.

MTTS 2004, May 17 - June 12, 2004
Level I - Group theory

B.Sury
I.S.I., Bangalore

This is an expanded version of the original notes. Some of the material is taken from my book "Group theory - Selected Problems" published in 2004 by the Universities Press and distributed by Orient-Longman Pvt.Ltd.

§0 Introduction

The presence of group theory can often be felt even when it is not seen. Let me explain the meaning of the first statement. Generally, what one sees is some geometrical object and the underlying group of symmetries is always there, guiding the whole structure but, to be seen, it has to take an avataar. Of course, the significance of a notion lies in the number of diverse situations in which it plays a substantial role. In that respect, the notion of a group is one of the most - if not the most - important in mathematics as well as in other sciences. We digress for a moment to ask a question. We have been saying that a group may be invisible but may be at work in a situation. Here is a problem which exemplifies this.

Consider all integers of the form $14x + 20y$ and all integers of the form $24x + 34y$ where x, y run through integers. Then, look at the lattice points $(14x + 20y, 24x + 34y)$ in the plane. This set intersects the Y-axis and the X-axis at times. What is the smallest natural number n for which both $(n, 0)$ and $(0, n)$ are in this set? What is the proportion of lattice points on the plane which lie on this set?

These questions can be easily and naturally answered by recognising the groups figuring here. We shall do so later but you are urged to think about it. I offer a small prize to whoever solves either question first.

Group as Latin square

All of you have probably already learnt a lot of group theory but let us start with some easy and familiar things. Let us start with a visual way of defining groups. This idea is due to Uri (not Sury !) Rimon of Hebrew University, Jerusalem.

View multiplication table of group as a Latin square with a distinguished vertex. That is to say, there is a square array of symbols (to represent the elements of the group) with each symbol occurring exactly once in each row and in each column. There is a distinguished symbol e such that whenever one draws a rectangle $eacb$ cornered at this symbol with ea along a column and eb along a row, then the symbol c diagonally opposite to it can be taken to represent the product ab. It is a nice exercise to deduce associativity and verify that such a table is equivalent to a group structure on the symbols. It is also a nice exercise to prove statements like $(ab)^{-1} = b^{-1}a^{-1}$.

A word of warning : Not every Latin square with a distinguished vertex arises from a group. It is possibly true that if the operation using rectangles as above is well-defined, then the Latin square indeed comes from a group.

As we remarked, it is fruitful to view the same notion in different guises

as it tells us some new aspect of its personality. One could think of a group as a special equivalence relation which is often how it arises. By the way, continuing in the same vein of visual viewing, one may think of a general equivalence relation in the following nice way. Given a set of d objects, a relation is just a $d \times d$ matrix R of 0's and 1's associated in an obvious manner. Reflexivitiy and symmetricity correspond, respectively, to the relation matrix having diagonal entries 1 and being symmetric.
Ex : What does transitivity correspond to?
Ans. $R_{ij} \neq 0$ implies $(R^2)_{ij} \neq 0$.

Without further ado, let us now start with the examples of groups that one comes across. We shall then study them in different ways; study them according to their size, study them according to the way they 'act', study them according to their internal structure etc. Needless to say, there are still many things to be properly understood. For instance, the finite simple groups have been classified but it cannot be said that the classification of finite simple groups (CFSG) is well-understood (indeed, no proof which is 5000 pages long can be so accepted).

Examples and basic notions.

- $\mathbb{Z}/n$ denotes the group of integers modulo n, under addition mod n.
- $(\mathbb{Z}/n)^* = \{a \leq n : (a,n) = 1\}$ under multiplication mod n.

The notations $\mathbb{Z}/n$ and $(\mathbb{Z}/n)^*$ are perhaps less familiar compared to $\mathbb{Z}_n$ etc.

- $S^1 = \{z \in \mathbf{C} : |z| = 1\}$ under multiplication.

All the above are abelian groups.

- S_n, the set of all permutations (i.e., bijections) of n symbols, under the composition of permutations. In fact, for any set X, the set $Sym(X)$ of all bijections on X is a group under composition. The subset $S_F(X)$ consisting of all those bijections which move only finitely many elements, is a subgroup. We shall use the convention that the permutation $\sigma\tau$ is obtained by applying σ first and τ later.
- $GL(n, \mathbf{Q})$, the set of all $n \times n$ rational matrices which have non-zero determinants (the symbol GL stands for 'general linear') is a group under matrix multiplication.

Similarly, one can define $GL(n, \mathbf{R})$ and $GL(n, \mathbf{C})$.
One has also the 'special linear' groups $SL(n, \mathbf{Q}), SL(n, \mathbf{R})$ etc. consisting of the matrices which have determinant 1.

• $SO(n) = \{g \in SL(n, \mathbf{R}) : gg^t = I_n\}$, is known as the special orthogonal group of degree n.
• $SU(n) = \{g \in SL(n, \mathbf{C}) : \bar{g}g^t = I_n\}$, is known as the special unitary group of degree n.
• $Sp(n) = \{g \in SL(2n, \mathbf{R}) : g^t \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix} g = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}\}$, is known as the symplectic group of degree n; here 0 and I_n denote $n \times n$ block matrices. These are nonabelian groups when $n > 1$.
• $B(n, \mathbf{Q})$, the upper triangular rational matrices with nonzero diagonal entries, $U(n, \mathbf{Q})$, the subset consisting of those upper triangular matrices which have all diagonal entries equal to 1 and $T(n, \mathbf{Q})$, the diagonal matrices with all diagonal entries nonzero rational numbers, are groups.
Note that $T(n, \mathbf{Q})$, for any n and, $U(2, \mathbf{Q})$ are abelian groups.
• $GL(n, \mathbb{Z}) = \{g$ an integral matrix : $det(g) = \pm 1\}$.
• $GL(n, \mathbb{Z}[1/p])$ where p is a prime number.
Here, $\mathbb{Z}[1/p]$ denotes the set of rational numbers whose denominators are divisible only by p and, the above group consists of all matrices whose determinants are $\pm$ a power of p.
Note that

