a left R-module and a right T-module.

Part 2

Soluble and nilpotent groups

CHAPTER 2

Series

For further details of the arguments of this section, see [**Rob96**].

In this section G will be a finite group.

2.1. Series

A *series* in G is a sequence of *distinct* subgroups

$$1 = H_0 \trianglelefteq H_1 \trianglelefteq \dots \trianglelefteq H_n = G, \tag{2.1.1}$$

each normal in the next.

One says that the series (2.1.1) is *normal* if each $H_i \trianglelefteq G$.

2.1.1. Exercise. *Show that the sequence*

$$1 < \langle\, (12)(34)\, \rangle < \langle\, (12)(34), (13)(24)\, \rangle < S_4$$

is a series which is not normal.

2.2. Ω-groups and Ω-series

The concept of an Ω-series is sometimes useful.

2.2.1. Definition. Let G be a group, Ω a set, and

$$\alpha : G \times \Omega \to G$$

a function.

A *right operator group* is a triple (G, Ω, α), such that for each $\omega \in \Omega$ the map

$$g \mapsto (g, \omega)\alpha$$

is an endomorphism of G.

One says that G is an Ω-group, and writes simply g^{ω} for $(g, \omega)\alpha$.

An Ω-subgroup is a subgroup $H \le G$ such that $h^{\omega} \in H$ for each $h \in H$ and $\omega \in \Omega$.

(1) If $\Omega = \emptyset$, we get simply a group, and the Ω-subgroups of G are just the subgroups of G.
(2) If $\Omega = \mathrm{Inn}(G)$, then the Ω-subgroups are the subgroups that are normal in G.
(3) If $\Omega = \mathrm{Aut}(G)$, then the Ω-subgroups are the so-called *characteristic* subgroups.
(4) If $\Omega = \mathrm{End}(G)$, then the Ω-subgroups are the so-called *fully invariant* subgroups.

2.2.2. Exercise. *Find examples of groups G and subgroups $H \le G$ such that*

(1) H is normal, but not characteristic in G.
(2) H is characteristic, but not fully invariant in G.

(Hint: For the first question, one can consider an elementary abelian group of order p^2, where p is a prime, that is, a group of the form $C_p \times C_p$, where C_p, where C_p is cyclic of order p.

For the second question, consider the group A_5, which is known to be simple, and a group $B = \langle\, b \,\rangle$ of order 2, say, and the product $G = A_5 \times B$. Then $B = Z(G)$ is characteristic in G. But the endomorphism φ, which has kernel A_5 and maps $b \mapsto (12)(34)$, does not map B into B.

Alternatively, take G to be the dihedral group of order 8. Show that it has a unique cyclic subgroup C of order 4, which is thus characteristic. Show that C is not fully invariant.)

An Ω-series is a series where each term is an Ω-subgroup.

It follows that a series is *normal* if and only if it is an $\mathrm{Inn}(G)$-series.

2.3. Ω-composition series

A series is a refinement of another if it can be obtained by inserting further subgroups. For instance

$$1 < \langle\, (12)(34) \,\rangle < \langle\, (12)(34), (13)(24) \,\rangle < S_4$$

refines

$$1 < \langle\, (12)(34), (13)(24) \,\rangle < S_4.$$

An Ω-series which has no proper refinement is called a *Ω-composition series.* If $\Omega = \emptyset$ one speaks simply of a *composition series.* If $\Omega = \mathrm{Inn}(G)$, one speaks of a *principal series.*

It follows from the third isomorphism theorem that the factors H_{i+1}/H_i of a composition series are simple groups. One sees that the factors of a principal series are *characteristically simple*, that is, they have no proper, non-trivial characteristic subgroups.

2.3.1. Proposition. *Let $G \neq 1$ be a finite group which is characteristically simple.*

Then there is a simple group S (abelian or non-abelian) and a positive integer such that

$$G \cong S^n = S \times \cdots \times S.$$

Proof. Let us first consider the special case when G is abelian. Let p be a prime dividing its order. Then a Sylow p-subgroups is normal, and thus characteristic in G, so that G is a p-group. The subgroup

$$\{\, g \in G : g^p = 1 \,\}$$

is characteristic in G, so that G is elementary abelian, so $G = S^n$ where S is cyclic of order p.

In the general case, let N be a minimal normal subgroup of G. Then for each automorphism φ of G such that $N \neq N^\varphi$ one has $N \cap N^\varphi = 1$, so that

$[N, N^{\varphi}] \le N \cap N^{\varphi} = 1$. (Note that $N^{\varphi} \trianglelefteq G$, as $N^{\varphi\iota(g)} = N^{\iota(g^{\varphi^{-1}})\varphi} = N^{\varphi}$, see the proof of Exercise 1.8.5.) Since $G = \langle\, N^{\varphi} : \varphi \in \operatorname{Aut}(G)\,\rangle$, one sees (argument below) that G is a direct product of some of the N^{φ}, including N. Thus if $1 \ne K \trianglelefteq N$, then $K \trianglelefteq G$, so $K = N$ and N is simple.

To see that G is a direct product of some of the N^{φ}, start with $M = N$. If $M = G$, we are done. Now let M be a direct product of some of the N^{φ}, including N, so that $M \trianglelefteq G$. If $M < G$, there is a $N^{\psi} \not\le M$. Since $M \cap N^{\psi} \ne N^{\psi}$, we have $M \cap N^{\psi} = 1$ by the minimality of N^{ψ}, so that $MN^{\psi} = M \times N^{\psi}$. □

And now for the converse.

2.3.2. Proposition. *A direct product of isomorphic simple groups is characteristically simple.*

2.3.3. Exercise. *Using Proposition 1.9.3, show that if G is an elementary abelian p-group, for a prime p, then its automorphism group acts transitively on the non-zero elements.*

(Hint: One has to show that given a finite-dimensional vector space V (over any field, actually), and two non-zero vectors $v, w \in V$, there is a linear map taking v to w.)

Proof. If the simple group S is abelian, then it has order a prime p, so that G is elementary abelian, and one can use Exercise 2.3.3.

So let S be non-abelian simple, so that $G = S^n = T_1 \times \dots T_n$ for some $n > 1$, the case $n = 1$ being trivial.

Let $L \ne 1$ be a *normal* subgroup of G, and let

$$1 \ne (s_1, s_2, \dots, s_n) \in L,$$

where we may assume $s_1 \ne 1$. Since $Z(S) = \{\,1\,\}$, there is $x \in S$ such that $s_1^x \ne s_1$. Therefore

$$(s_1, s_2, \dots, s_n)^{-1} \cdot (s_1, s_2, \dots, s_n)^{(x,1,\dots,1)} = (y, 1, \dots, 1)$$

for $y = s_1^{-1} s_1^x \ne 1$. Thus $L \cap T_1 \ne 1$. Since T_1 is minimal normal, we have $L \ge T_1$. Note that we have proved so far is that the T_i are the unique minimal normal subgroups of G. This is a special case of the Krull-Remak-Schmidt theory, see [**Rob96**, p. 80].

Now suppose L is indeed *characteristic* in G. It remains to note that there is a subgroup of $\operatorname{Aut}(G)$ isomorphic to S_n, which permutes the T_i. (See Exercise 2.3.4 just below.) Therefore $L = G$. □

2.3.4. Exercise. *Let S be a set, $n \ge 1$ an integer, $G = S^n$ be the direct product of n copies of S.*

(1) *Show that the assignment, for $\sigma \in S_n$,*

$$(s_1, \dots, s_n)^{\sigma} = (s_{1\sigma^{-1}}, \dots, s_{n\sigma^{-1}})$$

defines a right *action of S_n on G.*

(Hint: This is slightly tricky: the inverse is needed to make this into a *right* action.)

Let now S be a group, so that G is a direct product.

(2) Show that for each $\sigma \in S_n$ the map σ' given by

$$(s_1, \dots, s_n) \mapsto (s_1, \dots, s_n)^\sigma$$

defines an automorphism of the group G, and actually $\sigma \mapsto \sigma'$ is a morphism $S_n \to \operatorname{Aut}(G)$.

2.4. Uniqueness of the factors of an Ω-composition series

2.4.1. Theorem. *Let G be a finite Ω-group. Suppose G has two Ω-composition series.*

Then the factors of the two series are pairwise isomorphic.

Proof. Let

$$1 = H_0 \trianglelefteq H_1 \trianglelefteq \dots \trianglelefteq H_{n-1} \trianglelefteq H_n = G, \tag{2.4.1}$$

and

$$1 = K_0 \trianglelefteq K_1 \trianglelefteq \dots \trianglelefteq K_{m-1} \trianglelefteq K_m = G \tag{2.4.2}$$

be the two Ω-composition series, and proceed by induction on the order of G. If $H_{n-1} = K_{m-1}$, we are done by induction. If $H_{n-1} \neq K_{m-1}$, consider the Ω-subgroup $L = H_{n-1} \cap K_{m-1}$, and refine the Ω-series

$$L \trianglelefteq H_{n-1} \trianglelefteq G, \qquad L \trianglelefteq K_{m-1} \trianglelefteq G$$

to two Ω-composition series $\mathcal{H}$ and $\mathcal{K}$, by taking the same refinement of L for both. Since $H_{n-1}K_{m-1}$ is an Ω-subgroup properly containing both H_{n-1} and K_{m-1}, we have $H_{n-1}K_{m-1} = G$. Therefore

$$\begin{aligned} G/H_{n-1} = H_{n-1}K_{m-1}/H_{n-1} \cong K_{m-1}/L, \quad &\text{and} \\ &G/K_{m-1} = H_{n-1}K_{m-1}/K_{m-1} \cong H_{n-1}/L. \end{aligned} \tag{2.4.3}$$

Proceeding by induction, the factors of (2.4.1) and $\mathcal{H}$ are pairwise isomorphic; they are the factors of the Ω-series for L, plus H_{n-1}/L and G/H_{n-1}. Also, the factors of (2.4.2) and $\mathcal{K}$ are pairwise isomorphic; they are the factors of the Ω-series for L, plus K_{m-1}/L and G/K_{m-1}. By (2.4.3), we are done. □

2.4.2. Exercise. *The factors of an Ω-composition series do not determine the group uniquely. For instance both C_6 and S_3 have two composition factors which are cyclic of orders 2, 3, and both C_4 and $C_2 \times C_2$ have two composition factors which are cyclic of orders 2.*

CHAPTER 3

Soluble groups

3.1. Commutators

3.1.1. Definition. Let G be a group, $a, b \in G$. The *commutator* of a, b is

$$[a, b] = (ba)^{-1}ab = a^{-1}b^{-1}ab.$$

The name is justified by the

3.1.2. Lemma. *Let G be a group, $a, b \in G$. The following are equivalent*

(1) $ab = ba$, and
(2) $[a, b] = 1$.

3.1.3. Exercise. *Show that $[a, b]^{-1} = [b, a]$.*

3.1.4. Definition. The subgroup

$$G' = \langle\, [a, b] : a, b \in G \,\rangle$$

of G generated by the commutator is referred to as the *derived subgroup* or the *commutator subgroup*.

3.1.5. Remark. In general not all elements of G' will be commutators, but just products of commutators. But it has been proved that if G is a finite, nonabelian simple group, then every element of G' is a commutator [**LOST10**].

3.1.6. Exercise. *Show that $S_n' = A_n$ for all $n > 1$.*
(Hint:

(1) Show that

$$(132)(123\ldots k) = (145\ldots k) \qquad (12)(34) = (123)(143).$$

(2) Show that A_n is generated by the 3-cycles.
(3) Show that $[(12), (23)] = (123)$.

)

3.1.7. Exercise. *Show that $A_n' = A_n$ for $n \geq 5$.*

More generally, if $A, B \leq G$, then we define the subgroup

$$[A, B] = \langle\, [a, b] : a \in A, b \in B \,\rangle\,.$$

3.1.8. Exercise.

(1) Prove the identities, for a, b, c in a group.

$$[a, bc] = [a, c][a, b]^c, \qquad [ab, c] = [a, c]^b[b, c].$$

(2) Prove that if A, B are subgroups of a group G, then $A, B \leq N_G([A, B])$.

Note the following

3.1.9. Lemma.

- *If G, H are groups, $\varphi : G \to H$ is a morphism, and $a, b \in G$, then $\varphi([a, b]) = [\varphi(a), \varphi(b)]$.*
- *G' is a fully invariant (and thus characteristic) subgroup of G, that is, $\varphi(G') = \varphi(G)' \leq G'$ for all $\varphi \in \operatorname{End}(G)$.*
- *If $K \leq G$, then the following are equivalent*
  (1) $K \trianglelefteq G$, and
  (2) $[K, G] \leq K$.

Proof. The first statement is clear, as a morphism respects products and inverses.

The second one follows immediately.

As to the third one, just note that for $a \in K$ and $b \in G$ we have

$$a^b = b^{-1}ab = a[a, b].$$

□

Note the following

3.1.10. Lemma. *Let G be a group, $N \trianglelefteq G$.*

- *If $H \trianglelefteq N$, then H is not necessarily normal in G.*
- *If H is characteristic in N, then $H \trianglelefteq G$.*

Proof. For the first claim, take $G = A_4$, $N = \langle\, (12)(34), (13)(24) \,\rangle$ to be the 2-Sylow subgroups of G, and $H = \langle\, (12)(34) \,\rangle$.

For the second one, consider for each $g \in G$, the map

$$N \to N$$
$$n \mapsto g^{-1}ng.$$

This is well defined, as $N \trianglelefteq G$, and it is an automorphism of N. Since H is characteristic in N, we have $g^{-1}hg \in H$ for all $g \in G$ and $h \in H$, that is, $H \trianglelefteq G$. □

It is easy to see that G/G' is abelian. More generally, we have

3.1.11. Proposition. *Let G be a group, $H \leq G$. The following are equivalent:*

(1) $H \trianglelefteq G$ and G/H is abelian, and
(2) $G' \leq H$.

Thus G' is the smallest normal subgroup of G with abelian quotient.

Proof. Assuming $H \trianglelefteq G$ and G/H abelian, we have for all $a, b \in G$

$$H = [aH, bH] = [a, b]H,$$

so that $[a, b] \in H$ and thus $G' \leq H$.

Conversely if $G' \leq H$, then $[H, G] \leq G' \leq H$, so $H \trianglelefteq G$, and for $a, b \in G$ we have

$$[aH, bH] = [a, b]H = H.$$

□

3.2. The derived series and soluble groups

3.2.1. Definition. Given a group G, one constructs its *derived sequence*

$$G^0 = G, G^{(1)} = G', \dots, G^{(n+1)} = (G^n)', \dots$$

G is said to be soluble (*solvable* in American English) if there is n such that $G^{(n)} = \{ 1 \}$. (So the derived sequence is a series, according to our definitions.)

3.2.2. Remark. The name comes form Galois theory, as an equation is soluble by radicals if and only if its Galois group is soluble.

3.2.3. Proposition. *Let G be a group. The following are equivalent:*

(1) G is soluble;
(2) there is a normal series $G = G_0 \geq G_1 \geq \cdots \geq G_m = \{ 1 \}$ with G_i/G_{i+1} abelian for all i;
(3) there is a series $G = G_0 \geq G_1 \geq \cdots \geq G_m = \{ 1 \}$ with G_i/G_{i+1} abelian for all i;

Proof. If the first condition holds, then $G_i = G^{(i)}$ satisfies the second one.

If the second condition holds, then the series G_i satisfies the third one.

Let now G_i be a series as in (3). Then $G_1 \trianglelefteq G$ and G/G_1 is abelian, so that $G' = G^{(1)} \leq G_i$. Proceeding by induction, assume $G^{(i)} \leq G_i$. Since $G_{i+1} \trianglelefteq G_i$, and G_i/G_{i+1} is abelian, we have $G^{(i+1)} = (G^{(i)})' \leq G_i' \leq G_{i+1}$, so that $G^{(n)} = \{ 1 \}$. □

3.2.4. Proposition. *Let G be a finite group. Then the following are equivalent.*

(1) G is soluble,
(2) there is a composition series whose factors are of prime order,
(3) the factors of any composition series are of prime order.