$$GL(n, \mathbb{Z}) \leq GL(n, \mathbb{Z}[1/p]) \leq GL(n, \mathbf{Q})$$

where we have used the notation $H \leq G$ to denote the fact that H is a subgroup of G. We shall also write $G \geq H$ at times. Also, $H < G$ for groups would mean that H is a proper subgroup.
Let us note that the set of matrices

$$G = \{\begin{pmatrix} a & a \\ 0 & 0 \end{pmatrix} : a \in \mathbf{R}^*\}$$

forms a group under matrix multiplication but it is *not* a subgroup of $GL(2, \mathbf{R})$.
We have

$$Perm(n) = \{P_\sigma : \sigma \in S_n\} \leq GL(n, \mathbb{Z})$$

where P_σ is the 'permutation matrix' whose rows are the rows of the identity matrix permuted according to σ i.e., the rows of P_σ are $R_{\sigma(1)}, \cdots, R_{\sigma(n)}$.
Note that trivially $\bigcap_I H_i \leq G$ if $H_i (i \in I) \leq G$ while the union of two subgroups may not be a subgroup. Further, if g is an element of a group G, then its *centraliser* $C_G(g) := \{x \in G : xg = gx\} \leq G$. Moreover, for any subset $S \subset G$, the subgroup $C_G(S) := \bigcap_{s \in S} C_G(s)$ is the *centraliser of* S. For $S = G$, the centraliser $C_G(G)$ is usually denoted by $Z(G)$, and is called

the *center* of G.
If $\mu(n)$ denotes the complex n-th roots of unity, we have a chain of subgroups

$$\mu(n) \leq \bigcup_n \mu(n) \leq S^1 \leq \mathbf{C}^*.$$

For any subset $S \subset G$, one denotes by $< S >$, the subgroup generated by S. In concrete terms, it consists of all finite products of elements of S and their inverses. For instance, for any group G and any positive integer n, the subset $G^n := \{g^n : g \in G\}$ gives a subgroup $< G^n >$. Note that, in any abelian group G, we have $< G^n >= G^n$.
Recall that a subgroup N is *normal* in G if $gxg^{-1} \in N$ for all $g \in G, x \in N$. This notion comes from Galois theory. Evidently, for any G and any n, the subgroup $< G^n >$ is normal in G; we write $< G^n > \trianglelefteq G$.
Given a normal subgroup N, one can give a group structure to the set G/N of all (left or right - both are same) cosets of the subgroup. There is a natural onto homomorphism from G to G/N whose kernel is N and, conversely, the kernel of any homomorphism from G to some group, is a normal subgroup of G.

Examples of homomorphisms :

• $i : H \to G$; inclusion of a subgroup H in G.
• For any n, $\mathbf{Z} \to \mathbf{Z}; a \mapsto na$.
• $\mathbf{Z} \to \mathbf{Z}/n; a \mapsto a$ mod n.
• $GL(n, \mathbf{C}) \to \mathbf{C}^*; g \mapsto det(g)$.
• $S_n \to Perm(n); \sigma \mapsto P_\sigma$.
• The composition of two homomorphisms.
Note that the composite of the last-mentioned homomorphism with the determinant homomorphism is the 'sign' of the permutation.
• $G \to G; g \mapsto g^n$ if G is abelian.
• If $0 \neq \lambda \in \mathbf{R}, \mathbf{Z} \to \mathbf{R}^*; n \mapsto \lambda^n$.
• $\mathbf{R} \to \mathbf{R}^*; t \mapsto e^t$.
• $\mathbf{R} \to U(2, \mathbf{R}); t \mapsto \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$.
• $S^1 \to SO(2) = \{g \in SL(2, \mathbf{R}) : gg^t = 1\}; e^{i\theta} \mapsto \begin{pmatrix} Cos\theta & Sin\theta \\ -Sin\theta & Cos\theta \end{pmatrix}$.
• $SO(2) \to S^1; g \mapsto$ last column.
• $SU(2) \to S^3; g \mapsto$ last column.
Here $S^{n-1} = \{(z_1, \cdots, z_n) \in \mathbf{R}^n : \sum_{i=1}^n z_i^2 = 1\}$, the unit sphere in n dimensional space. The above two are isomorphisms. Some of the other homomorphisms above are isomorphisms as well.

One says that a group G is *finitely generated* if there exists a finite subset $S \subset G$ such that $G =< S >$.

Given a subset $S \subset G$, one also defines the *normal subgroup generated by S* as the smallest normal subgroup $< S >^N$ of G which contains S. In concrete terms, it consists of all finite products of conjugates of elements of S and their inverses.

For any subset $S \subset G$, we define the *normaliser*

$$N_G(S) = \{g \in G : gSg^{-1} = S\}.$$

Observe $< S > \trianglelefteq N_G(S) \leq G$ and $N_G(S)$ is the largest subgroup of G containing S in which $< S >$ is normal.

For any group G, and $x, y \in G$, we denote $[x, y] = xyx^{-1}y^{-1}$. Such elements are called commutators in G. If S is the subset of all commutators of G, one obtains the *commutator subgroup* of G, denoted by $[G, G]$. Sometimes, one also denotes $[G, G]$ by $D(G)$ and refers to it as the *derived group of G*.

More generally, for $H, K \leq G$, $[H, K]$ denotes the subgroup of G generated by $\{[h, k] : h \in H, k \in K\}$.

A group G is said to be *solvable* - the terminology comes from Galois theory - if the chain

$$G \geq D^1(G) := D(G) \geq D^2(G) := D(D(G)) \geq \cdots\cdots$$

becomes the trivial group in a finite number of steps. Evidently, any abelian group is solvable. A nonabelian example is the group $B(n, \mathbf{Q})$, the upper triangular rational matrices with nonzero diagonal entries for any $n \geq 2$.

A complementary concept is that of a *perfect* group; a group G is said to be *perfect* if $G = [G, G]$.

A group G is said to be *nilpotent* if the lower central series

$$G \geq C^1(G) := D(G) \geq C^2(G) \geq C^3(G) \cdots$$

becomes trivial in a finite number of steps where $C^{n+1}(G) = [G, C^n(G)]$.

It is not only evident that abelian groups are nilpotent but also that nilpotent groups are solvable because, it is seen by induction on n that $D^n(G) \leq C^n(G)$ for any G. The group $U(n, \mathbf{Q})$ is nilpotent but nonabelian when $n > 2$.

In any group G, let x^y denote yxy^{-1} and $[x, y]$ denote the commutator $xyx^{-1}y^{-1}$. It is also convenient to define x^{-y} to be $yx^{-1}y^{-1}$. Also, one defines inductively

$$[x_1, \cdots, x_n] := [[x_1, \cdots, x_{n-1}], x_n].$$

Free groups and presentations

F_n, the free group of rank n, is defined to be the set of all *reduced (finite) words* in n symbols $x_1, \cdots, x_n$ and their 'inverse symbols' denoted by $x_1^{-1}, \cdots, x_n^{-1}$. Here, a word is said to be reduced if it does not contain two symbols x_i and x_i^{-1} in juxtaposition. The empty word is also counted and is the identity element. The group multiplication is by concatenation of two reduced words and cancelling off all consecutive symbols of the form $x_i x_i^{-1}$ or $x_i^{-1} x_i$.

F_n is nonabelian if $n \geq 2$. It is evident that any group which can be generated by n elements can be identified upto isomorphism with a quotient group of F_n.

A group G is said to be *finitely presentable* if $G \cong F_n/K$ for some n where $K =< R >^N$ for some finite subset R of F_n. The choice of finite sets $X \subset G$ with $|X| = n$ and $R \subset F_n$ with $F_n/ < R >^N \cong G$ is called a *finite presentation of* G.

If $G =< X|R >$ and $H =< Y|S >$, then their *free product* is defined to be the group $< X \sqcup Y|R \cup S >$ and is denoted by $G * H$. Note that F_n is a free product of $\mathbb{Z}$ with itself n times (taking free products is an associative operation). We shall see via some problems that familiar groups like $SL(2, \mathbb{Z})/\{\pm I\}$ are free products.

Group actions

Let us recall the concept of group actions which makes groups one of the most powerful mathematical tools.

A group G is said to *act on a set* S if there is a homomorphism from G to the permutation group $Sym(S)$ of S. The action is said to be *faithful* if the homomorphism is injective. It is customary to write $g.s$ for the element of S that $s \in S$ is sent to by the permutation corresponding to $g \in G$. For $s \in S$, the subset $G.s := \{g.s : g \in G\}$ is called the *orbit of* s. For any $g \in G$, the subset $S^g := \{s \in S : g.s = s\}$ is called the set of fixed points under g. For $s \in S$, the subgroup $G_s := \{g \in G : g.s = s\}$ is called the isotropy subgroup at s. Note that, for $s \in S$, the map $g \mapsto g.s$ is a well-defined bijection from the set G/G_s of left cosets of G_s to the orbit $G.s$ of s. Thus, orbits under a finite group action have cardinalities which are divisors of the order of the group.

In the particular case of a group acting on itself by conjugation, we denote the orbit of an element g (this is the conjugacy class of g) by $G(g)$. One also writes $g \sim h$ to mean that g and h are in the same orbit; that is, they are conjugate.

One calls an action of G on a set S *transitive* if each orbit is the whole set S.
For instance, S_n acts transitively on $\{1, 2, \cdots, n\}$.
The group $GL(n, \mathbf{R})$ acts transitively on $\mathbf{R}^n - (0)$.
A group G is said to *act r-transitively* on a set S if G acts transitively on the set

$$\{(s_1, \cdots, s_r) \in S^r : g_i \ \ distinct\ \}.$$

Automorphism group

For an element g in a group G, one usually denotes the automorphism $x \mapsto g^{-1}xg$ of G as Int (g). The group Aut (G) of all automorphisms of G is defined by means of the composition in the following order: $g(\sigma\tau) = (g(\sigma))\tau$; this is consistent with our convention of multiplying permutations. In this convention, the set of automorphisms $x \mapsto Int(x)$ can be identified with a subgroup of Aut G because

$$g(Int(xy)) = (xy)^{-1}g(xy) = g(Int(x))(Int(y)).$$

This is in fact, a normal subgroup.