Proof. If G is soluble, refine the derived series to a composition series.

If a composition series has all factors of prime order, then by Theorem 2.4.1 this holds for any composition series.

A composition series with factors of prime order satisfies (3) of Proposition 3.2.3. □

In a similar manner one proves

3.2.5. Proposition. *Let G be a finite group. Then the following are equivalent.*

(1) G is soluble,
(2) there is a principal series whose factors are elementary abelian.,
(3) the factors of any principal series are elementary abelian.

3.2.6. Theorem. *Let G be a group.*

(1) If G is soluble and $H \leq G$, then H is soluble.
(2) If G is soluble and $N \trianglelefteq G$, then G/N is soluble.
(3) If G is soluble, and $\varphi : G \to K$ is a morphism, then $\varphi(K)$ is soluble.
(4) If $N \trianglelefteq G$, and both N and G/N are soluble, then G is soluble.

Proof. If G is soluble and $H \leq G$, just note that $H^{(i)} \leq G^{(i)}$ for all i.

If G is soluble and $N \trianglelefteq G$, consider a series G_i as in Proposition 3.2.3(2). Then $G_i N \trianglelefteq G$ for all i, so that

$$\frac{G_i N}{N} \trianglelefteq \frac{G}{N},$$

and

$$\left(\frac{G_i N}{N}\right) / \left(\frac{G_{i+1} N}{N}\right) \cong \frac{G_i N}{G_{i+1} N} = \frac{G_i G_{i+1} N}{G_{i+1} N} = \frac{G_i}{G_i \cap G_{i+1} N} \cong \left(\frac{G_i}{G_{i+1}}\right) / \left(\frac{G_i \cap G_{i+1} N}{G_{i+1}}\right)$$

is abelian, as a quotient of the abelian group G_i / G_{i+1}.

(3) follows from the first isomorphism theorem.

As to (4), let X_i / N be a series with abelian quotients for G/N, and Y_i be a series with abelian quotients for G. Since

$$\frac{X_{i-1}/N}{X_i/N} \cong \frac{X_{i-1}}{X_i},$$

we have that the series obtained by starting with the X_i, and continuing with the Y_i, is a series with abelian quotients for G. □

CHAPTER 4

Nilpotent groups

4.1. Central series and nilpotent groups

4.1.1. Definition. A *chain* $G = G_1 \geq G_2 \geq \dots \geq G_n \geq \dots$ is said to be *central* if for each i, we have $[G_i, G] \leq G_{i+1}$.

If $G_n = \{ 1 \}$ for some n, that we speak of a *central series.*

Note that the condition $[G_i, G] \leq G_{i+1}$ implies $[G_i, G] \leq G_i$, that is $G_i \trianglelefteq G$ for all i.

4.1.2. Definition. A group G is said to be *nilpotent* if it has a central series.

4.1.3. Lemma. *A nilpotent group is soluble.*

Proof. The quotients of a central series are abelian, as $[G_i, G_i] \leq [G_i, G] \leq G_{i+1}$. □

4.1.4. Exercise. *Show that S_3, A_4, S_4 are soluble but not nilpotent.*

4.1.5. Lemma. *For a group G and a series $G = G_0 \geq G_1 \geq \dots \geq G_n = \{ 1 \}$, the following are equivalent*

(1) the series is central, and

(2) for each i we have $G_i \trianglelefteq G$, and

$$\frac{G_{i-1}}{G_i} \leq Z\left(\frac{G}{G_i}\right).$$

Recall that for a group G

$$Z(G) = \{ z \in G : zx = xz \text{ for all } x \in G \}$$

is the centre (center) of G.

4.1.6. Exercise. *Show that the two conditions of Lemma 4.1.5 are equivalent, noting that $z \in Z(G)$ iff $[z, x] = 1$ for all $x \in G$.*

4.1.7. Definition. The *lower central chain* of the group G is defined by $\gamma_1(G) = G$, and $\gamma_{i+1}(G) = [\gamma_i(G), G]$, for $i \geq 1$. If $\gamma_n(G) = \{ 1 \}$ for some n, we speak of the *lower central series.*

The *upper central chain* of the group G is defined by $Z_0(G) = \{ 1 \}$, and

$$\frac{Z_i(G)}{Z_{i-1}(G)} = Z\left(\frac{G}{Z_{i-1}(G)}\right)$$

for $i \geq 1$. If $Z_n(G) = G$ for some n, we speak of the *upper central series.*

4.1.8. Exercise.

(1) *Show that these are indeed two central series.*
(2) *Show that $\gamma_2(G) = G'$ and $Z_1(G) = Z(G)$.*

4.1.9. Theorem. *Let G be a group, and*

$$G = G_1 \geq G_2 \geq \cdots \geq G_n = \{ 1 \}$$

a central series.

Then for each i one has

$$\gamma_i(G) \leq G_i \leq Z_{n-i}(G).$$

Proof. We have $\gamma_1(G) = G = G_1$, and then proceeding by induction

$$\gamma_{i+1}(G) = [\gamma_i(G), G] \leq [G_i, G] \leq G_{i+1}.$$

We have $G_{n-1}/G_n \leq Z(G/G_n) = Z_1(G)/\{ 1 \}$, so that $G_{n-1} \leq Z_1(G) = Z(G)$. Proceeding by backward induction, $[G_i, G] \leq G_{i+1} \leq Z_{n-i-1}(G)$, so that

$$G_i Z_{n-i-1}(G)/Z_{n-i-1}(G) \leq Z(G/Z_{n-i-1}(G)) = Z_{n-i}(G)/Z_{n-i-1}(G),$$

that is, $G_i \leq Z_{n-i}(G)$. □

4.1.10. Corollary. *Let G be a group. The following are equivalent*

(1) *G is nilpotent, that is, it has a central series,*
(2) *the lower central series terminates at $\{ 1 \}$,*
(3) *the upper central series terminates at G.*

4.1.11. Exercise. *Compute the lower and upper central chains for S_3, A_4, S_4.*

4.1.12. Theorem. *Let G be a group.*

(1) *If G is nilpotent and $H \leq G$, then H is nilpotent.*
(2) *If G is nilpotent, and $\varphi : G \to K$ is a morphism, then $\varphi(G)$ is nilpotent.*
(3) *If G is nilpotent and $N \trianglelefteq G$, then G/N is nilpotent.*

4.1.13. Exercise. *Show that item* (4) *of Theorem 3.2.6 does not hold with soluble replaced by nilpotent.*

Proof of Theorem 4.1.12. If $H \leq G$, we have $\gamma_i(H) \leq \gamma_i(G)$ for all i.

We have already seen that $\varphi(\gamma_2(G)) = \varphi(G') = \varphi(G)' = \gamma_2(\varphi(G))$. Proceeding by induction, we find

$$\varphi(\gamma_{i+1}(G)) = \varphi([\gamma_i(G), G]) = [\varphi(\gamma_i(G)), \varphi(G)] = [\gamma_i(\varphi(G)), \varphi(G)] = \gamma_{i+1}(\varphi(G)).$$

□

4.1.14. Lemma. *Let G be a nilpotent group. If $H < G$, then $H < N_G(H)$.*

Recall that

$$N_G(H) = \{ x \in G : h^x \in H \text{ for all } h \in H \} = \{ x \in G : [h, x] \in H \text{ for all } h \in H \}$$

is the largest subgroup of G in which H is normal.

Proof. Let i the the smallest number such that $\gamma_i(G) \leq H$, so that $\gamma_{i-1}(G) \not\leq H$. Then $[\gamma_{i-1}(G), H] \leq [\gamma_{i-1}(G), G] = \gamma_i(G) \leq H$, so that $\gamma_{i-1}(G) \leq N_G(H)$. □

4.2. Finite nilpotent groups

4.2.1. Lemma (Frattini argument).

(1) Suppose the group G acts on the set Ω, and $H \le G$ acts transitively on Ω. Then for $\alpha \in \Omega$ we have $G = G_\alpha H$.

(2) Let X be a finite group, $H \trianglelefteq X$ and let S a Sylow p-subgroup of H. Then $X = N_X(S)H$.

(3) Let G be a finite group, and S be a Sylow p-subgroup of G. Then $N_G(N_G(S)) = N_G(S)$.

Proof. If $g \in G$, then since H is transitive there is $h \in H$ such that $\alpha = (\alpha^g)^h = \alpha^{gh}$, so that $gh \in G_\alpha$ and $g \in G_\alpha H$.

Let Ω be the set of Sylow p-subgroup of H. Now X acts by conjugation on Ω, as $H \trianglelefteq X$; by Sylow's theorems, H acts transitively on Ω; the stabiliser of S is $N_X(S)$.

Let $H = N_G(S)$, and $X = N_G(N_G(S))$. Then $H \trianglelefteq X$, and thus $X = N_X(S)N_G(S) = N_G(S)$. □

4.2.2. Exercise. *Let G be a* finite *group. Show that if for each $H < G$ we have $H < N_G(H)$, then G is nilpotent.*

4.2.3. Theorem.

(1) A direct product of finitely many nilpotent groups is nilpotent.

(2) A finite p-group is nilpotent.

(3) A finite group G is nilpotent if and only if each p-Sylow subgroup is normal, so that G is the direct product of its distinct Sylow subgroups.

Proof. Note that if $G = A \times B$, and $a_1, a_2 \in A$, $b_1, b_2 \in B$, then $[a_1b_1, a_2b_2] = [a_1, a_2][b_1, b_2]$. It follows that if $G = H_1 \times \cdots \times H_n$, and k is large enough so that $\gamma_k(H_i) = \{ 1 \}$ for all i, then $\gamma_k(G) = \{ 1 \}$.

By the arguments of Section 1.7, if $P \ne \{ 1 \}$ is a p-group, then $Z(P) \ne \{ 1 \}$. It follows by induction that the upper central series terminates at P.

If G is nilpotent, and S is a Sylow p-subgroup, then we have $N_G(S) = N_G(N_G(S))$. It follows that $N_G(S) = G$, that is, $S \trianglelefteq G$. □

4.3. Nilpotent groups are soluble

We have already seen that a nilpotent group is soluble. Let us look at the relations between the lower central series and the derived series in a (nilpotent) group.

Let G be a group. We have $\gamma_2(G) = G^{(1)}$. Then $\gamma_3(G) = [\gamma_2(G), G] \ge [G^{(1)}, G^{(1)}] = G^{(2)}$. Proceeding by induction, if $\gamma_i(G) \ge G^{(i-1)}$, then $\gamma_{i+1}(G) = [\gamma_i(G), G] \ge [G^{(i-1)}, G^{(i-1)}] = G^{(i)}$.

4.3.1. Proposition (Hall-Witt Identity). *Let G be a group, $a, b, c \in G$. Then*

$$[a, b^{-1}, c]^b [b, c^{-1}, a]^c [c, a^{-1}, b]^a = 1.$$

The formula can be thought of as a group-theoretic version of the Jacobi identity in Lie algebras. It has wonderful geometric interpretations, see the blog post of the Fields medalist Terence Tao [**Tao12**], and also [**Cal11**].

Proof. Note that each factor comes from the previous one by the cyclic permutation $a \mapsto b \mapsto c \mapsto a$.

$$[a, b^{-1}, c]^b = b^{-1}[a, b^{-1}]^{-1}c^{-1}[a, b^{-1}]cb =$$
$$= b^{-1}[b^{-1}, a]c^{-1}[a, b^{-1}]cb = b^{-1}ba^{-1}b^{-1}ac^{-1}a^{-1}bab^{-1}cb =$$
$$= (a^{-1}b^{-1}ac^{-1}a^{-1})(bab^{-1}cb) = (aca^{-1}ba)^{-1}(bab^{-1}cb).$$

Set

$$U = aca^{-1}ba, V = bab^{-1}cb, W = cbc^{-1}ac.$$

Note that each element is obtained from the previous one by the cyclic permutation $a \mapsto b \mapsto c \mapsto a$. Thus

$$[a, b^{-1}, c]^b = U^{-1}V, [b, c^{-1}, a]^c = V^{-1}W, [c, a^{-1}, b]^a = W^{-1}U,$$

and the formula follows. □

4.3.2. Theorem (Hall's three-subgroup Lemma). *Let G be a group, $A, B, C \leq G$, and $N \trianglelefteq G$.*

If $[A, B, C], [B, C, A] \leq N$, then $[C, A, B] \leq N$

Proof. Let $a \in A, b \in B, c \in C$. Then

$$[c, a^{-1}, b]^a = [a, b^{-1}, c]^{-b}[b, c^{-1}, a]^{-c} \in N.$$

□

4.3.3. Corollary. $[\gamma_i(G), \gamma_j(G)] \leq \gamma_{i+j}(G)$.

It can be shown that an analogous formula does *not* hold for the upper central series.

Proof. Argue by induction on j, the case $j = 1$ following from the definition. If $j > 1$ we have

$$[\gamma_i(G), \gamma_j(G)] = [\gamma_i(G), [\gamma_{j-1}(G), G]] = [\gamma_{j-1}(G), G, \gamma_i(G)]$$

Now

$$[\gamma_i(G), \gamma_{j-1}(G), G] \leq [\gamma_{i+j-1}(G), G] = \gamma_{i+j}(G),$$
$$[G, \gamma_i(G), \gamma_{j-1}(G)] = [\gamma_{i+1}(G), \gamma_{j-1}(G)] \leq \gamma_{i+j}(G).$$

Therefore $[\gamma_i(G), \gamma_j(G)] \leq \gamma_{i+j}(G)$. □

4.3.4. Corollary. $G^{(i)} \leq \gamma_{2^i}(G)$.

Proof. We have $G^{(1)} = \gamma_2(G)$. Proceeding by induction on i,

$$G^{(i+1)} = [G^{(i)}, G^{(i)}] \leq [\gamma_{2^i}(G), \gamma_{2^i}(G)] \leq \gamma_{2^{i+1}}(G).$$

□

Part 3

Representations and Characters

CHAPTER 5

Representations

5.1. Permutation representations

Let G be a group G, Ω a non-empty set, and denote by $S(G)$ the group of permutations (bijective maps) on Ω. The following are well-known to be equivalent:

(1) A (right) action of G on Ω, and
(2) a morphism $\varphi : G \to S(\Omega)$.

The second instance will be called a *permutation representation* of G on Ω.

Given a group G, two important permutation representations, which occur in Cayley's Theorem, are

(1) The right regular representation
$$\begin{aligned} \mathfrak{r} : G &\to S(G) \\ g &\mapsto (x \mapsto xg), \end{aligned}$$
and
(2) the left regular representation
$$\begin{aligned} \mathfrak{l} : G &\to S(G) \\ g &\mapsto (x \mapsto g^{-1}x). \end{aligned}$$

5.2. Linear representations

Let G be a finite group, V a vector space of finite dimension n over the complex numbers $\mathbf{C}$. We will usually implicitly assume that $V = \mathbf{C}^n$ is the space of row vectors.

A *(linear) representation* of degree n of G is a morphism
$$\rho : G \to \mathrm{GL}(V) = \mathrm{GL}(n, \mathbf{C}).$$
So the group $\rho(G)$ is a group of matrices.

In the rest of these notes, the term *representation* will always stand for a *linear* representation.

As usual we have $\rho(g^{-1}) = \rho(g)^{-1}$. If $n = |\,G\,|$, we have $g^n = 1$ for $g \in G$, so that $\rho(g)^n = I$, where I is the identity map on V. This $\rho(g)$ has minimal polynomial dividing $x^n - 1$. It follows that the eigenvalues of $\rho(g)$ are n-th roots of unity, and they are distinct, so that $\rho(g)$ is diagonalizable.

5.3. Inner products

See Section 1.6 for the proper definitions.

We now show

5.3.1. Lemma. *There is an inner product on V such that the $\rho(g)$ are unitary matrices.*

Proof. In fact, let $\langle\,\cdot,\cdot\,\rangle$ be the standard inner product on V. Define, for $x, y \in V$

$$\langle\!\langle x, y \rangle\!\rangle = \sum_{g \in G} \langle x\rho(g), y\rho(g) \rangle \,.$$

It is easy to show that $\langle\!\langle \cdot, \cdot \rangle\!\rangle$ is again an inner product. Moreover for $h \in G$ we immediately have

$$\langle\!\langle x\rho(h), y\rho(h) \rangle\!\rangle = \sum_{g \in G} \langle x\rho(h)\rho(g), y\rho(h)\rho(g) \rangle = \sum_{k \in G} \langle x\rho(k), y\rho(k) \rangle = \langle\!\langle x, y \rangle\!\rangle,$$

so that each $\rho(h)$ is unitary with respect to $\langle\!\langle \cdot, \cdot \rangle\!\rangle$. In particular, with respect to a basis which is orthonormal for $\langle\!\langle \cdot, \cdot \rangle\!\rangle$, we will have

$$\rho(g)^{-1} = \rho(g)^* = \overline{\rho(g)^t},$$

as for $x, y \in V$ and $g \in G$ we have

$$\langle\!\langle x\rho(g)^*, y \rangle\!\rangle = \langle\!\langle x, y\rho(g) \rangle\!\rangle = \langle\!\langle x\rho(g)^{-1}\rho(g), y\rho(g) \rangle\!\rangle = \langle\!\langle x\rho(g)^{-1}, y \rangle\!\rangle,$$

and then since when $x, y \in V$ are written with respect to such a basis we have

$$\langle\!\langle x, y \rangle\!\rangle = \overline{x} \cdot y^t,$$

then

$$\langle\!\langle x, y\rho(g) \rangle\!\rangle = \overline{x} \cdot \rho(g)^t y^t, = \overline{x\overline{\rho(g)^t}} \cdot y^t, = \langle\!\langle x\overline{\rho(g)^t}, y \rangle\!\rangle.$$

□

5.4. From permutation representations to linear representations

To every permutation representation of a group G on a finite set Ω one can associate a linear representation, which we will also refer to as a permutation representation.

Assuming first for simplicity that $\Omega = \{\,1, 2, \dots, n\,\}$, let V be a space of dimension n, with basis $v_1, v_2, \dots, v_n$, and suppose G acts on Ω, that is, we have a permutation representation of G on Ω. Then

$$\begin{aligned} \rho : G &\to \mathrm{GL}(V) \\ g &\mapsto (v_i \mapsto v_{ig^{-1}}) \end{aligned}$$

is a (linear) representation of G.

In general, given a permutation representation of the finite group G on teh finite set Ω, let V be a vector space with base v_α, for $\alpha \in \Omega$, then

$$\begin{aligned} \rho : G &\to \mathrm{GL}(V) \\ g &\mapsto (v_\alpha \mapsto v_{\alpha g^{-1}}) \end{aligned}$$

is a (linear) representation of G.

In the particular case when G acts on itself by right multiplication (so that the permutation representation is $\mathfrak{r}$), we will also refer to the associated linear representation as the (right) *regular representation* of G.

5.5. Subrepresentations

If W is a subspace of V of dimension m, it may happen that W is *invariant under* $\rho(G)$, that is, for all $w \in W$ and $g \in G$, then $w\rho(g) \in W$. Then we get a *subrepresentation* of G

$$\rho^W : G \to \mathrm{GL}(W) = \mathrm{GL}(m, \mathbf{C}).$$

5.6. Examples

As an example, consider the cyclic group $G = \langle\, a \,\rangle \le S_3$ of order three, where $a = (132)$, so that G acts naturally on $\Omega = \{\, 1, 2, 3 \,\}$. Then

$$\rho(a) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}.$$

Consider $w_0 = e_1 + e_2 + e_3$. (Here the e_i are the elements of the standard basis of $\mathbf{C}^3$.) Then

$$w_0\rho(a) = e_1\rho(a) + e_2\rho(a) + e_3\rho(a) = e_2 + e_3 + e_1 = w_0,$$

so that $U = \langle\, w_0 \,\rangle$ is invariant under $\rho(G)$. Let $\omega = \exp(i\frac{2\pi}{3})$ be a primitive third root of unity. Consider the elements

$$\begin{cases} w_1 = e_1 + \omega e_2 + \omega^2 e_3 \\ w_2 = e_1 + \omega^2 e_2 + \omega e_3 \end{cases}$$

Then

$$w_1\rho(a) = e_2 + \omega e_3 + \omega^2 e_1 = \omega^2 w_1$$

and

$$w_2\rho(a) = e_2 + \omega^2 e_3 + \omega e_1 = \omega w_2.$$

It follows that $W = \langle\, w_1, w_2 \,\rangle$ is also $\rho(G)$-invariant.

5.6.1. Exercise.

(1) *Show that the w_i are a basis of V.*
(Hint: The matrix

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & \omega & \omega^2 \\ 1 & \omega^2 & \omega \end{bmatrix}$$

is a Vandermonde matrix.)

(2) *Show that $V = U \oplus W$.*

(3) *Show that in the basis w_1, w_2 of W, the (sub)representation $\rho^W : G \to \mathrm{GL}(W)$ is given by*

$$\rho^W(a) = \begin{bmatrix} \omega^2 & 0 \\ 0 & \omega \end{bmatrix}.$$

5.7. The group algebra

There is another way of describing representations. Let G be a finite group, and consider the set $\mathbf{C}[G] = \mathbf{C}^G$ of maps from G to $\mathbf{C}$, which is a vector space of dimension $|\,G\,|$ over $\mathbf{C}$. Define on $\mathbf{C}[G]$ a convolution product by

$$a * b(g) = \sum_{x,y \in G, xy=g} a(x) b(y).$$

This can be rewritten of course as

$$a * b(g) = \sum_{x \in G} a(x) b(x^{-1} g), \tag{5.7.1}$$

but the previous symmetric form is handier for the proofs. Note that this is similar to the product of polynomials.

With this operation, $\mathbf{C}[G]$ turns out to be an *algebra*, that is, a vector space over $\mathbf{C}$ endowed with an associative, bilinear product. Let us check associativity.

$$\begin{aligned} ((a * b) * c)(g) &= \sum_{t,z \in G, tz=g} ((a * b)(t)) c(z) \\ &= \sum_{t,z \in G, tz=g} \left(\sum_{x,y \in G, xy=t} a(x) b(y) \right) c(z). \\ &= \sum_{x,y,z \in G, xyz=g} a(x) b(y) c(z). \end{aligned}$$

One obtains the very same result with $a * (b * c)$.

5.7.1. Exercise. *Try and do this with the asymmetric form* (5.7.1).

5.7.2. Exercise. *Check the remaining properties.*

The vector space $\mathbf{C}[G]$ has a basis given by the δ_g, for $g \in G$, given by

$$\delta_g(x) = \begin{cases} 1 & \text{if } x = g \\ 0 & \text{otherwise,} \end{cases}$$

as for $a \in \mathbf{C}[G]$ we have uniquely

$$a = \sum_{g \in G} a(g) \delta_g.$$

Note that

$$\delta_g * \delta_h(x) = \sum_{y,z \in G, yz=x} \delta_g(y) \delta_h(z),$$

and the only non-zero summand is when $y = g$, $z = h$ so that $x = gh$. It follows that $\delta_g * \delta_h = \delta_{gh}$. It is therefore customary to write δ_g as g, and say that

$$\mathbf{C}[G] = \left\{ \sum_{g \in G} a_g g : a_g \in \mathbf{C} \right\} \tag{5.7.2}$$

is a vector space with basis the elements of G, and product that extends by linearity that of G.

The group algebra has the following universal property.

5.7.3. Proposition. *Let G be a group, A a $\mathbf{C}$-algebra.*

Let $\rho : G \to A$ be a morphism of groups. Then there is a unique morphism of algebras with unity $\rho' : \mathbf{C}[G] \to A$ that extends ρ.

Proof. ρ' is uniquely determined by (5.7.2). □

Therefore if $\rho : G \to \mathrm{GL}(V)$ is a representation, this extends uniquely to a morphism of algebras with unity $\rho' : \mathbf{C}[G] \to \mathrm{End}(V)$ given by

$$\rho'(\sum_{g \in G} a_g g) = \sum_{g \in G} a_g \rho(g).$$

Note that for $a \in \mathbf{C}$ one has $\rho'(a \cdot 1) = a \cdot I$, where $1 \in G$, and I is the identity map on V. And conversely, given such an algebra morphism ρ', its restriction to G is a representation of G.

Now the morphism ρ' defines the structure of a right $\mathbf{C}[G]$-module on V, as per Section 1.10. We have obtained

5.7.4. Proposition. *Let G be a finite group, and V be a finite dimensional $\mathbf{C}$-vector space. The following data are equivalent:*

(1) a linear representation of G on V, and

(2) a $\mathbf{C}[G]$-module structure on V.

Sometimes we will just speak of a *G-module* instead of a $\mathbf{C}[G]$-module. And when speaking of modules, we will simply write vg for $\rho(g)$, for $v \in V$ and $g \in G$.

Note that the right regular representation can be regarded as a representation $G \to \mathrm{GL}(\mathbf{C}[G])$.

Later we will need

5.7.5. Proposition. *The centre $Z(\mathbf{C}[G])$ of $\mathbf{C}[G]$ consists of the elements*

$$\sum_{g \in G} f(g) g$$

where $f : G \to \mathbf{C}$ is a class function, that is $f(g^h) = f(g)$ for all $g, h \in G$.

Proof. Immediate, just conjugate an element in the centre by $h \in G$. □

5.8. Maschke's Theorem

5.8.1. Theorem (Maschke). *Let $\rho : G \to \mathrm{GL}(V)$ be a representation of G.*

Suppose U is a $\rho(G)$-invariant subspace of V.

Then there is a $\rho(G)$-invariant subspace W of V such that

$$V = U \oplus W.$$

Since this is a fundamental result, we will give two proofs of it.

First proof of Maschke's theorem. The key to this is an averaging argument.

Let us choose an arbitrary subspace X of V such that $V = U \oplus X$. Of course X need not be $\rho(G)$-invariant.

Let $\pi : V \to U$ be the projection of V onto U along X, that is, if we write $v \in V$ as $v = u + x$, with $u \in U$ and $x \in X$, then $\pi(v) = u$.

Consider the following linear map on V

$$\psi = \frac{1}{|\,G\,|} \sum_{g \in G} \rho(g)^{-1} \pi \rho(g).$$

Thus for $v \in V$

$$v\psi = \frac{1}{|\,G\,|} \sum_{g \in G} ((v\rho(g)^{-1})\pi)\rho(g).$$

Note first that $V\psi \subseteq U$, as π maps V onto U, and U is $\rho(G)$-invariant. But in fact $V\psi = U$: if $u \in U$, then

$$u\psi = \frac{1}{|\,G\,|} \sum_{g \in G} ((u\rho(g)^{-1})\pi)\rho(g) = \frac{1}{|\,G\,|} \sum_{g \in G} (u\rho(g)^{-1})\rho(g) = \frac{1}{|\,G\,|} \sum_{g \in G} u = u,$$

as $u\rho(g)^{-1} \in U$ for all $g \in G$. Now $\psi^2 = \psi$, since $v\psi^2 = (v\psi)\psi = v\psi$, as $v\psi \in U$, and we have just shown that ψ restricts to the identity on U. By Lemma 1.1.1,

$$V = U \oplus \ker(\psi).$$

Now note that for $v \in V$ and $h \in G$ one has

$$\begin{aligned} v\rho(h)\psi &= v\rho(h)\frac{1}{|\,G\,|} \sum_{g \in G} \rho(g^{-1})\pi\rho(g) \\ &= v\frac{1}{|\,G\,|} \sum_{g \in G} \rho(hg^{-1})\pi\rho(g) \\ &= v\frac{1}{|\,G\,|} \sum_{g \in G} \rho(gh^{-1})^{-1}\pi\rho(g) \\ &= v\frac{1}{|\,G\,|} \sum_{k \in G} \rho(k)^{-1}\pi\rho(kh) \\ &= v\psi\rho(h) \end{aligned} \tag{5.8.1}$$

This implies that $\ker(\psi)$ is $\rho(G)$-invariant, as for $v \in \ker(\psi)$ and $g \in G$ we have

$$(v\rho(g))\psi = (v\psi)\rho(g) = 0.$$

□

Second proof of Maschke's theorem. We employ the inner product of Lemma 5.3.1. Let W be the orthogonal of U with respect to $\langle\langle \cdot, \cdot \rangle\rangle$, that is,

$$W = U^{\perp} = \{\, v \in V : \langle\langle u, x \rangle\rangle = 0 \text{ for all } u \in U \,\}.$$

We claim that W is $\rho(G)$-invariant. In fact if $w \in W$, then for all $g \in G$ and $u \in U$ we have

$$\langle\langle u, w\rho(g) \rangle\rangle = \langle\langle u\rho(g)^{-1}, w \rangle\rangle = 0,$$

as U is $\rho(G)$-invariant. □

5.8.2. Remark. If one looks at the proof that $V = U \oplus U^{\perp}$ of Section 1.6, where one employs a projection p onto U along $U^{\perp}$, one can see the similarity between the two proofs of Maschke's Theorem.

5.8.3. Definition. If $\rho_i : G \to \mathrm{GL}(V_i)$ are representations of the group G, for $i = 1, \dots, n$, then the *direct sum representation* is defined as

$$\begin{aligned}\rho_1 \oplus \rho_2 \oplus \dots \oplus \rho_n : G \to \mathrm{GL}(V_1) \times \mathrm{GL}(V_2) \times \dots \times \mathrm{GL}(V_n) \leq \mathrm{GL}(V_1 \oplus V_2 \oplus \dots \oplus V_n)\\ g \mapsto (v_1 \oplus v_2 \oplus \dots \oplus v_n \mapsto v_1\rho_1(g) \oplus v_2\rho_2(g) \oplus \dots \oplus v_n\rho_n(g))\end{aligned}$$

5.9. Irreducible representations

5.9.1. Definition. A representation $\rho : G \to \mathrm{GL}(V)$ is said to be *irreducible* if the only $\rho(G)$-invariant subspaces of V are $\{ 0 \}$ and V itself.

Maschke's Theorem implies immediately (by induction on the dimension of the vector space)

5.9.2. Theorem. *Every finite-dimensional representation over* $\mathbf{C}$ *is a direct sum of irreducible ones.*

One should compare this result to the fact that two groups may have the same composition factors, without being isomorphic. For instance, as noted earlier, both C_6 and S_3 have two composition factors that are cyclic groups of order 2 and 3. So this *simple* constituents alone do not determine a group uniquely, as they can be put together in different ways. (In this particular case, there are two non-isomorphic semidirect products of C_3 by C_2.)

With a representation, instead, if you know its irreducible subrepresentations, the representation is uniquely determined as the direct sum of them, there is only one (trivial) way of gluing them together.

5.9.3. Exercise. *The statement is not true anymore over fields F whose characteristic divides the order of G. The simplest example is given by $G = \langle\, a \,\rangle$ cyclic of order 2, and by the representation over the field F with two elements defined by*

$$\begin{aligned}\rho : G &\to F^2\\ a &\mapsto \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.\end{aligned}$$

5.10. Morphisms and modules over the group algebra

5.10.1. Definition. Let $\rho_i : G \to \mathrm{GL}(V_i)$ be representations of G, for $i = 1, 2$. A *morphism* of representations is a linear map $f : V_1 \to V_2$ such that for $g \in G$

$$\rho_1(g) f = f \rho_2(g).$$

This is a property we had already noted for the projection ψ in (5.8.1) in the course of the proof of Maschke's Theorem. The definition becomes clearer once one keeps in mind the equivalence of group representations and group algebra modules described in Section 5.7.

This yields

- $\rho(G)$-invariant subspaces are nothing else but $\mathbf{C}[G]$-submodules, and
- morphisms of representations are nothing else but morphisms of $\mathbf{C}[G]$-modules.