Jordan-Hölder theorem

The last important notion we recall is that of composition series and the Jordan-Holder theorem. Recall that a *normal series* of a group G is simply a finite sequence of subgroups

$$\{1\} = G_0 \lhd G_1 \cdots \lhd G_n = G.$$

Of course, we know that the G_i's need not be normal in the whole of G; they are usually called *subnormal* subgroups of G. A normal series

$$\{1\} = H_0 \lhd H_1 \cdots \lhd H_m = G$$

is said to be a *refinement* of a normal series

$$\{1\} = G_0 \lhd G_1 \cdots \lhd G_n = G$$

if each G_i is some H_j; it is called a proper refinement if $m > n$.
It is elementary to show that any two normal series have refinements which are *isomorphic*; that is, refinements in which any successive factor G_i/G_{i-1} of one is isomorphic to a successive factor of the other and vice versa.
One says that a normal series is a *composition series* if it has no proper

refinements. Note that this is equivalent to the successive factors being simple groups.
For instance, any normal series for $\mathbb{Z}$ will have a proper refinement because the smallest nontrivial term in such a series is an infinite cyclic group and will have proper nontrivial (normal) subgroups. Therefore, $\mathbb{Z}$ has *no* composition series.
The theorem of Jordan-Holder asserts that for a *finite* group, each normal series admits of a refinement which is a composition series and that any two composition series are isomorphic.
Finally, we recall Zorn's lemma which is equivalent to the axiom of choice. The use of Zorn's lemma becomes unvoidable when we need to prove facts like every vector space has a basis. Let $(S, \leq)$ be any non-empty partially ordered set. This means the three properties : (i) $s \leq s$ for all $s \in S$; (ii) $s \leq t, t \leq s \Rightarrow s = t$ and, (iii) $s \leq t \leq u \Rightarrow s \leq u$.
A chain is a subset T of S in which, for any two elements $s, t \in T$, either $s \leq t$ or $t \leq s$. Zorn's lemma asserts that if every chain T in S has an upper bound (that is, an element $s_0 \in S$ such that $t \leq s_0$ for all $t \in T$), then S has a maximal element (that is, an element $m \in S$ which is not $\leq$ any element of S other than itself).
Finally, we recall that a group is said to have exponent d if d is the smallest natural number for which each element g satisfies $g^d = 1$.
One word of convention - a subgroup M of a group G is *maximal* if M is a proper subgroup which is not contained in any other proper subgroup of G. Also, the identity element is always denoted by 1 unless there is an abelian group written additively when it is denoted by 0. Also, we write $O(g)$ for the order of an element g of a group G while the order of the group itself is written as $|G|$. The notation $p^r \| n$ means that p^r is the highest p-power dividing n.

Now, we begin with the systematic study of groups. After studying cyclic and abelian groups, we discuss group actions and permutation groups. Then, we shall study nilpotent and solvable groups. Following that we study groups of matrices and groups figuring in geometry. Finally, we discuss miscellaneous results on groups which do not fit exactly into any of the above titles. Each of these discussions will be interspersed with several related results and miscellaneous comments. **Throughout the notes, we use the definitions and notations introduced in section §0.**

§1 Cyclic groups

A group generated by a single element is called *cyclic.*
The only infinite cyclic group, upto isomorphism, is $\mathbf{Z}$ and for any n, the only finite cyclic group of order n is, upto isomorphism, the set $\mathbf{Z}/n$ of integers modulo n considered under addition modulo n. Indeed, for finite cyclic G, $g \mapsto \bar{1}$ is an isomorphism onto $\mathbb{Z}/|G|$.
Evidently, Subgroups and quotient groups of cyclic groups are cyclic as well but it is not true that a subgroup of an n-generated group is n-generated although a quotient of an n-generated group is n-generated.

Lemma 1.1
(i) The direct product $G \times H$ of nontrivial cyclic groups G, H is cyclic if, and only if, both G, H are finite and their orders m, n are coprime. (ii) A finite group is cyclic if, and only if, it has a unique subgroup of each order dividing the order of the group. Hence $n = \sum_{d|n} \phi(d)$ for each natural number n.

Lemma 1.2
Let G be a finite group in which, for every $n/|G|$, the set $\{g : g^n = e\}$ has at most n elements. Then, G is cyclic and the set $\{g : g^n = e\}$ has exactly n elements for each $n||G|$. In particular, any finite subgroup of K^* is cyclic for any field K.

Note that for the application to field theory, we need only prove the result for abelian G. Later, we will discuss this again and prove it as an application of Sylow's theorems.

Proof of 1.2
Let $|G| = n$ and $d|n$.
Consider $N(d) = \#\{g \in G : O(g) = d\}$.
If $N(d) \neq 0$, look at some element g with $O(g) = d$. As $e, g, g^2, \cdots, g^{d-1}$ are distinct and are solutions of $x^d = e$, these are *all* the solutions of the equation $x^d = e$. As elements of order d in G are among these and are $\phi(d)$ in number, we have proved that $N(d) = \phi(d)$ if $N(d) \neq 0$. As every element of G has some order d dividing n, we have $n = \sum_{d|n} N(d)$. Since $n = \sum_{d|n} \phi(d) \geq \sum_{d|n} N(d) = n$, we must have the equality $N(d) = \phi(d)$ for all $d|n$. In particular, $N(n) = \phi(n) \neq 0$.

All of us are familiar with the division algebra H of Hamilton's quaternions

$$H := \{a_0 + a_1 i + a_2 j + a_3 k : a_i \in \mathbf{R}\}$$

where the mulltiplication is defined by $i^2 = j^2 = k^2 = -1$, and $ij = k = -ji$. Note that H is a 4-dimensional vector space over $\mathbf{R}$ and is central over $\mathbf{R}$. In general, a finite-dimensional division algebra D over a field K is a finite-dimensional vector space D over K such that D has a multiplication under

which the set D^* of all non-zero elements forms a group; the mutliplication is also required to satisfy $x(y+z) = xy + xz$ and $(x+y)z = xz + yz$ and, $\lambda(xy) = (\lambda x)y$ for all $\lambda \in K$ and all $x, y, z \in D$. Note that D is a field if, and only if, D^* is an abelian group.
Further, D is said to be *central* over K if the centre of D^* is K^*.
A very similar construction to H above produces many examples of central division algebras over $\mathbf{Q}$.