If a morphism f of representations (modules) is bijective, we call it of course an *isomorphism.* So ρ_1, ρ_2 are isomorphic if there is f such that

$$\rho_2(g) = f^{-1}\rho_1(g)f, \qquad \text{for all } g \in G.$$

5.11. Schur's Lemma

5.11.1. Theorem (Schur's Lemma). *Let $\rho_i : G \to \mathrm{GL}(V_i)$ be irreducible representations of G, for $i = 1, 2$, and $f : V_1 \to V_2$ a morphism.*

Then

- *either $f = 0$,*
- *or f is an isomorphism (that is, f is bijective).*

Proof. One sees easily that $\ker(f) \subseteq V_1$ is $\rho_1(G)$-invariant, and $V_1 f \subseteq V_2$ is $\rho_2(G)$-invariant.

If $V_1 f = \{0\}$, then $f = 0$. If $V_1 f \neq \{0\}$, then $V_1 f = V_2$, as ρ_2 is irreducible, and $\ker(f) \neq V_1$, so that $\ker(f) = \{0\}$, as ρ_1 is irreducible. □

5.11.2. Corollary. *Let $\rho : G \to \mathrm{GL}(V)$ be irreducible. If $f : V \to V$ is an isomorphism, then there is $\lambda \in \mathbf{C}^*$ such that $vf = \lambda v$ is multiplication by the scalar λ.*

Proof. Let λ be an eigenvalue of f. (We are exploiting for the first time the fact that $\mathbf{C}$ is algebraically closed.)

Then from

$$\rho_1(g)f = f\rho_2(g), \qquad \text{and} \qquad \rho_1(g)(\lambda I) = (\lambda I)\rho_2(g),$$

where I is the identity on V, we get

$$\rho_1(g)(f - \lambda I) = (f - \lambda I)\rho_2(g).$$

Thus $f - \lambda I$ is also a morphism of representations. Since it is singular, it must be zero, so that $f = \lambda I$ is scalar multiplication by λ. □

5.11.3. Lemma. *Let $\rho_i : G \to \mathrm{GL}(V_i)$ be representations of G, for $i = 1, 2$. Let $f : V_1 \to V_2$ be any linear map. Then*

$$f' = \sum_{g \in G} \rho_1(g^{-1}) f \rho_2(g)$$

is a morphism of representations $V_1 \to V_2$.

Proof. Very much as in Maschke's Theorem. □

5.12. Orthogonality Relations

5.12.1. Proposition. *Let $\rho_i : G \to \mathrm{GL}(V_i)$, for $i = 1, 2$, be non-isomorphic, irreducible representations. Take V_1, V_2 to be spaces of row vectors, with standard bases $u_1, \dots, u_n$ and $v_1, \dots, v_m$. Let $\rho_i^{jk}(g)$ denote the (j, k)-component of $\rho_i(g)$.*

Then for all j, s, t, l we have

$$\sum_{g \in G} \rho_1^{js}(g^{-1}) \rho_2^{tl}(g) = 0 \tag{5.12.1}$$

5.12.2. REMARK. By choosing the bases to be orthonormal as in Lemma 5.3.1, then (5.12.1) can be rewritten as

$$\sum_{g\in G} \overline{\rho_1^{sj}(g)}\rho_2^{tl}(g) = 0$$

because

$$\rho_1(g^{-1}) = \rho_1(g)^{-1} = \rho_1(g)^* = \overline{\rho_1(g)^t}.$$

PROOF. Let $E_{st} : V_1 \to V_2$ be the linear map that sends all u_i to 0, except $u_s E_{s,t} = v_t$.

Then

$$\begin{aligned} u_j\rho_1(g^{-1})E_{s,t}\rho_2(g) &= \sum_{k=1}^{n} \rho_1^{jk}(g^{-1})u_k E_{s,t}\rho_2(g) \\ &= \rho_1^{js}(g^{-1})v_t\rho_2(g) \\ &= \sum_{l=1}^{m} v_l\rho_1^{js}(g^{-1})\rho_2^{tl}(g). \end{aligned}$$

Summing over $g \in G$ we get

$$0 = \sum_{l=1}^{m} v_l \sum_{g\in G} \rho_1^{js}(g^{-1})\rho_2^{tl}(g),$$

so that

$$\sum_{g\in G} \rho_1^{js}(g^{-1})\rho_2^{tl}(g) = 0$$

for all j, s, t, l. □

5.12.3. PROPOSITION. *Let ρ be an irreducible representation. Take V to be spaces of row vectors, with standard basis $u_1, \ldots, u_n$. Let $\rho^{jk}(g)$ denote the (j, k)-component of $\rho(g)$.*

Then for all j, s, t, l we have

$$\sum_{g\in G} \overline{\rho^{sj}(g)}\rho^{tl}(g) = \begin{cases} \dfrac{|\,G\,|}{n} & \textit{if } j = l \textit{ and } s = t \\ 0 & \textit{otherwise.} \end{cases}$$

PROOF. In the notation of the previous proof, take first $s \neq t$. Then

$$\operatorname{trace}(\rho(g^{-1})E_{s,t}\rho(g)) = \operatorname{trace}(E_{s,t}) = 0,$$

so that

$$\sum_{g\in G} \rho^{js}(g^{-1})\rho^{tl}(g) = 0$$

as in the previous proof.

When $s = t$, we have

$$\operatorname{trace}(\rho(g^{-1})E_{s,s}\rho(g)) = \operatorname{trace}(E_{s,s}) = 1,$$

so that

$$\operatorname{trace}(\sum_{g\in G} \rho(g^{-1})E_{s,s}\rho(g)) = |\,G\,|,$$

and thus $\sum_{g\in G}\rho(g^{-1})E_{s,s}\rho(g)$ has to be an isomorphism $V \to V$. By Corollary 5.11.2, we have that

$$\sum_{g\in G}\rho(g^{-1})E_{s,s}\rho(g)$$

is scalar multiplication by $\dfrac{|\,G\,|}{n}$. The result follows as in the previous proof. □

5.12.4. Remark. The formula of Proposition 5.12.3 also holds if instead of a single irreducible representation ρ one considers two (possibly different) irreducible representation isomorphic representations ρ_1, ρ_2. This is because by definition there is a bijective linear map f such that for all $g \in G$ one has $\rho_2(g) = f^{-1}\rho_1(g)f$, so that

$$\operatorname{trace}(\rho_1(g^{-1})E_{s,t}\rho_2(g)) = \operatorname{trace}(\rho_1(g^{-1})E_{s,t}f^{-1}\rho_1(g)f) = \operatorname{trace}(E_{s,t}).$$

CHAPTER 6

Characters

6.1. Characters

6.1.1. Definition. The *character* of a representation ρ is the map

$$\begin{aligned}\chi : G &\to \mathbf{C}\\ g &\mapsto \operatorname{trace}(\rho(g)).\end{aligned}$$

One says that ρ *affords* χ.

We have seen in Subsection 5.9 that two representations ρ_1, ρ_2 are isomorphic if and only if there is a bijective linear map such that

$$\rho_2(g) = f^{-1}\rho_1(g)f, \qquad \text{for all } g \in G.$$

Therefore

$$\operatorname{trace}(\rho_2(g)) = \operatorname{trace}(f^{-1}\rho_1(g)f) = \operatorname{trace}(\rho_1(g)).$$

In other words

6.1.2. Proposition. *The character of a representation only depends on the isomorphism type of the representation.*

6.1.3. Remark. If $a, b : G \to \mathbf{C}$ are two functions, then setting

$$(a, b) = \frac{1}{|\,G\,|}\sum_{g\in G}\overline{a(g)}b(g)$$

we obtain an inner product in the space of all such functions.

6.1.4. Theorem (Orthogonality relations).

(1) *If χ, ψ are the characters of two non-isomorphic, irreducible representations, then*

$$(\chi, \psi) = 0.$$

(2) *If χ is the character of an irreducible representation, then*

$$(\chi, \chi) = 1.$$

(3) *Summing it up, the characters of the irreducible representations are orthonormal.*

Proof. The first statement is clear from Proposition 5.12.1. The second one follows from Proposition 5.12.1, as

$$\begin{aligned}(\chi, \chi) &= \frac{1}{|\,G\,|} \sum_{g \in G} \left(\sum_{i=1}^{n} \overline{\rho^{ii}(g)} \right) \left(\sum_{j=1}^{n} \rho^{jj}(g) \right) \\ &= \frac{1}{|\,G\,|} \sum_{i=1}^{n} \sum_{g \in G} \overline{\rho^{ii}(g)} \rho^{ii}(g) \\ &= \frac{1}{|\,G\,|} \cdot n \cdot \frac{|\,G\,|}{n} = 1.\end{aligned}$$

□

We obtain an important fact

6.1.5. Theorem. *The character of a representation determines the representation up to isomorphism.*

In case you are interested in how to get back from a character to the (unique) representation it affords it, check [**Spe10**].

Proof. Let ρ be a representation, and decompose it as

$$\rho = n_1 \rho_1 \oplus n_2 \rho_2 \oplus \cdots \oplus n_t \rho_t,$$

where the ρ_i are pairwise non-isomorphic irreducible representations, and n_i denotes the number of times ρ_i occurs.

Taking the trace, we get

$$\chi = n_1 \chi_1 \oplus n_2 \chi_2 \oplus \cdots \oplus n_t \chi_t,$$

where χ is the character of ρ, and χ_i is the character of ρ_i.

Now

$$(\chi_i, \chi) = n_i.$$

Thus the character χ determines the n_i.

It follows in particular that the n_i are only determined by the isomorphism type of ρ. □

This allows for the following

6.1.6. Definition. A character is said to be *irreducible* if the corresponding representation is irreducible.

The set of the irreducible characters of G is denoted by $\mathrm{Irr}(G)$.

6.2. Decomposing the regular representation

Let $\rho : G \to \mathrm{GL}(\mathbf{C}[G])$ be the right regular representation, and ψ its character. Since $x\rho(g) = xg = x$ only if $g = 1$, we get

$$\psi(g) = \begin{cases} |\,G\,| & \text{if } g = 1 \\ 0 & \text{otherwise.} \end{cases}$$

Let χ be an irreducible character. Then

$$(\psi, \chi) = \frac{1}{|\,G\,|} \sum_{g \in G} \overline{\psi(g)} \chi(g) = \frac{1}{|\,G\,|} \psi(1)\chi(1) = \frac{1}{|\,G\,|} |\,G\,| \chi(1) = \chi(1).$$

6.2.1. Definition. The *degree* of a character χ is $\chi(1) = \operatorname{trace}(I)$, which equals the dimension of the vector space associated to representation that affords χ.

It follows

6.2.2. Theorem.

(1) *Every irreducible representation with character χ occurs $\chi(1)$ times in the decomposition of the regular representation as a sum of irreducible representation.*

(2) *If ψ is the character of the regular representation, then*

$$\psi = \sum_{\chi} \chi(1)\chi,$$

where χ ranges over the irreducible characters of G.

(3)

$$|\,G\,| = \sum_{\chi} \chi(1)^2, \tag{6.2.1}$$

where χ ranges over the irreducible characters of G.

Proof. It remains only to prove the important last statement, which follows by evaluating (2) at 1. □

6.2.3. Corollary. *Let ρ_i be the pairwise non-isomorphic, irreducible representations, and ρ_i^{jk} be their components. Then the ρ_i^{jk} are an orthogonal basis for the space of functions $G \to \mathbf{C}$.*

Proof. We have see in Subsection 5.12 that these functions are orthogonal. Now (6.2.1) shows that their number equals the dimension of the space of functions $G \to \mathbf{C}$. □

6.3. Number of irreducible characters

Here I am following [**Ser16**].

If $\rho_i : G \to \mathrm{GL}(V_i)$ are the pairwise non-isomorphic irreducible representations of G, then we have an algebra morphism

$$r : \mathbf{C}[G] \to \sum_i \operatorname{End}(V_i).$$

The two spaces have the same dimension, by (6.2.1). We will show that r is injective, and this will imply that r is an algebra isomorphism. But an element of $\ker(r)$ acts as 0 in every irreducible representation, thus in the regular representation, which is faithful (action on $1 \in \mathbf{C}[G]$).

Now the centre of the algebra $\sum_i \operatorname{End}(V_i)$ has dimension the number of distinct irreducible characters. The centre of $\mathbf{C}[G]$ has a basis given by the sums over the conjugacy classes, and thus the dimension of the centre equals the number of conjugacy classes. We obtain

6.3.1. Theorem. *The number of distinct irreducible characters (that is, of pairwise non-isomorphic representations) of G equals the number of conjugacy classes of G.*

The isomorphism r induces an isomorphism from $Z(\mathbf{C}[G])$ to the centre of $\sum_i \operatorname{End}(V_i)$, which is a sum of copies of $\mathbf{C}$ (scalar matrices for each $\operatorname{End}(V_i)$). Choose one of the V_i, associated to the irreducible representation ρ with character χ. Consider the map r_χ, which is the composition of r with the projection on $\operatorname{End}(V_i)$, and which is simply given by

$$r_\chi(\sum_{g\in G} f(g)g) = \sum_{g\in G} f(g)\rho(g) \tag{6.3.1}$$

Let now $\alpha = \sum_{g\in G} f(g)g \in Z(\mathbf{C}[G])$; this goes under r_χ to a scalar matrix S in $Z(\operatorname{End}(V_i))$, which is multiplication by some a. Since $a = \operatorname{trace}(S)/\chi(1)$, (6.3.1) yields that the value of a is

$$\frac{1}{\chi(1)} \cdot \operatorname{trace}(\sum_{g\in G} f(g)\rho(g)) = \frac{1}{\chi(1)} \sum_{g\in G} f(g)\chi(g) = \frac{|\,G\,|}{\chi(1)}(f, \overline{\chi}).$$

6.3.2. Theorem. *Let $\sum_{g\in G} f(g)g \in Z(\mathbf{C}[G])$.*
Then $r_\chi(\sum_{g\in G} f(g)g)$ is the scalar matrix that is multiplication by

$$\frac{|\,G\,|}{\chi(1)}(f, \overline{\chi}).$$

6.4. Representation and characters from quotients

If N is a normal subgroup of the finite group G, and ρ is a representation of G/N, then $\rho(g) = \rho(gN)$, for $g \in G$, defines a representation of G, whose character is $\chi'(g) = \chi(gN)$, if χ is the character of ρ.

6.5. Products of representations and characters

If $\rho_i : G \to \operatorname{GL}(V_i)$ are two representations of G, for $i = 1, 2$, then (1.14.4) shows that

$$\begin{aligned} \rho_1 \otimes \rho_2 : G &\to \operatorname{GL}(V_1 \otimes_{\mathbf{C}} V_2) \\ g &\mapsto \rho_1(g) \otimes \rho_2(g) \end{aligned}$$

is a representation. If χ_i is the character of ρ_i, then (1.14.3) shows that the character χ of $\rho_1 \otimes \rho_2$ is the product of χ_1 and χ_2,

$$\chi(g) = \chi_1(g)\chi_2(g).$$

There is a special case of this that does not require the technicalities of the tensor product. Recall that a character λ is linear if $\lambda(1) = 1$, so that $\lambda : G \to GL(1, \mathbf{C})$ coincides with the representation that affords it.

Now, if $\rho : G \to GL(V)$ is a representation with character χ, and λ is a linear character, then

$$\begin{aligned} \lambda \cdot \rho : G &\to \mathrm{GL}(V) \\ g &\mapsto \lambda(g)\rho(g) \end{aligned}$$

is visibly also a representation, whose character is $g \mapsto \lambda(g)\chi(g)$.

6.6. Kernels and centres

If $\rho : G \to \mathrm{GL}(V)$ is a representation of degree n, with character χ, and $g \in \ker(\rho)$, then $\chi(g) = n$. Conversely, if $\chi(g) = n$, then Lemma 1.11.13 shows that $g \in \ker(\rho)$. So we can write

$$\ker(\chi) = \{ g \in G : \chi(g) = \chi(1) \},$$

and $\ker(\chi) = \ker(\rho)$.

6.6.1. Lemma. *Let χ be a character of G, and $\chi = \sum n_i \chi_i$, with $\chi \in \mathrm{Irr}(G)$. Then*

$$\ker(\chi) = \cap \{ \ker(\chi_i) : n_i > 0 \}. \tag{6.6.1}$$

Taking χ to be the regular character, we see that the intersection of the kernels of all irreducible characters is $\{ 1 \}$.