Remark 1.3
Let D be any finite-dimensional central division algebra over a field F of characteristic $p > 0$. Then, every finite subgroup of D^* is again cyclic.
On the other hand, the above Hamilton quaternion algebra

$$H := \{a_0 + a_1 i + a_2 j + a_3 k : a_i \in \mathbf{R}\}$$

is such that $H^* := H \setminus (0)$ is a group which contains a finite noncyclic group - the quaternion group $Q = \{\pm 1, \pm i, \pm j, \pm k\}$.
To see how the first remark follows, let $F = Z(D)$ be the centre of D. Note that $F \supseteq \mathbf{F}_p$, the field of p elements. If $G \leq D^*$ is a finite subgroup, then the set

$$K := \{\sum_{g \in G} \alpha_g g : \alpha \in \mathbf{F}_p\}$$

is a finite-dimensional $\mathbf{F}_p$-vector space. So, it is a finite division algebra and, therefore, it must be a field K and $G \leq K^*$. Therefore, by lemma 1.2, G must be cyclic.

Remark 1.4 (groups vs congruences)
A re-statement of Lagrange's theorem for the groups $(\mathbb{Z}/p)^*$ and $(\mathbb{Z}/n)^*$, one has Fermat's little theorem asserting $a^{p-1} \equiv 1 \bmod p$ for prime p and $(a,p) = 1$ and Euler's congruence asserting $a^{\phi(n)} \equiv 1 \bmod n$ for $(a,n) = 1$.
In fact, we have the assertion $n | \phi(a^n - 1)$ for natural numbers a, n. This is because, in the group $(\mathbb{Z}/(a^n-1))^*$, the integer a has order n.
The Wilson congruence that for a prime p, one has $(p-1)! \equiv -1 \bmod p$ follows by looking at the product of all elements in $\mathbb{Z}_p^*$. In the product, each element cancels with its inverse except for those elements which are their inverses. The only elements of $(\mathbb{Z}/p)^*$ which are their own inverses are 1 and $p-1$ because if $i^2 \equiv 1 \bmod p$, then $p|(i-1)$ or $p|(i+1)$; so the resultant product is $p-1$.
We have a more general congruence :
Let a be a natural number and p be a prime. Suppose $p|(a^n - 1), p^2 \not| (a^n - 1)$ for some $n \geq 1$. Then, $a^{p-1} \not\equiv 1$ modulo p^2.

To prove this, let $\theta : (\mathbb{Z}/p^2)^* \to (\mathbb{Z}/p)^*$ be the canonical homomorphism. Considering a as an element of $(\mathbb{Z}/p^2)^*$, it follows that $a^n \in \ker\,\theta$. Clearly, Ker θ has order p and a^n is a nontrivial element of $(\mathbb{Z}/p^2)^*$ since $p^2 \not| (a^n - 1)$ by hypothesis. Therefore, a^n has order p in $(\mathbb{Z}/p^2)^*$. This means p divides the order of a in $(\mathbb{Z}/p^2)^*$ and, thus $a^{p-1} \not\equiv 1$ modulo p^2.

Theorem 1.5
$(\mathbb{Z}/n)^* = \{a \leq n : (a,n) = 1\}$ is cyclic if, and only if, $n = 2, 4, p^r$ or $2p^r$ for some odd prime p.
Any generator for the group for such n is called a primitive root modulo n in number-theoretic parlance.

Exercise 1.6
Are $\mathbb{Z}[1/2]$ and $\mathbb{Z}[1/3]$ isomorphic?

§2 **Abelian groups**

Lemma 2.1
A finite abelian group A is isomorphic to the direct sum of cyclic p-groups for various primes p dividing its order.

A free abelian group of rank n is defined to be a group isomorphic to $\mathbf{Z}^n$, the set of integral n-tuples under co-ordinatewise addition. Equivalently, it is an abelian group G with a basis of n elements $g_1, \cdots, g_n$ i.e., $\sum_i a_i g_i = 0$ implies $a_i = 0$ for all i and $G = \{a_1 g_1 + \cdots + a_n g_n : a_i \in \mathbf{Z}\}$. The rank n is uniquely determined since the number of homomorphisms from G to $\mathbf{Z}/2$ is 2^n.

One of the most important results on abelian groups is the next theorem which, in turn, follows from the one following it :

Theorem 2.2 (Structure theorem for finitely generated abelian groups)
A finitely generated abelian group is isomorphic to $\mathbf{Z}^m \times \mathbf{Z}_{d_1} \times \cdots \times \mathbf{Z}_{d_r}$ for some $m \geq 0$ and d_i dividing d_{i+1}. The integer m as well as all the d_i's (upto sign) are uniquely determined.

Theorem 2.3 (Invariant factor theorem)
If H is a subgroup of a free abelian group G of rank n, then H is free abelian of rank $r \leq n$. Further, there are bases $\{e_1, \cdots, e_n\}$ of G and $\{d_1 e_1, \cdots, d_r e_r\}$ of H respectively where d_i divides d_{i+1} for $i < r$. The integers d_i are uniquely determined upto sign and are called the invariant factors of H.

The proof is carried out by induction on n using the division algorithm as follows. It is clear for $n = 1$. Assume $n > 1$ and that the theorem holds for $m < n$. Corresponding to any basis of G, there is a positive integer with the property that it is the smallest positive integer that occurs as a coefficient in the expression of elements of H in terms of this basis. This positive integer can depend on the basis and let l_1 be the smallest such with respect to all bases of G. Let $v_1, \cdots, v_n$ be a corresponding basis for G such that $v = l_1 v_1 + \sum_{i=2}^{n} a_i v_i \in H$. Dividing all the a_i by l_1, we have $a_i = q_i l_1 + r_i$ with $0 \leq r_i < l_1$. Evidently, $v = l_1(v_1 + \sum_{i=2}^{n} q_i v_i) + \sum_{i=2}^{n} r_i v_i$ and $v_1 + \sum_{i=2}^{n} q_i v_i, v_2, \cdots, v_n$ is another basis of G. By the minimality of l_1, we must have $r_i = 0$ for all $i \geq 2$. Thus, writing w_1 for $v_1 + \sum_{i=2}^{n} q_i v_i$, $v = l_1 w_1 \in H$. Look at the subset H_0 of H which have coefficients of w_1 to be zero in terms of the basis $w_1, v_2, \cdots, v_n$ of G. Clearly, H_0 is a subgroup of H such that $H_0 \cap \mathbf{Z}v = \{0\}$. Also, if $h \in H$, write $h = b_1 w_1 + \sum_{i=2}^{n} b_i v_i$. Once again, dividing the b_i's by l_1, say, $b_i = m_i l_1 + s_i$ with $0 \leq s_i < l_1$, we have $h - m_1 v = s_1 w_1 + \sum_{i=2}^{n} b_i v_i \in H$. Thus, by the minimality of l_1 we get $s_1 = 0$ i.e., $h - m_1 v \in H_0$. Thus, $H = H_0 \oplus \mathbf{Z}v$. Now, H_0 is contained in the subgroup $G_0 = \sum_{i=2}^{n} \mathbf{Z} v_i$. By induction hypothesis, G_0 has a basis $w_2, \cdots, w_n$ and there exists $r \leq n$ such that H_0 has a basis of the form $d_2 w_2, \cdots, d_r w_r$ with $d_2 | d_3 | \cdots | d_n$. Clearly, therefore, H itself has rank $r \leq n$ and $l_1 w_1, d_2 w_2, \cdots, d_n w_n$ is a basis for H. We have only to show that $l_1 | d_2$. Once again, writing $d_2 = c l_1 + d$ with $0 \leq d < l_1$, we notice $l_1 w_1 + d_2 w_2 = l_1(w_1 + c w_2) + d w_2 \in H$ where $w_1 + c w_2, w_2, \cdots, w_n$ is a basis of G. Thus, minimality of l_1 forces $d = 0$ i.e., $l_1 | d_2$. The proof is complete.

Exercise

(i) Deduce 2.2 from 2.3.

(ii) Give an example of a group G and a subgroup H such that G can be generated by some number n of elements while H cannot be n-generated.

(iii) Solve the prize problem posed in the beginning.

Lemma 2.4

The existence of bases as in the invariant factor theorem is equivalent to the following statement about matrices :

Given any $A \in M_{m,n}(\mathbb{Z})$ of maximum possible rank, there exist $P \in GL(m, \mathbb{Z})$ and $Q \in GL(n, \mathbb{Z})$ such that PAQ is a matrix whose'diagonal' entries are $d_1, d_2, \cdots$ where $d_i | d_{i+1}$.

Further, $\mathrm{GL}(n, \mathbb{Z})$ is generated by elementary matrices $I + E_{ij}$.

Proof

Suppose the matrix statement holds. Let H be a subgroup of a free abelian group G of rank n. Then, H is also free abelian of rank $m \leq n$ (this we are assuming known through other arguments). Let $\alpha : \mathbb{Z}^m \to H$ and $\beta : G \to \mathbb{Z}^n$ be isomorphisms. If $i : H \leq G$ denotes the inclusion map, we have the composite map $\beta \circ i \circ \alpha$ corresponds to a matrix $A \in M_{n,m}(\mathbb{Z})$ with respect to the canonical ordered bases of $\mathbb{Z}^m$ and $\mathbb{Z}^n$. The matrix statement gives us $P \in GL(n, \mathbb{Z})$ and $Q \in GL(m, \mathbb{Z})$ such that

$$AQ = P \ . \begin{pmatrix} d_1 & \cdots & 0 \\ \ddots & \ddots & \ddots \\ 0 & \cdots & d_m \\ 0 & \cdots & 0 \\ 0 & \cdots & 0 \end{pmatrix}$$

where $d_i | d_{i+1}$.
Hence, the bases

$$\{v_1, \cdots, v_n\} = \{Pe_1, \cdots, Pe_n\}$$

of $\mathbb{Z}^n$ and

$$\{w_1, \cdots, w_m\} = \{Qe_1, \cdots, Qe_m\}$$

of $\mathbb{Z}^m$ are so that

$$\{\beta^{-1}(v_1), \cdots, \beta^{-1}(v_n)\}$$

is a basis for G and $\{\alpha(w_1), \cdots, \alpha(w_m)\}$ is a basis for H.
Now, note that the matrix identity above implies that $AQ(e_i) = P(d_i e_i)$ where e_i on the left side are in $\mathbb{Z}^m$ and those on the right side are in $\mathbb{Z}^n$.
That is, $\beta\alpha(Qe_i) = d_i P(e_i)$.
So, we have $\beta\alpha(w_i) = d_i v_i$, which means that the bases $\{\beta^{-1}(v_1), \cdots, \beta^{-1}(v_n)\}$ of G and $\{\alpha(w_1), \cdots, \alpha(w_m)\}$ of H are as asserted in the invariant factor theorem.
Conversely, let us assume that the invariant factor theorem holds. Consider any $A \in M_{n,m}(\mathbb{Z})$ of rank $\max(m, n)$. Without loss of generality, we shall take $m \leq n$ for, otherwise, we could take the transpose. Now, A defines a homomorphism

$$T_A : \mathbb{Z}^m \to \mathbb{Z}^n \ ; v \mapsto Av.$$

Now the image of T_A is a free abelian group generated by the n vectors $Ae_1, \cdots, Ae_m$.
Since the matrix A has rank m, the vectors $Ae_1, \cdots, Ae_m$ are linearly independent vectors over $\mathbf{Q}$. Therefore, they are linearly independent over $\mathbb{Z}$ also. In other words, Image T_A is free abelian subgroup of $\mathbb{Z}^n$ of rank m.
By the invariant factor theorem, let us choose bases $\{v_1, \cdots, v_n\}$ of $\mathbb{Z}^n$ and $\{d_1 v_1, \cdots, d_m v_m\}$ of Image T_A such that $d_i | d_{+1}$. Call $Aw_i = d_i v_i$ for all