Proof. The $\supseteq$ inclusion in (6.6.1) is clear.

For the reverse inclusion $\subseteq$, write ρ, ρ_i for the representations affording χ, χ_i. If $g \in \ker(\chi) = \ker(\rho)$, then $\rho(g)$ is the identity matrix. This is independent of the choice of a basis, so this means each $\rho_i(g)$ are identity matrices, so $g \in \ker(\rho_i) = \ker(\chi_i)$ for all i. □

The *centre* of a character χ is

$$Z(\chi) = \{ g \in G : |\, \chi(g) \,| = \chi(1) \} \geq \ker(\chi).$$

6.6.2. Lemma. *Let χ be a character of the representation ρ of the group G. Then*

(1) $Z(\chi) = \{ g \in G : \rho(g)$ is scalar$\}$;
(2) $Z(\chi) \leq G$;
(3) $\chi_{Z(\chi)} = \chi(1)\lambda$ for a linear character λ of $Z(\chi)$;
(4) $Z(\chi)/\ker(\chi)$ is cyclic;
(5) $Z(\chi)/\ker(\chi) \leq Z(G/\ker(\chi))$.

If $\chi \in \mathrm{Irr}(G)$ we have also

(6) $Z(\chi)/\ker(\chi) = Z(G/\ker(\chi))$.

Proof. By Lemma 1.11.13, $|\, \chi(g) \,| = \chi(1)$ if and only if $\rho(g)$ is scalar.

Let $\lambda : Z(\chi) \to \mathbf{C}$ be defined by $\rho(g) = \lambda(g)I$, according to the previous point. We have $\lambda(gh)I = \rho(gh) = \rho(g)\rho(h) = \lambda(g)\lambda(h)I$, so that $Z(\chi)$ is a subgroup, and λ is a linear character.

Since $\ker(\chi) = \ker(\lambda)$, we have that $Z(\chi)/\ker(\chi)$ is isomorphic to a finite subgroup of $\mathbf{C}^*$, and thus is cyclic. Also, $\ker(\chi) = \ker(\rho)$, and $\rho(Z(\chi))$ is in the centre of $\rho(G) \cong G/\ker(\chi)$, as it is made of scalar matrices.

Finally, if $\chi \in \mathrm{Irr}(G)$, then (to be completed, but straightforward). □

6.7. Character tables

6.7.1. Definition. The *character table* of G is a square matrix, which the rows labelled by the distinct irreducible characters, and the columns labelled by the conjugacy classes.

The χ, a^G entry is $\chi(a)$.

6.8. Characters of abelian groups

6.8.1. Theorem. *Let G be a finite abelian group.*

Then each irreducible character χ of G is linear*, that is, it has $\chi(1) = 1$.*

It follows that χ coincides with the representations which affords it, so that $\chi : G \to \mathbf{C}^*$ is a morphism of groups.

This follows from Section 1.5: a finite number of commuting, diagonalizable matrices can be diagonalized simultaneously, that is, there is a basis with respect to which they are all diagonal.

Alternatively, an abelian group of order n has n conjugacy classes, and thus n irreducible characters. Since $n = \sum_{\chi \in \mathrm{Irr}(G)} \chi(1)^2 \geq n$, each $\chi(1)$ must be 1.

Let $G = \langle\, g \,\rangle$ be cyclic of order n. Let ω be a primitive n-th root of unity. Then the irreducible characters of G are

$$\begin{aligned} \chi_k : G &\to \mathbf{C}^* \\ g &\mapsto \omega^k \end{aligned}$$

for $0 \leq k < n$, so that $\chi_k(g^j) = \omega^{jk}$.

The character table of a cyclic group is a Vandermonde matrix, for instance when $n = 3$, it is

	1	g	g^2
χ_0	1	1	1
χ_1	1	ω	ω^2
χ_2	1	ω^2	ω

Note that $\chi_h \cdot \chi_k(g) = \omega^{h+k} = \chi_{h+k}(g)$, so the characters form a group isomorphic to G.

This holds more generally for the *dual group* of characters of a finite abelian group G. In fact, if

$$G \cong \mathbf{Z}/n_1\mathbf{Z} \times \cdots \times \mathbf{Z}/n_t\mathbf{Z},$$

with, say $n_1 \geq n_2 \geq \cdots \geq n_t > 0$, then the irreducible characters are

$$\begin{aligned} \chi_{k_1,\dots,k_t} : G &\to \mathbf{C}^* \\ (x_1, \dots, x_t) &\mapsto \omega_1^{k_1 x_1} \cdot \cdots \cdot \omega_t^{k_t x_t} \end{aligned}$$

where ω_j is a primitive n_j-th root of unity. We have

$$\chi_{h_1,\dots,h_t} \cdot \chi_{k_1,\dots,k_t} = \chi_{h_1+k_1,\dots,h_t+k_t}.$$

Recall from Section 6.4 that if $N \trianglelefteq G$, then any representation $\rho : G/N \to \mathrm{GL}(V)$ of the quotient group G/N lifts to a representation $\rho' : G \to G/N \to \mathrm{GL}(V)$ of G, and so does the corresponding character.

In particular, the linear characters of a group G are $|\, G/G' \,|$, the liftings of the irreducible characters of the abelian group G/G'.

6.9. Induced Characters

6.9.1. Introduction. Let G be a group, $H \le G$, and $\rho : G \to \mathrm{GL}(V)$ be a representation of G. Clearly $\rho_H : H \to \mathrm{GL}(V)$ is a representation of H.

However, if $\sigma : H \to \mathrm{GL}(V)$ is a representation of H, it is not always the case that $\sigma = \rho_H$, for a representation ρ of G. For instance, let $g = (123) \in S_3$. The cyclic group $H = \langle\, g \,\rangle$ of order 3 has a linear representation σ such that $\sigma() = \omega$ is a primitive third root of unity. However, $G = S_3 \ge H$ has no (linear) representation ρ which takes the value ω on G.

The method of induced representations remedy in some sense the situation. Given a representation σ of H on the vector space W , the method yields a representation σ^G of G on a larger vector space V which is natural in some sense, and with the property that if σ is the restriction to H of an irreducible representation ρ of G, then ρ is a constitutent of σ^G.

6.9.2. Serre's heuristic approach. Let us start with the *heuristic* approach of [**Ser78**, 3.3].

Let G be a finite group, and V a $\mathbf{C}[G]$-module. (The language of modules turns handy here, but it can be translated any time in terms of representations.)

Let $H \le G$, and let W be a $\mathbf{C}[H]$-submodule of V. Let $\mathcal{T}$ be a complete set of representatives for the right cosets of H in G, that is, every such coset can be written as Ht, for a unique $t \in \mathcal{T}$. Then note that then set $Wt = W(Ht)$ only depends on the coset Ht, and not on the choice of a particular representative. Moreover each Wt is a subspace of V, and it is a $\mathbf{C}[H^t]$ module, as $wt(t^{-1}ht) = (wh)t \in Wt$ for $w \in W$ and $h \in H$.

Consider the sum $S = \sum_{t \in \mathcal{T}} Wt$ (this is just the set of all sums from the summands). This is a $\mathbf{C}[G]$-submodule of V, as for $g \in G$ we will have $tg = ht'$ for some $h \in H$ and $t' \in \mathcal{T}$, so that $(Wt)g = (Wh)t' = Wt' \subseteq S$.

We will say that the $\mathbf{C}[G]$-module (and the corresponding representation) is *induced* from the $\mathbf{C}[H]$-module W if there is a direct sum decomposition

$$V = \bigoplus_{t \in \mathcal{T}} Wt.$$

6.9.1. Theorem. *Let G be a finite group, and $H \le G$.*

Let W be a $\mathbf{C}[H]$-module.

Then there exists a $\mathbf{C}[G]$-module induced by W, and this is unique up to isomorphism.

Proof. One can see that it is possible to reduce to the case when W is irreducible.

Then W is a submodule of the regular module $W = \mathbf{C}[H]$ for the group algebra $\mathbf{C}[H]$.

We claim that $V = \mathbf{C}[G]$ is the $\mathbf{C}[G]$-module induced by W. In fact an arbitrary element of $\mathbf{C}[G]$ can be written as

$$\sum_{g \in G} f(g)g = \sum_{t \in \mathcal{T}} \left(\sum_{h \in H} f(ht)h \right) t \in \sum_{t \in \mathcal{T}} Wt,$$

where $f : G \to \mathbf{C}$, and this representation is clearly unique.

It remains to show that if the $\mathbf{C}[G]$-module is induced by the $\mathbf{C}[H]$-module W, and W' is a $\mathbf{C}[H]$-submodule of W, then

$$V' = \sum_{t \in \mathcal{T}} W't$$

is a $\mathbf{C}[G]$-submodule of V, which is induced by W', but this is pretty straightforward.

We skip uniqueness for the moment. □

If $(Wt)g \neq Wt$ for all $t \in \mathcal{T}$, then $\psi(g) = 0$, as the blocks Wt are permuted without fixed points. Now $Wtg = Wt$

6.9.3. Same problem in terms of characters. Taken from [**Isa06**]. There is some duplication here.

If ρ is a representation of G, and $H \leq G$, then the restriction ρ_H of ρ to H is clearly a representation of H.

If we have a representation of $H \leq G$, can we make it into a representation of G? This is possibly best understood in terms of characters. If χ is a character on H, then this is a class function on H, but it need not be a class function on G.

6.9.2. Exercise. *Let $G = A_4$, $H = \{ 1, (12)(34), (13)(24), (14)(23) \}$. Show that $\rho : H \to \mathrm{GL}(1, \mathbf{C}) = \mathbf{C}^*$ defined by*

$$\rho((12)(34)) = 1, \qquad \rho((13)(24)) = -1$$

is a representation/character for H. Show that this is not (the restriction of) a class function on G, as $(12)(34))$ and $((13)(24)$ are conjugate in G.

We can fix this as follows. First extend χ to χ°, which is zero outside H. Then define

$$\chi^G(x) = \frac{1}{|\,H\,|} \sum_{g \in G} \chi^\circ(x^g).$$

This is clearly a class function, and $\chi^G(1) = \dfrac{|\,G\,|}{|\,H\,|}\chi(1)$.

To show that χ^G is character of G, we appeal to

6.9.3. Proposition (Frobenius Reciprocity). *Let $H \leq G$. Let φ be a class function on H, and ϑ a class function on G.*

Then

$$(\vartheta, \varphi^G)_G = (\vartheta_H, \varphi)_H,$$

where the two scalar products are in G and H respectively.

PROOF.

$$\begin{aligned}
(\vartheta, \varphi^G)_G &= \frac{1}{|G|} \sum_{x \in G} \overline{\vartheta(x)} \varphi^G(x) \\
&= \frac{1}{|G|} \frac{1}{|H|} \sum_{x, g \in G} \overline{\vartheta(x)} \varphi^\circ(x^g) \\
&= \frac{1}{|G|} \frac{1}{|H|} \sum_{y, g \in G} \overline{\vartheta(y^{g^{-1}})} \varphi^\circ(y) \\
&= \frac{1}{|H|} \sum_{y \in G} \overline{\vartheta(y)} \varphi^\circ(y) \\
&= \frac{1}{|H|} \sum_{y \in H} \overline{\vartheta(y)} \varphi(y) \\
&= (\vartheta_H, \varphi)_H.
\end{aligned}$$

□

6.9.4. COROLLARY. *If φ is a character of $H \leq G$, then φ^G is a character of G.*

PROOF. We have seen that $\varphi^G(1) = \frac{|G|}{|H|}\varphi(1)$, so $\varphi^G \neq 0$. As a class function on G, φ^G is a $\mathbf{C}$-linear combination of irreducible characters of G. Now if $\chi \in \mathrm{Irr}(G)$, we have $(\varphi^G, \chi)_G = (\varphi, \chi_H)$, and since χ_H is a character of H, the latter is a non-negative integer. It follows φ^G is a character of G. □

6.9.5. COROLLARY. *Let $H \leq G$, and $\varphi \in \mathrm{Irr}(H)$. Then φ is a constituent of χ_H, for some $\chi \in \mathrm{Irr}(G)$.*

6.9.6. DEFINITION. An irreducible character φ is a *constituent* of a character ψ is $(\varphi, \psi) \neq 0$.

PROOF. Let χ be an irreducible constituent of φ^G. Then

$$0 \neq (\varphi^G, \chi) = (\varphi, \chi_H),$$

so that φ is a constituent of χ_H. □

6.9.7. COROLLARY. *Let G be an abelian group, and $H \leq G$.*

Then every irreducible character of H is the restriction to H of an irreducible character of G.

PROOF. By the previous result, if $\varphi \in \mathrm{Irr}(H)$, then φ is a constituent of χ_H, for some $\chi \in \mathrm{Irr}(G)$. Since both characters have degree 1, they must be equal. □

6.9.4. Tensor products. We now mention how induced characters arise from tensor products. Let G be a finite group, and $H \leq G$. Let $V = V_{\mathbf{C}[H]}$ be a finite-dimensional $\mathbf{C}[H]$ module. Since $\mathbf{C}[G] = {}_{\mathbf{C}[H]}\mathbf{C}[G]_{\mathbf{C}[G]}$ is a $(\mathbf{C}[H], \mathbf{C}[G])$-bimodule,

$$W = V_{\mathbf{C}[H]} \otimes_{\mathbf{C}[H]} {}_{\mathbf{C}[H]}\mathbf{C}[G]_{\mathbf{C}[G]}$$

becomes a right $\mathbf{C}[G]$-module.

Let T be a *complete set of representative of the right cosets of H in G*. The idea is, let $G/H = \{ Hg : g \in G \}$ be the *set* of right cosets of H in G. The map

$$\begin{aligned} G &\to G/H \\ g &\mapsto Hg \end{aligned}$$

is surjective, hence it has a one-sided inverse (actually, may), that is, there are maps $\tau : G/H \to G$ such that $Hg = H\tau(Hg)$ for all $g \in G$. Every such map thus selects a representative for each coset. The image T of any such map τ is called a *complete set of representative of the right cosets of H in G*. A coset can be written uniquely as Ht, for $t \in T$, and an element of G can be written uniquely as ht, for $h \in H, t \in T$.

Note that if $g = ht$, for $h \in H, t \in T$, then $H^g = g^{-1}Hg = t^{-1}h^{-1}Hht = t^{-1}Ht = H^t$, so all conjugates of H are of the latter form.

We then have that as a vector space W decomposes as

$$(V \otimes t_1) \oplus \cdots \oplus (V \otimes t_n),$$

as

$$v \otimes (\sum_{g \in G} a_g g) = \sum_{g \in G} a_g v \otimes g = \sum_{h \in H} \sum_{t \in T} a_{ht} v \otimes ht = \sum_{t \in T} (\sum_{h \in H} a_{ht} vh) \otimes t.$$

If $v_1, \dots, v_n$ is a basis of V, one can see that the $v_i \otimes t$ are a basis for W, for $i = 1, \dots, n$ and $t \in T$.

6.9.8. Remark. This is related to a map called *transfer*, see [**Rob96**, Ch. 10] or [**Ser16**, Ch. 7] or [**Hup67**] under *Verlagerun*.

What is the character ψ of the G-representation W, in terms of the character χ of the H-representation V? If $t \in T$, and $g \in G$, we will have

$$tg = ht'$$

for some unique $h \in H$ and $t' \in T$. Therefore for $i = 1, \dots, n$ we will have

$$(v_i \otimes t)g = v_i \otimes (tg) = v_i \otimes (ht') = (v_i h) \otimes t'.$$

A non-zero diagonal coefficient can thus occur only if $t' = t$, that is $tg = ht$, or $g = t^{-1}ht \in H^t$. And then, such a contribution would be a_{ii}, if

$$v_i h = \sum_{i=1}^{n} a_{ij} v_j.$$

Since $\sum_{i=1}^{n} a_{ii}$ is the value on h of the character of the representation determined by the $\mathbf{C}[H]$-module V, we obtain the above formula for the induced character.