$i \leq m$.
Let $P \in GL(n, \mathbb{Z})$ denote the matrix effecting the change of basis from the canonical basis to the v_i's. Similarly, let $Q \in GL(m, \mathbb{Z})$ be the matrix effecting the change of basis from the canonical basis to the w_i's.
Then, $P^{-1}AQ(e_i) = d_i v_i$ for all $i \leq m$. In other words, $P^{-1}AQ$ has the form asserted.
The above proof of the invariant factor theorem clearly shows the generation of $GL(n, \mathbb{Z})$ by the elementary matrices.

Exercise 2.5
Prove that $SL(n, \mathbb{Z})$ is perfect for $n \geq 3$.

Lemma 2.6
For any $A \in M_{m,n}(\mathbb{Z})$ define $h_i(A)$ to be the GCD of all $i \times i$ minors of A. If A has maximal rank, then for any $P \in GL(m, \mathbb{Z})$ and $Q \in GL(n, \mathbb{Z})$, the numbers $h_i(A) = h_i(PA) = h_i(AQ)$ for all i. The invariant factors of a matrix $A \in M_{m,n}(\mathbb{Z})$ are $h_1(A), \frac{h_2(A)}{h_1(A)}, \frac{h_3(A)}{h_2(A)}, \cdots$ etc.
We know that $GL(n, \mathbb{Z})$ is generated by the matrices of the form $X_{ij} = I + E_{ij}; i \neq j$ and the matrices $diag(\pm 1, \cdots, \pm 1)$. elsewhere. We shall check for each r that

$$h_r(AX_{ij}) = h_r(X_{ij}A)$$

for all $i \neq j \leq n$.
By the previous lemma, we need to consider only A of the 'diagonal' form with non-zero entries $d_1, \cdots, d_m$ with $d_i | d_{i+1}$.
Therefore, it is clear that $h_r(AD) = h_r(DA)$ for $D = diag(\pm 1, \cdots, \pm 1)$.
Now, for such A, we have, if $i > m$ that $AX_{ij} = A$ and, if $i \leq m$, $AX_{ij} = A + A'$ where A' is a matrix whose only nonzero entry is d_i at the (i, j)-th place.
Clearly, $h_r(AX_{ij}) = h_r(A)$.
Similarly, we see also that $h_r(X_{ij}A) = h_i(A)$. Therefore, we have the first assertion. For the second, we merely note that for 'diagonal' matrices A as above, with $d_i | d_{i+1}$, the numbers $h_i(A) = d_1 \cdots d_i$. Thus, the invariant factors are successive quotients of the h_i's.

Exercise 2.7
If S is any set of generators of the additive group of $\mathbf{Q}$, then S contains a proper subset of generators. Further, show $\mathbf{Q}$ does not have proper maximal subgroups.

§3 Group actions and permutation groups

One can use group actions to prove many of the basic results on groups as follows.

Cauchy's & Fermat's little theorems 3.1
Let G be any group of order n and let p any prime number. Consider the subset S of $G \times \cdots \times G$ (p times) defined by

$$S = \{(g_1, \cdots, g_p) : g_1 g_2 \cdots g_p = e\}.$$

Evidently, $|S| = n^{p-1}$. For each tuple $(g_1, \cdots, g_p)$ in S, there are exactly p distinct tuples $(g_2, \cdots, g_p, g_1)$, $(g_3, \cdots, g_p, g_1, g_2)$ etc. in S unless $g_1 = g_2 = \cdots = g_p$ (here is where we use the fact that p is prime). Note that $g_1 = g_2 = \cdots = g_p$ if, and only if, $g_1^p = e$. Thus, we have $|S| \equiv \#\{g : g^p = e\}$ mod p.
If $p|n$, then p divides $n^{p-1} = |S| \equiv \#\{g : g^p = e\}$ mod p. In this case, (since $e^p = e$), there are at least $p - 1$ elements of order p in G. This proves Cauchy's theorem.
If $p \not| n$, then one has $g^p = e$ for some g if, and only if, $e = g^{(n,p)} = g$. Thus, $|S| \equiv 1$ mod p. This proves Fermat's little theorem.

We have the following strikingly novel proofs of Fermat's famous theorem asserting that a prime number of the form $4n + 1$ is a sum of two squares. The proofs are due to Zagier, HeathBrown and Mohan Nair and are variants of the same argument.