6.10. Characters of permutation representations

We have seen that if the finite group G acts on the finite set $\Omega = \{ 1, \dots, n \}$, this induces a linear representation of G on a vector space of basis $v_1, \dots, v_n$. What is its characters χ? There is a contribution "1" to $\chi(g)$ from every i such that $i^{g^{-1}} = i$, that is,

6.10.1. Lemma. *$\chi(g) = F(g)$.*

Here $F(g)$ is the number of fixed points of g acting on Ω.

6.10.2. Lemma. *The number of orbits of a permutation representation is*

$$\frac{1}{|\, G \,|} \sum_{g \in G} F(g).$$

Proof. Consider the set $\Omega \times G$, and its subset

$$\Delta = \{ (\alpha, g) : \alpha^g = \alpha \} .$$

In a typical *double counting* argument, we can count by rows, that is

$$|\, \Delta \,| = \sum_{\alpha \in \Omega} |\, \{ g \in G : \alpha^g = \alpha \} \,| = \sum_{\alpha \in \Omega} |\, G_\alpha \,| , \tag{6.10.1}$$

and we can count by column

$$|\, \Delta \,| = \sum_{g \in G} |\, \{ \alpha \in \Omega : \alpha^g = \alpha \} \,| = \sum_{g \in G} F(g). \tag{6.10.2}$$

For each orbit α^G we have, by orbit-stabiliser

$$\sum_{\beta \in \alpha^G} |\, G_\beta \,| = \sum_{\beta \in \alpha^G} \frac{|\, G \,|}{|\, \alpha^G \,|} = |\, G \,| .$$

Therefore (6.10.1) yields that $|\, \Delta \,|$ is $|\, G \,|$ times the number of orbits. Comparing to (6.10.2), we get the formula. □

From the two lemmas we get

6.10.3. Proposition. *Let χ be the character of a permutation representation.*

(1) $(1, \chi)$ is the number of orbits of G.
(2) G acts transitively if and only if $(1, \chi) = 1$

If G acts transitively, then we have $\chi = 1 + \psi$, for a character ψ not having 1 as a constituent.

If G acts on the set Ω, then it acts on $\Omega^2 = \Omega \times \Omega$ by $(\alpha, \beta)^g = (\alpha^g, \beta^g)$. Since $(\alpha, \beta)^g = (\alpha, \beta)$ if and only if $\alpha^g = \alpha$ and $\beta^g = \beta$, we will have that the permutation character for the action on Ω^2 will be χ^2.

6.10.4. Proposition. *Let G act transitively on Ω, with $|\, \Omega \,| > 1$, let χ be the corresponding permutation character, and $\chi = 1 + \psi$.*

The following are equivalent

(1) G acts 2-transitively on Ω,
(2) G has two orbits on Ω^2, and

(3) $\psi \in \operatorname{Irr}(G)$.

Proof of Proposition 6.10.4 . Since $|\,\Omega\,| > 1$, G has at least two orbits on Ω^2, namely

$$\{\, (\alpha, \alpha) : \alpha \in \Omega \,\} \qquad \text{and} \qquad \{\, (\alpha, \beta) : \alpha, \beta \in \Omega, \alpha \neq \beta \,\},$$

The first set is a an orbit, as G is transitive, and the second one will be an orbit precisely when G is 2-transitive.

Since the character of the action on Ω^2 is χ^2, G acts 2-transitively on Ω if and only if

$$2 = (1, \chi^2) = (\chi, \chi) = 1 + (\psi, \psi),$$

where we have used the facts that $(1, \psi) = 0$, and that χ has integer, and thus real, values. □

6.11. Character tables of small groups

6.11.1. S_3**.**

#	1	2	3
	1	(123)	(12)
	1	1	1
	1	1	-1
χ	2	-1	0

The first two characters are the linear ones from $S_3/A_3 \cong C_2$.

The non-linear character χ can be deduced from the orthogonality relations, or from the standard permutation character.

It is also easy to compute the representation ρ corresponding to χ. The eigenvalues of $\rho((123))$ are of the form ω^j, where ω is a primitve 3-rd root of unity. Since (123) is conjugate to its inverse, if ω^j is an eigenvalue, so is ω^{-j}. If the eigenvalues are both 1, then we should have $\chi((123)) = 2$, which is not the case. Then the eigenvalues are ω, ω^{-1} (and in fact $\omega + \omega^{-1} = -1$. Now since $(123)^{(12)} = (123)^{-1}$, if $v\rho((123)) = \omega v$, we have (omitting the ρ, that is, thinking in terms of modules)

$$v(123)(12) = (v(12))\omega = (v(12))(123)^{-1}),$$

so that $(v(12))(123) = (v(12))\omega^{-1}$. It follows that (12) exchanges the eigenspaces of (123) relative to the two eigenvalues, so that

$$\rho((123)) = \begin{bmatrix} \omega & \\ & \omega^{-1} \end{bmatrix}, \qquad \rho((12)) = \begin{bmatrix} & \lambda \\ \lambda^{-1} & \end{bmatrix},$$

for some $\lambda \neq 0$, the λ^{-1} coming from the fact that $\rho((12))^2 = 1$.

6.11.2. A_4**.**

#	1	3	4	4
	1	(12)(34)	(123)	(132)
	1	1	1	1
	1	1	ω	ω^{-1}
	1	1	ω^{-1}	ω
χ	3	-1	0	0

The first three characters are the linear ones from $A_4/V \cong C_3$, where $V = \{\,1, (12)(34), (13)(24), (14)(23)\,\}$.

The non-linear character χ can be deduced from the orthogonality relations, or from the standard permutation character. To compute the associated representation ρ, note that we must have

$$\rho((12)(34)) = \begin{bmatrix} 1 & & \\ & -1 & \\ & & -1 \end{bmatrix}, \rho((14)(23)) = \begin{bmatrix} -1 & & \\ & -1 & \\ & & 1 \end{bmatrix}, \rho((13)(24)) = \begin{bmatrix} -1 & & \\ & 1 & \\ & & -1 \end{bmatrix},$$

with respect to a suitable basis. Since $\rho((123))$ must permute the three cyclically, one can take

$$\rho((123)) = \begin{bmatrix} & & 1 \\ 1 & & \\ & 1 & \end{bmatrix}.$$

6.11.3. S_4.

#	1	3	6	8	6
	1	(12)(34)	(12)	(123)	(1234)
	1	1	1	1	1
	1	1	-1	1	-1
	2	2	0	-1	0
	3	-1	1	0	-1
	3	-1	-1	0	1

The first two characters are the linear ones from $S_4/A_4 \cong C_2$.

The first three characters come from those of $S_3 \cong S_4/V$.

The fourth character come from the standard permutation character.

The last one is the previous one times the *sign character*, that is, the second one.

6.11.4. S_5. (Here we mainly follow [**Bla19**].)

First, linear characters and standard permutation representation.

#	1	15	10	20	20	30	24
	1	(12)(34)	(12)	(123)	(123)(45)	(1234)	(12345)
	1	1	1	1	1	1	1
	1	1	-1	1	-1	-1	1
	4	0	2	1	-1	0	-1
	4	0	-2	1	1	0	-1

Now some congruences show that the next characters must have

#	1	15	10	20	20	30	24
	1	(12)(34)	(12)	(123)	(123)(45)	(1234)	(12345)
	1	1	1	1	1	1	1
sign	1	1	-1	1	-1	-1	1
	4	0	2	1	-1	0	-1
	4	0	-2	1	1	0	-1
	5						
	5						
	6		0		0	0	

withe the second 5 being the previous one times sign, and the zeroes of the 6 being due to the same reason.

Let ρ be an irreducible representation of degree 5, and χ its character.

Consider the eigenvalues of $\rho((12345))$. Since all 5-cycles are conjugate, either these eigenvalues are all 1 (but then in $120 = \lvert S_5 \rvert = \sum_{g \in G} \overline{\chi(g)}\chi(g)$ this would contribute $24 \cdot \chi((12345))^2 = 24 \cdot 25$, which is too much), or they must be $1, \omega, \omega^2, \omega^3, \omega^4$, where ω is a primitive 5-th root of 1, so that $\chi((12345)) = 0$.

If we consider a 4-cycle, one sees the eigenvalues $\pm i$ cancel out in pairs, so we are left with an odd number of ± 1. If $\chi((1234)) = \pm 3$, as above we have a contribution of $20 \cdot 9$. So $\chi((1234)) = \pm 1$.

An analogue argument with the 3-cycles, keeping in mind the numbers we already got, shows that the eigenvalues must be $1, \varphi, \varphi^2, \varphi, \varphi^2$, where φ is a primitive 3-rd root of unity.

	1	15	10	20	20	30	24
	1	(12)(34)	(12)	(123)	(123)(45)	(1234)	(12345)
	1	1	1	1	1	1	1
sign	1	1	-1	1	-1	-1	1
	4	0	2	1	-1	0	-1
	4	0	-2	1	1	0	-1
	5			-1		1	0
	5			-1		-1	0
	6		0		0	0	

Now (23) inverts (123), and thus exchanges the ψ and ψ^2 eigenspaces, so it must have a ± 1 on the one-dimensional eigenspace relative to the eigenvalue 1, and two $\left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]$ blocks, so that $\chi((23)) = \pm 1$. Since (45) commutes with (123), we see that the values on (123)(45) are the same.

	1	15	10	20	20	30	24
	1	(12)(34)	(12)	(123)	(123)(45)	(1234)	(12345)
	1	1	1	1	1	1	1
sign	1	1	-1	1	-1	-1	1
	4	0	2	1	-1	0	-1
	4	0	-2	1	1	0	-1
	5		1	-1	1	1	0
	5		-1	-1	-1	-1	0
	6		0		0	0	

Orthogonality then yields $\chi((12)(34)) = 1$.

Now we may get 6 simply by subtracting from the regular representation. For (12)(34) the value of the character is

$$\frac{1}{6}(-1-1-5-5) = -2.$$

For (123) the value of the character is

$$\frac{1}{6}(-1-1-4-4+5+5) = 0.$$

For (123)(45) the value of the character is

$$\frac{1}{6}(-1+1-4+4+5-5) = 0.$$

For (12345) the value of the character is

$$\frac{1}{6}(-1-1+4+4) = 1.$$

We get

	1	15	10	20	20	30	24
	1	(12)(34)	(12)	(123)	(123)(45)	(1234)	(12345)
	1	1	1	1	1	1	1
sign	1	1	-1	1	-1	-1	1
	4	0	2	1	-1	0	-1
	4	0	-2	1	1	0	-1
	5	1	1	-1	1	1	0
	5	1	-1	-1	-1	-1	0
	6	-2	0	0	0	0	1

6.11.5. A_5**.** Let us start from S_5, where we get

1	15	20	12	12
1	(12)(34)	(123)	(12345)	(13524)
1	1	1	1	1
4	0	1	-1	-1
5	1	-1	0	0
6	-2	0	1	1

For the last character χ we have $(\chi, \chi) = 2$, the other remaining irreducible. Thus χ splits as the sum of two irreducible characters of degree 3 each (check sum of squares of degrees).

Note here that (12345) is conjugate to its inverse, but not to its square. So the eigenvalues (which cannot be all 1, lest overcounting) must be $1, \omega, \omega^{-1}$ for (12345), and $1, \omega^2, \omega^{-2}$ for (13524).

A similar argument gets the value zero on (123), and then orthogonality with the trivial character does it for (12)(34).

1	15	20	12	12
1	(12)(34)	(123)	(12345)	(13524)
1	1	1	1	1
4	0	1	-1	-1
5	1	-1	0	0
3	-1	0	$-\omega - \omega^{-1}$	$-\omega^2 - \omega^{-2}$
3	-1	0	$-\omega^2 - \omega^{-2}$	$-\omega - \omega^{-1}$

6.12. A non-linear character vanishes somewhere

Taken from [**Isa06**, p. 40].

6.12.1. Lemma. *Let G be a cyclic group of order n, and write $S \subseteq G$ for the set of elements of G of order n.*

Let χ be a character of G such that $\chi(s) \neq 0$ for all $s \in S$.

Then

$$\sum_{s \in S} |\, \chi(s) \,|^2 \geq |\, S \,|.$$

Proof. Let $E/\mathbf{Q}$ be the splitting field of $x^n - 1$, and $H = \mathrm{Gal}(E/\mathbf{Q})$ its Galois group. An element $h \in H$ takes an n-th root of unity ω to a power ω^m, for some m coprime to n. Since

$$\chi(s) = \omega_1 + \dots + \omega_t$$

for some n-th roots of unity ω_i, we will have

$$\chi(s)^h = \omega_1^m + \dots + \omega_t^m.$$

For m coprime to n, the map $x \to x^m$ is a permutation of G (actually an isomorphism), and in particular a bijection on S. If ρ affords χ, we have for $s \in S$

$$\rho(s^m) = \rho(s)^m = \begin{bmatrix} \omega_1 & & & \\ & \omega_2 & & \\ & & \ddots & \\ & & & \omega_t \end{bmatrix}^m = \begin{bmatrix} \omega_1^m & & & \\ & \omega_2^m & & \\ & & \ddots & \\ & & & \omega_t^m \end{bmatrix},$$

so that

$$\chi(s)^h = \chi(s^m). \tag{6.12.1}$$

Now H is an abelian group, isomorphic to $(\mathbf{Z}/n\mathbf{Z})^*$. By Lemma 1.11.10, we have for $\alpha \in E$ and $h \in H$

$$(|\, \alpha \,|^2)^h = \left|\, \alpha^h \,\right|^2. \tag{6.12.2}$$

We have just seen that for m coprime to n, the map $x \mapsto x^m$ is a permutation of S. It follows that $\prod_{s \in S} |\, \chi(s) \,|^2$ is invariant under H by (6.12.1) and (6.12.2), and thus rational. Since it is an algebraic integer (see Proposition 1.11.7 and Proposition 6.13.1 below), it is an integer. Since χ does not vanish on S, we have

$$\prod_{s \in S} |\, \chi(s) \,|^2 \geq 1.$$

By the arithmetic/geometric means inequality, we obtain

$$\frac{1}{|\, S \,|} \sum_{s \in S} |\, \chi(s) \,|^2 \geq (\prod_{s \in S} |\, \chi(s) \,|^2)^{1/|\, S \,|} \geq 1.$$

□

6.12.2. Theorem (Burnside). *Let G be a finite group, and χ be an irreducible, non-linear character of G.*

Then there is $g \in G$ such that $\chi(g) = 0$.

Proof. Suppose the irreducible character χ satisfies $\chi(g) \neq 0$ for all $g \in G$.

Consider the equivalence relation on G given by

$$aRb \quad \text{iff} \quad \langle\, a \,\rangle = \langle\, b \,\rangle\,.$$

The equivalence class S of an element a, of some order n, is thus given by the set of elements of order n of the cyclic group $\langle\, a \,\rangle$.

Lemma (6.12.1) yields that for each such class S we have $\sum_{s \in S} |\, \chi(s) \,|^2 \geq |\, S \,|$. Summing over all equivalence classes of non-identity elements, we get

$$\sum_{1 \neq g \in G} |\, \chi(g) \,|^2 \geq |\, G \,| - 1.$$

Therefore

$$|\, G \,| = |\, G \,| \, (\chi, \chi) = \sum_{g \in G} |\, \chi(g) \,|^2 \geq |\, G \,| - 1 + \chi(1)^2,$$

which yields $\chi(1) \leq 1$, so that χ is linear. □

6.13. Integrality

If ρ is a linear representation of G, and $g \in G$ has order n, then

$$1 = \rho(1) = \rho(g^n) = \rho(g)^n.$$

It follows that $\rho(g)$ is a root of $x^n - 1$, so that the minimal polynomial of $\rho(g)$ is a divisor of $x^n - 1$, and thus the eigenvalues of $\rho(g)$ are n-th roots of unity. We obtain

6.13.1. Proposition. *Character values, as sum of roots of unity, are algebraic integers.*

6.13.2. Exercise.

(1) *Show that is ρ is the right regular representation, and $g \in G$ has order n, then the minimal polynomial of $\rho(g)$ is $x^n - 1$.*

(2) Show that the above need not hold for an arbitrary representation. Avoid the trivial case when $\rho(g) = 1$, and try and find an example in which $|\, g \,| = |\, \rho(g) \,|$.

6.13.3. Theorem. *Let $f : G \to \mathbf{C}$ be a class function on the group G, whose values are algebraic integers, so that $\alpha = \sum_{g \in G} f(g) g \in Z(\mathbf{C}[G])$.*

(1) α is integral.

(2) If χ is an irreducible character of G, then

$$\frac{1}{\chi(1)} \sum_{g \in G} f(g)\chi(g)$$

is an algebraic integer.

Proof. Let S be the subset of the commutative ring $Z(\mathbf{C}[G])$ consisting of the $\sum_{g \in G} f(g) g$, where f is a class function with integer values. The elements $\sum C = \sum_{g \in \mathcal{C}} g$, for $\mathcal{C}$ a conjugacy class, are a basis of S as a $\mathbf{Z}$-module.

We claim that S is a subring of $Z(\mathbf{C}[G])$. In fact, let C, D be two conjugacy classes. Since $Z(\mathbf{C}[G])$ is a subring of $\mathbf{C}[G]$, we will have

$$(\sum C) \cdot (\sum D) = \sum_{E} \lambda(C, D, E) E,$$

where E ranges over the set of conjugacy classes, and $\lambda_E \in \mathbf{C}$. But $\lambda(C, D, E)$ counts how many times a fixed element $e \in E$ occurs as a product cd, for $c \in C$ and $d \in D$, and this is a (non-negative) integer. (See the example following this proof.)

Since the ring S is finitely generated as a $\mathbf{Z}$-module, all of its elements are integral.

Consider now the subring T of $Z(\mathbf{C}[G])$ consisting of the $\alpha = \sum_{g \in G} f(g) g$, where f is a class function with algebraic integer values. Since sums and product of integral elements are integral, it follows that α is integral.

Applying r_χ to α, and appealing to Theorem 6.3.2, we obtain the second part, since clearly the image under a ring morphism of an integral element is integral, and thus that number is an algebraic integer. □

6.13.4. Example. Let $G = S_3$, $C = \{\, (12), (13), (23) \,\}$, $D = \{\, (123), (132) \,\}$. We have the products

	(123)	(132)
(12)	(13)	(23)
(13)	(23)	(12)
(23)	(12)	(13)

Thus

$$(\sum C)(\sum D) = 2 \sum C.$$

6.13.5. Theorem (Burnside). *If $C = x^G$, $D = y^G$ and $E = z^G$, then*

$$\lambda(C, D, E) = \frac{|\, G \,|}{|\, C_G(x) \,| \cdot |\, C_G(y) \,|} \sum_{\chi \in \mathrm{Irr}(G)} \frac{\chi(x)\chi(y)\chi(z^{-1})}{\chi(1)}.$$

Proof. Recall that in the isomorphism

$$\mathbf{C}[G] \to \sum_{\chi \in \mathrm{Irr}(G)} M_{\chi(1)}(\mathbf{C})$$

an element $g \in G$ goes in the element

$$(\rho(g))_\chi$$

whose χ-th component is just $\rho(g)$, where ρ affords χ.

So if $g^G = \{ g^{x_1}, \dots, g^{x_n} \}$, with the g^{x_i} distinct, we have

$$\sum g^G = g^{x_1} + \dots + g^{x^n} \mapsto (\rho(g^{x_1}) + \dots + \rho(g^{x^n}))_\chi$$

In an argument that we have already seen in Theorem 6.3.2, $\rho(g^{x_1}) + \dots + \rho(g^{x^n})$ is a scalar matrix aI (where $a \in \mathbf{C}$, and I is a suitable identity matrix), so that

$$\left| g^G \right| \chi(g) = \mathrm{trace}(\rho(g^{x_1}) + \dots + \rho(g^{x^n})) = \mathrm{trace}(aI) = a\chi(1),$$

as the trace is a class function.

So if we now consider the isomorphism of rings

$$\mathbf{Z}(\mathbf{C}[G]) \to \sum_{\chi \in \mathrm{Irr}(G)} \mathbf{C},$$

we have

$$\sum g^G \mapsto \left(\frac{\left| g^G \right| \chi(g)}{\chi(1)} \right)_\chi$$

Thus

$$(\sum g^G)(\sum h^G) \mapsto \left(\frac{\left| g^G \right| \chi(g) \left| h^G \right| \chi(h)}{\chi(1)^2} \right)_\chi$$

We want to find the integers $\lambda(g, h, w)$ such

$$\left(\frac{\left| g^G \right| \chi(g) \left| h^G \right| \chi(h)}{\chi(1)^2} \right)_\chi = \sum_w \lambda(g, h, w) \left(\frac{\left| w^G \right| \chi(w)}{\chi(1)} \right)_\chi, \tag{6.13.1}$$

where w ranges over a set of representatives of the conjugacy classes of G.

Multiply both sides of (6.13.1) componentwise by $\chi(z^{-1})\chi(1)$, for a fixed z, and sum over $\chi \in \mathrm{Irr}(G)$. By Lemma 6.13.7 below, the right-hand side becomes

$$\sum_w \left| w^G \right| \lambda(g, h, w) \sum_{\chi \in \mathrm{Irr}(G)} \chi(z^{-1})\chi(w) = \lambda(g, h, z) \left| z^G \right| \cdot \left| C_G(z) \right| = \lambda(g, h, z) \left| G \right|,$$

where we have used orbit-stabiliser, so that we obtain

$$\sum_{\chi \in \mathrm{Irr}(G)} \frac{\left| g^G \right| \chi(g) \left| h^G \right| \chi(h)\chi(z^{-1})}{\chi(1)} = \lambda(g, h, z) \left| G \right|,$$

which yields the claim, keeping in mind that by orbit-stabiliser

$$\frac{\left| g^G \right| \left| h^G \right|}{\left| G \right|} = \left| G \right| \frac{\left| g^G \right|}{\left| G \right|} \frac{\left| h^G \right|}{\left| G \right|} = \frac{\left| G \right|}{\left| C_G(g) \right| \left| C_G(h) \right|}.$$

□

6.13.6. Example (Example 6.13.4 revisited). We compute $(\sum C)(\sum D)$ using Theorem 6.13.5 and the table of Subsection 6.11.1.

$$\lambda(C, D, \{\,1\,\}) = \frac{6}{3 \cdot 2} \sum_{\chi} \frac{\chi(12)\chi(123)\chi(1)}{\chi(1)} = 1 - 1 + 0 \cdot (-1) = 0.$$

$$\lambda(C, D, C) = \sum_{\chi} \frac{\chi(12)\chi(123)\chi(12)}{\chi(1)} = 1 + (-1)^2 + \frac{0 \cdot (-1) \cdot 0}{2} = 2.$$

$$\lambda(C, D, D) = \sum_{\chi} \frac{\chi(12)\chi(123)\chi(123)}{\chi(1)} = 1 + (-1) \cdot 1 \cdot 1 + \frac{0 \cdot (-1) \cdot (-1)}{2} = 0.$$

6.13.7. Lemma (The other orthogonality relations). *For $z, w \in G$*

$$\sum_{\chi \in \operatorname{Irr}(G)} \chi(z^{-1})\chi(w) = \begin{cases} |\, C_G(z) \,| & \textit{if } z \textit{ and } w \textit{ are conjugate,} \\ 0 & \textit{otherwise.} \end{cases}$$

Proof. Consider the extended character table Y, whose (χ, g) entry is $\left| \, g^G \, \right|^{1/2} \chi(g)$. The orthogonality relations yield $\overline{Y} Y^t = |\, G \,|\, I$, for a suitable identity matrix I. Therefore $Y^t \overline{Y} = |\, G \,|\, I$, and this means that for $z, w \in G$

$$\sum_{\chi \in \operatorname{Irr}(G)} \left| \, z^G \, \right|^{1/2} \chi(z^{-1}) \left| \, w^G \, \right|^{1/2} \chi(w) = \begin{cases} |\, G \,| & \text{if } z \text{ and } w \text{ are conjugate,} \\ 0 & \text{otherwise.} \end{cases}$$

When z and w are conjugate, we have thus

$$\sum_{\chi \in \operatorname{Irr}(G)} \chi(z^{-1})\chi(z) = \frac{|\, G \,|}{|\, z^G \,|} = |\, C_G(z) \,| \,.$$

When z and w are not conjugate, we have $\sum_{\chi \in \operatorname{Irr}(G)} \chi(z^{-1})\chi(w) = 0$. □

Note the following consequence of Lemma 6.13.7, which states that the size of a centraliser does not increase when going from a group to a quotient group.

6.13.8. Proposition. *Let G be a finite group, $N \trianglelefteq G$, $x \in G$. Then*

$$|\, C_G(x) \,| \geq \left| \, C_{G/N}(xN) \, \right|.$$

Proof via characters. We know that every (irreducible) representation of G/N corresponds to a(n irreducible) representation of G with N in its kernel. Witn a slight abuse of notation, then

$$\left| \, C_{G/N}(xN) \, \right| = \sum_{\chi \in \operatorname{Irr}(G/N)} |\, \chi(xN) \,|^2 \leq \sum_{\chi \in \operatorname{Irr}(G)} |\, \chi(x) \,|^2 = |\, C_G(x) \,| \,.$$

□

Direct proof. By orbit-stabiliser

$$\lvert C_G(x) \rvert = \frac{\lvert G \rvert}{\lvert x^G \rvert}, \quad \left\lvert C_{G/N}(xN) \right\rvert = \frac{\lvert G/N \rvert}{\lvert (xN)^{G/N} \rvert} = \frac{\lvert G \rvert}{\lvert N \rvert \cdot \lvert (xN)^{G/N} \rvert}.$$

We have thus to prove that

$$\left\lvert x^G \right\rvert \leq \lvert N \rvert \cdot \left\lvert (xN)^{G/N} \right\rvert. \tag{6.13.2}$$

We have

$$(xN)^{G/N} = \left\{ (xN)^{gN} : g \in G \right\} = \{ x^g N : g \in G \}. \tag{6.13.3}$$

Consider the union of the $\left\lvert (xN)^{G/N} \right\rvert$ cosets in the left-hand side of (6.13.3). This union has $\lvert N \rvert \cdot \left\lvert (xN)^{G/N} \right\rvert$ elements. Now the union of the cosets in the right-hand side of equation (6.13.3) is $x^G N \supseteq x^G$. We obtain (6.13.2). □

Coming back to our mainline, we have

6.13.9. Corollary (to Theorem 6.13.3). *If χ is an irreducible character, then $\chi(1)$ divides the order of G.*

Proof. In part (2) of Theorem 6.13.3, take $f = \overline{\chi}$ to get that $\lvert G \rvert / \chi(1)$ is a rational number which is an algebraic integer, and thus it is an integer. □

6.13.10. Proposition. *Let ρ be an irreducible representation, and χ be its character. Let $g \in G$. We have*

(1)

$$\frac{\left\lvert g^G \right\rvert \cdot \chi(g)}{\chi(1)}$$

is an algebraic integer.

Suppose now $\gcd(\left\lvert g^G \right\rvert, \chi(1)) = 1$. *Then*

(2) $\chi(g)/\chi(1)$ is an algebraic integer.

(3) If $\chi(g) \neq 0$, then $\rho(g)$ is a scalar matrix.

Proof. Let $f : G \to \mathbf{C}$ be the class function that is zero everywhere, except on the conjugacy class of g, where its value is 1, so that

$$\sum_{x \in G} f(x)\chi(x) = \left\lvert g^G \right\rvert \chi(g).$$

By Theorem 6.13.3(2), $\left\lvert g^G \right\rvert \cdot \chi(g)/\chi(1)$ is an algebraic integer.

If $\gcd(\left\lvert g^G \right\rvert, \chi(1)) = 1$, by Bézout, there are $a, b \in \mathbf{Z}$ such that $a \left\lvert g^G \right\rvert + b\chi(1) = 1$, so that

$$\frac{\chi(g)}{\chi(1)} = a \left\lvert g^G \right\rvert \cdot \frac{\chi(g)}{\chi(1)} + b\chi(g).$$

Since both $\left\lvert g^G \right\rvert \cdot \chi(g)/\chi(1)$ and $\chi(g)$ are algebraic integers, so is $\chi(g)/\chi(1)$.

$\chi(g)/\chi(1)$ is of the form of (1.11.1) of Lemma 1.11.13, therefore $\rho(g)$ is a scalar. □

CHAPTER 7

Applications

7.1. Burnside $p^a q^b$

In this section we report the celebrated theorem of Burnside [**Bur04**], that shows that finite group of order divisible of at most two primes are soluble. (A_5 is non-abelian simple, of order $2^2 \cdot 3 \cdot 5$.)

Burnside's proof makes use of characters, whose theory he had contributed to developing. Burnside published his proof in 1904. One had to wait until the 1970's for proofs not using characters [**Gol70, Ben72, Mat73**].

7.1.1. Lemma. *Let G be a finite group, $1 \ne g \in G$ such that its conjugacy class has order a power of a prime p.*

Then there is a proper normal subgroup N of G such that $gN \in Z(G/N)$.

Proof. From Lemma 6.13.7 we have

$$\sum_{\chi \in \operatorname{Irr}(G)} \chi(1)\chi(g) = 0,$$

and thus

$$\sum_{1 \ne \chi \in \operatorname{Irr}(G)} \frac{\chi(1)\chi(g)}{p} = -\frac{1}{p}.$$

Since the right-hand side is not an algebraic integer, there is $1 \ne \chi \in \operatorname{Irr}(G)$ such that $\chi(1)\chi(g)/p$ is not an algebraic integer, so that $\chi(g) \ne 0$, and $p \nmid \chi(1)$. Since $\left| g^G \right|$ is a power of p, it follows that $\gcd(\left| g^G \right|, \chi(1)) = 1$. Now Proposition 6.13.10(3) yields that $\rho(g)$ is scalar, where ρ is a representation that affords χ. Since $\chi \ne 1$, we have that $\ker(\rho)$ is a proper normal subgroup of G, and the first isomorphism theorem yields $G/\ker(\rho) \cong \rho(G)$. Since $\rho(g)$ is scalar, it is in the centre of $\rho(G)$, and thus gN is in the centre of G/N. □

7.1.2. Theorem (Burnside $p^a q^b$). *A group of order $p^a q^b$, where p and q are primes, is soluble.*

Proof. Note first of all that groups of order the power of a prime are nilpotent (and thus soluble), as follows from the class equation. Therefore we may assume that both p and q divide the order of G.

Considering a composition series of G, we see that we may assume G to be non-abelian simple. Were $\{ 1 \}$ the only conjugacy class of order not divisible by q, then the order of G would congruent to 1 modulo q, a contradiction to the fact that q divides the order of G.

So there is an element $g \in G$ whose conjugacy class g^G has order not divisibile by q and thus, by orbit-stabiliser, g^G has order a power of p. By Lemma 7.1.1, $g \in Z(G)$, a contradiction. □

7.2. Frobenius groups and the Frobenius kernel

A *Frobenius group* is a finite group which acts transitively on a finite set Ω, such that the stabilisers are non-trivial (so the group does not act regularly) and pairwise disjoint, that is, for $\alpha, \beta \in \Omega$, with $\alpha \neq \beta$, we have $G_\alpha \cap G_\beta = \{ 1 \}$. A typical example is S_3.

Note that if $g \in G \setminus G_\alpha$, then $\alpha^g \neq \alpha$, and thus $G_\alpha \cap G_{\alpha^g} = \{ 1 \}$. Since $G_{\alpha^g} = G_\alpha^g$, we obtain in particular, $N_G(G_\alpha) = G_\alpha$. More strongly, if $h \in G_\alpha$ and $g \in G$ is such that $h^g \in G_\alpha$, then $h^g \in G_\alpha \cap G_{\alpha^g}$, so that $\alpha = \alpha^g$ and $g \in G_\alpha$.

Since G is transitive, the stabilisers are all conjugate. An abstract characterisation of Frobenius groups is thus as the finite groups G which have a subgroup $H \neq \{ 1 \}$ such that $H \cap H^g = \{ 1 \}$ for $g \in G \setminus H$. In fact, letting G act on the cosets Hg of H in G by right multiplication, we have that the action is transitive, and the stabiliser of Hg is H^g.

The subgroup H is called a *Frobenius complement.*

For the following theorem, no character-free proof is known.

7.2.1. Theorem. *Let G be a Frobenius group with respect to the subgroup H.*

Then $N = \{ 1 \} \cup (G \setminus \bigcup_{g \in G} H^g)$ is a normal subgroup of G, so that G is the semidirect product of N by H.

The subgroup N is called the *Frobenius kernel.* It is clear that N is a normal *set.* The point is proving that it is a subgroup.

Note that

$$| N | = | G | - (| G : H | \cdot (| H | - 1)) = | G | - | G | + | G : H | = | G : H | ,$$

so that $G = HN$, and N is a transitive subgroup.

Proof 1, from [**Isa06**]. The idea of this proof is the following. If N exists, then every character of H extends to a character of G which has N in its kernel. Since for $h \in H$ the fact that $\chi(h) = \chi(1)$ for all $\chi \in \mathrm{Irr}(H)$ implies $h = 1$, we obtain that the intersection of all these kernels is exactly N. Now our goal is exactly to prove the existence of N. This we will do by extending every irreducible character of H to a character of G, and obtaining N as the intersection of the kernels of all these extensions. Details below.

Let ϑ be a class function on H such that $\vartheta(1) = 0$. We claim that

$$(\vartheta^G)_H = \vartheta. \tag{7.2.1}$$

Let $h \in H$, $h \neq 1$. Then

$$\vartheta^G(h) = \frac{1}{| H |} \sum_{x \in G} \vartheta^\circ(xhx^{-1}).$$

If $\vartheta^{\circ}(xhx^{-1}) \neq 0$ for some x, then $1 \neq xhx^{-1} \in H \cap H^{x^{-1}}$, so that $x \in H$, and $\vartheta^{\circ}(xhx^{-1}) = \vartheta(h)$, as ϑ is a class function on H. Therefore

$$\vartheta^G(h) = \frac{1}{|\,H\,|} \sum_{x \in H} \vartheta(h) = \vartheta(h).$$

We have then

$$\vartheta^G(1) = \frac{|\,G\,|}{|\,H\,|} \vartheta(1) = 0,$$

so that (7.2.1) holds.

Let now $1 \neq \varphi \in \operatorname{Irr}(H)$, and write $\vartheta = \varphi - \varphi(1) 1_H$, so that ϑ is a class function on H with $\vartheta(1) = 0$.

By Frobenius reciprocity and (7.2.1), we have

$$(\vartheta^G, \vartheta^G)_G = (\vartheta, (\vartheta^G)_H)_H = (\vartheta, \vartheta)_H,$$

from which it follows that $(\vartheta^G, \vartheta^G)_G = 1 + \varphi(1)^2$.

Now $(\vartheta^G, 1_G)_G = (\vartheta, 1_H)_H = -\varphi(1)$, so that $\vartheta^G = \varphi^* - \varphi(1) 1_G$, where φ^* is a class function on G such that $(\varphi^*, 1_G) = 0$. Since $1 + \varphi(1)^2 = (\vartheta^G, \vartheta^G)_G = (\varphi^*, \varphi^*)_G + \varphi(1)^2$, we get $(\varphi^*, \varphi^*)_G = 1$.

ϑ is a difference of characters, so that ϑ^G also is, and thus so is $\varphi^* = \vartheta^G + \varphi(1) 1_G$. Writing φ^* as a linear combination with integer coefficients of irreducible characters, we see that $\pm\varphi^* \in \operatorname{Irr}(G)$. But since for $h \in H$ one has

$$\varphi^*(h) = \vartheta^G(h) + \varphi(1) = \vartheta(h) + \varphi(1) = \varphi(h),$$

we have $\varphi^*(1) > 0$, so that $\varphi^* \in \operatorname{Irr}(G)$.

A brief aside on kernels, which reprises an argument at the beginning of Section 6.6. Let σ be a representation of a group G, and τ its character. Write

$$\ker(\tau) = \{\, g \in G : \tau(g) = \tau(1) \,\}.$$

Clearly $\ker(\sigma) \subseteq \ker(\tau)$. Conversely, if $g \in \ker(\tau)$, by Lemma 1.11.12, we have that $\sigma(g) = \sigma(1)$, so that $g \in \ker(\sigma)$.

For every $1_H \neq \varphi \in \operatorname{Irr}(H)$ we have obtained an extension $\varphi^* \in \operatorname{Irr}(G)$. Consider the intersection of all of their kernels

$$M = \bigcap_{\varphi} \ker(\varphi^*).$$

If $x \in M \cap H$, then $\varphi(x) = \varphi^*(x) = \varphi^*(1) = \varphi(1)$ for all $\varphi \in \operatorname{Irr}(H)$, so that $x = 1$.

Now note that if M is a normal subgroup of G such that $M \cap H = 1$, then $M \cap H^x = 1$ for all x, and thus $M \subseteq N$.

Conversely, if $1 \neq g \in N$, then

$$\varphi^*(g) - \varphi^*(1) = \varphi^*(g) - \varphi(1) = \vartheta^G(g) = 0,$$

so that $g \in M$, and $N = M$ is a normal subgroup of G. □

Proof 2, from [**Ser16**]. Let ψ be any class function on H. Then there is a unique class function ψ' on G which extends ψ and is constant on N.

If a conjugacy class of G does intersect H trivially, then it is contained in N, whence the uniqueness.

As to existence,

$$\psi'(x) = \begin{cases} \psi(1) & \text{if } x \in N \\ \psi(h) & \text{if } x = h^g\text{, for some } h \in H \text{ and } g \in G. \end{cases}$$

Note that ψ' is well-defined. In fact if $h_1^{g_1} = h_2^{g_2}$, then $h_1^{g_1 g_2^{-1}} = h_2 \in H$, whence $g_1 g_2^{-1} \in H$, and h_1, h_2 are conjugate in H.

In the rest of the proof, we will use the scalar product

$$(\alpha, \beta)_G = \sum_{g \in G} \alpha(g) \beta(g^{-1})$$

on G, and the analogue on H.

Let now ϑ be a class function on G. We claim that

$$(\vartheta, \psi')_G = (\vartheta_H, \psi)_H + \psi(1)(\vartheta, 1)_G - \psi(1)(\vartheta_H, 1)_H. \tag{7.2.2}$$

Note this expression is linear in ψ. It holds true when $\psi \equiv k \equiv \psi'$ is an integer k, thus it suffices to consider the case when $\psi(1) = 0$, so that $\psi \equiv 0$ on N, when the expression becomes

$$(\vartheta, \psi')_G = (\vartheta_H, \psi)_H.$$

Let T be a left transversal of H in G (that is, a complete set of representatives of the left cosets of H in G), so that the H^t are the conjugates of H, for $t \in T$, and every conjugate different from 1 of an element of H can be written uniquely as h^t, for $h \in H$ and $t \in T$.

$$\begin{aligned} (\vartheta, \psi')_G &= \frac{1}{|\,G\,|} \sum_{g \in G} \vartheta(g) \psi'(g^{-1}) \\ &= \frac{1}{|\,G\,|} \sum_{(t,h) \in T \times H} \vartheta(h) \psi'(t^{-1} h^{-1} t) \\ &= \frac{|\,T\,|}{|\,G\,|} \sum_{h \in H} \vartheta(h) \psi'(h^{-1}) \\ &= (\vartheta_H, \psi)_H. \end{aligned}$$

Now we claim that for ψ_i class functions on H, we have

$$(\psi_1, \psi_2)_H = (\psi_1', \psi_2')_G,$$

that is, the map $\psi \mapsto \psi'$ is an isometry.

In fact, setting $\vartheta \equiv 1$ in (7.2.2) we get $(1, \psi')_G = (1, \psi)_H$. If we define $\psi^*(g) = \psi(g^{-1})$, we have

$$(\psi_1', \psi_2')_G = (\psi_1' \psi_2^{*\prime}, 1)_G = ((\psi_1 \psi_2^*)', 1)_G = (\psi_1 \psi_2^*, 1)_H = (\psi_1, \psi_2)_H.$$

Now we claim that if ψ is a character of H and ϑ is a character of G, then $(\psi', \vartheta)_G$ is an integer. This is because ϑ_H is a character of H, and thus every term of the right-hand side of (7.2.2) is an integer.

We now claim that if $\chi \in \operatorname{Irr}(H)$, then $\chi' \in \operatorname{Irr}(G)$. We have $\chi' = \sum_{\vartheta \in \operatorname{Irr}(G)} c_\vartheta \vartheta$, for some $c_\vartheta \in \mathbf{C}$. By the previous step, $c_\vartheta \in \mathbf{Z}$. Since $\sum_{\vartheta \in \operatorname{Irr}(G)} c_\vartheta^2 = (\chi', \chi')_G =$

$(\chi, \chi)_H = 1$, all c_i are zero except for one c_{ϑ_0}, which is ± 1. If $c_{\vartheta_0} = -1$, then $\chi' = -\vartheta_0$, contradicting $\chi'(1) = \chi(1) > 0$ and $\vartheta_0(1) > 0$. Thus $\chi' = \vartheta_0$.

It follows that is χ is a character of H, then χ' is a character of G.

Finally, let ρ be a faithful representation of H, such as the right regular one, and χ be its character. By the previous step, χ' is the character of a representation ρ' of G. If $g \in G$ is conjugate to an element different from 1 of H, then $\chi'(g) = \chi(h) \neq \chi(1) = \chi'(1)$, so that $\rho'(g) \neq 1$. If $g \in N$, then $\chi'(g) = \chi(1) = \chi'(1)$, so that $\rho'(g) = 1$. It follows that N coincides with $\ker(\rho')$, and thus it is a (normal) subgroup of G. □

7.3. Groups with an abelian Sylow p-subgroup

This is taken from [**Isa06**, p. 63].

7.3.1. THEOREM. *Let G be a finite group, p a prime dividing the order of G. Suppose a p-Sylow subgroup is abelian.*
Then $G' \cap Z(G)$ is not divisible by p.

PROOF. Suppose, by way of contradiction, that there is a subgroup $U \leq G' \cap Z(G)$ of order p, and let P be a Sylow p-subgroup cotaining U.

Let $\lambda \neq 1_U$ be an irreducible character of U. By Corollary 6.9.7, $\lambda = \mu_U$ for some irreducible character μ of P.

If

$$\mu^G = \sum a_\chi \chi, \tag{7.3.1}$$

for $\chi \in \operatorname{Irr}(G)$, then

$$\mu^G(1) = |\, G : P \,| \cdot \mu(1) = |\, G : P \,|$$

is coprime to p. It follows there is $\chi \in \operatorname{Irr}(G)$, which occurs in (7.3.1) with a nonzero coefficient a_χ, such that $p \nmid \chi(1)$. Thus $0 \neq a_\chi = (\mu^G, \chi) = (\mu, \chi_P)$, that is, μ is a consituent of χ_P, and thus $\lambda = \mu_U$ is a constituent of χ_U. Since $U \leq Z(G)$, we have $\chi_U = \chi(1)\lambda$ and $\det(\chi)_U = \lambda^{\chi(1)}$. Since $U \leq G'$, we have $\det(\chi)_U = 1_U$ and $\lambda^{\chi(1)} = 1_U$. Therefore $p \nmid \chi(1)$, $\lambda \neq 1_U$ and $|\, U \,| = p$, a contradiction. □

Bibliography

[Ben72] Helmut Bender, *A group theoretic proof of Burnside's p^aq^b-theorem*, Math. Z. **126** (1972), 327–338. MR 0322048

[Bla19] Chris Blair, *Character table of S_5*, online, Trinity College Dublin, May 2019, `https://www.maths.tcd.ie/~cblair/notes/s5.pdf`.

[Bur04] W. Burnside, *On Groups of Order $p^\alpha q^\beta$*, Proc. London Math. Soc. (2) **1** (1904), 388–392. MR 1576790

[Cal11] Danny Calegari, *The Hall-Witt identity*, November 2011, `https://lamington.wordpress.com/2011/11/20/the-hall-witt-identity/`.

[Car19] A. Caranti, *Alcune note per un corso di teoria dei gruppi*, online, Università degli Studi di Trento, 2019, `http://www.science.unitn.it/~caranti/Didattica/Gruppi/static/Note/Note_Gruppi.pdf`.

[CM19] A. Caranti and S. Mattarei, *Note di algebra per un corso da 12 crediti*, online, Università degli Studi di Trento, 2019, `http://www.science.unitn.it/~caranti/Didattica/Algebra/static/Note/Algebra.pdf`.

[Gol70] David M. Goldschmidt, *A group theoretic proof of the p^aq^b theorem for odd primes*, Math. Z. **113** (1970), 373–375. MR 0276338

[Gor80] Daniel Gorenstein, *Finite groups*, second ed., Chelsea Publishing Co., New York, 1980. MR 569209

[Hup67] B. Huppert, *Endliche Gruppen. I*, Die Grundlehren der Mathematischen Wissenschaften, Band 134, Springer-Verlag, Berlin-New York, 1967. MR 0224703 (37 #302)

[Isa06] I. Martin Isaacs, *Character theory of finite groups*, AMS Chelsea Publishing, Providence, RI, 2006, Corrected reprint of the 1976 original [Academic Press, New York; MR0460423]. MR 2270898

[Jac85] Nathan Jacobson, *Basic algebra. I*, second ed., W. H. Freeman and Company, New York, 1985. MR 780184

[Lan02] Serge Lang, *Algebra*, third ed., Graduate Texts in Mathematics, vol. 211, Springer-Verlag, New York, 2002. MR 1878556

[LOST10] Martin W. Liebeck, E. A. O'Brien, Aner Shalev, and Pham Huu Tiep, *The Ore conjecture*, J. Eur. Math. Soc. (JEMS) **12** (2010), no. 4, 939–1008. MR 2654085

[Mac12] Antonio Machì, *Groups*, Unitext, vol. 58, Springer, Milan, 2012, An introduction to ideas and methods of the theory of groups. MR 2987234

[Mar18] Daniel A. Marcus, *Number fields*, Universitext, Springer, Cham, 2018, Second edition of [MR0457396], with a foreword by Barry Mazur. MR 3822326

[Mat73] Hiroshi Matsuyama, *Solvability of groups of order 2^ap^b*, Osaka J. Math. **10** (1973), 375–378, `http://projecteuclid.org/euclid.ojm/1200694311`. MR 0323890

[Rob96] Derek J. S. Robinson, *A course in the theory of groups*, second ed., Graduate Texts in Mathematics, vol. 80, Springer-Verlag, New York, 1996. MR 1357169 (96f:20001)

[Rot95] Joseph J. Rotman, *An introduction to the theory of groups*, fourth ed., Graduate Texts in Mathematics, vol. 148, Springer-Verlag, New York, 1995. MR 1307623

[Ser77] Jean-Pierre Serre, *Linear representations of finite groups*, Springer-Verlag, New York-Heidelberg, 1977, Translated from the second French edition by Leonard L. Scott, Graduate Texts in Mathematics, Vol. 42. MR 0450380

[Ser78] ———, *Représentations linéaires des groupes finis*, revised ed., Hermann, Paris, 1978. MR 543841

[Ser16] ———, *Finite groups: an introduction*, Surveys of Modern Mathematics, vol. 10, International Press, Somerville, MA; Higher Education Press, Beijing, 2016, With assistance in translation provided by Garving K. Luli and Pin Yu. MR 3469786

[Spe10] David E. Speyer, *Recovering representation from its character*, MathOverflow, 2010, `https://mathoverflow.net/q/32846(version:2010-07-21)`.

[Tao12] Terence Tao, *Cayley graphs and the algebra of groups*, May 2012, `https://terrytao.wordpress.com/tag/hall-witt-identity/`.