Lemma 3.2
Let p any prime number of the form $4n + 1$. Consider the finite set

$$S = \{(x, y, z) \in \mathbf{N} \times \mathbf{N} \times \mathbf{N} : x^2 + 4yz = p\}.$$

Define the Zagier map $\sigma : S \to S$ by mapping (x, y, z) to

$$(x + 2z, z, y - x - z) \;\; if \quad x \leq y - z \ ,$$

$$(2y - x, y, x + z - y) \;\; if \quad y - z < x < 2y \ ,$$

$$(x - 2y, x + z - y, y) \;\; if \quad x \geq 2y \ .$$

Then σ is a permutation of order 2 and has a unique fixed point. Hence, any prime number p of the form $4n + 1$ is a sum of two squares.
First, note that in the definition of σ, one could have taken the $<$ sign wherever $\leq$ appears; the reason is that $x = y - z$ and $x = 2y$ are impossible to hold in S.
Now, it is clear that σ has the unique fixed point $(1, 1, n)$. Now, we note

the general fact that for any permutation τ of order 2 on a set, the non-fixed points can be paired off and thus the number of fixed points is of the same parity as $|S|$. Applying this to σ, we have that $|S|$ must be odd. Turning this around and applying the above observation to the permutation $\tau : (x, y, z) \mapsto (x, z, y)$, it follows that τ must have an odd number of fixed points. Therefore, τ does have at least one fixed point. Any fixed point of τ is a tuple (x, y, y) which means that $p = x^2 + 4y^2$.

Exercise 3.3
(i) Let p be as before but define

$$S_1 = \{(x, y, z) \in \mathbb{Z} \times \mathbf{N} \times \mathbf{N} : x^2 + 4yz = p\}$$

and $S_2 = \{(x, y, z) \in S_1 : z > x + y\}$. Consider the Nair maps

$$\alpha : S_2 \to S_2; (x, y, z) \mapsto (-x - 2y, y, z - x - y)$$

and $\beta : S_2 \to S_2; (x, y, z) \mapsto (-x, y, z)$ or (x, z, y) according as to whether $z > y - x$ or $z < y - x$.
Then, prove α is an involution with a unique fixed point and draw the conclusion about p using β.
(ii) Let p, S_1 and S_2 be as above. Consider the subset $S_3 = \{(x, y, z) \in S_1 : z < x + y\}$. Prove that if no element of S_1 is of the form (x, y, y), then all the elements of S_1 can be collected in groups of 4 where exactly 2 are in S_2 and two in S_3. Consider the Heathbrown-Nair map

$$\theta : S_2 \to S_2; (x, y, z) \mapsto (-x - 2y, y, z - x - y)$$

to conclude that $|S_2|$ is odd and arrive at a contradiction.

Exercise 3.4
(i) For any n, prove that the permutations $\sigma = (1\ 2)$ and $\tau = (1\ 2 \cdots n)$ generate the whole of S_n.
Further, if p is a prime, show that any transposition and any p-cycle generate S_p.
(ii) For general n, and for a transposition σ and any n-cycle τ, find a necessary and sufficient condition for S_n to be generated by σ and τ.

Since any group acts faithfully on its underlying set by left multiplications, one has Cayley's theorem asserting the fact that any group is a group of permutations. More generally, if H is a subgroup of G, then G acts by left translations on the set of left cosets of H.

Lemma 3.5
If H has some finite index n in G, then H contains a normal subgroup N

whose index is a divisor of $n!$. In particular, if G is a finite group and p is the smallest prime dividing $O(G)$ and, if $H \leq G$ has index p, then it is normal in G. Thus, subgroups of index 2 are normal.

Proof.

Take N to be the kernel of the action. Thus $N \leq H$ and G/N is isomorphic to a subgroup of $Sym(G/H) = S_n$.

Now, $\exists N \leq H \leq G$ with N normal in G and $[G : N]|p!$ Since $[G : N]|O(G)$ as well, $[G : N]$ divides the GCD of $O(G)$ and $p!$ which is p since p is the smallest prime dividing $O(G)$. But, H contains N, which implies $p|[G : N]$. Thus, $[G : N] = p = [G : H]$.

Group actions are naturally used to prove Sylow's theorems and their generalisations as follows.

Theorem 3.6 (Frobenius)

Let G be a finite group and p^r be a prime power dividing the order of G. Then, there exist subgroups of order p^r in G and that these are $\equiv 1$ mod p in number.

Theorem 3.7 (Snapper)

Let G be a finite group and p be a prime such that $p^n||G|$. Suppose $H \leq G$ has order $p^m \leq p^n$. Then, :

(i) there exist subgroups of order p^n containing H and,

(ii) the number of subgroups of order p^n which contain H, is $\equiv 1$ modulo p.

Proof of 3.6

Consider the action by left multiplication of G on the set Ω of all *subsets* of G with p^r elements. Let us write $|G| = p^n d$ with $n \geq r$ and $p \not| d$. The cardinality of Ω is the binomial coefficient $\binom{p^n d}{p^r}$. It can be shown by elementary number theory that $|\Omega| \equiv p^{n-r}d$ mod p^{n-r+1} but as we observe below this also follows from some group-theoretic counting. Break up Ω into disjoint orbits, say, $\Omega = \bigcup_{i=1}^{t} G.S_i$ where $S_i \in \Omega$. We claim that for any $S \in \Omega$, $G = \cup_{g \in G} gS$. To see that this claim holds, take any $S \in \Omega$ and any $s \in S$. Now, $1 = s^{-1}s \in s^{-1}S \subset G.S$. Further, if $g \in G$, then $g = g.1 \in g.S$. Thus, the claim is true. Hence, in any orbit, the number of subsets is at least $p^{n-r}d$ and divides $|G| = p^n d$. This means that exactly one of the two things happen: either an orbit has exactly $p^{n-r}d$ elements or it has a multiple of p^{n-r+1} number of elements. Note that $|\Omega| \not\equiv 0$ mod p^{n-r+1}. Hence, not each orbit can have cardinality a multiple of p^{n-r+1}. This already proves that there are orbits with exactly $p^{n-r}d$ elements. Note that, obviously, the orbit of a set S in Ω has exactly $p^{n-r}d$ elements if, and only if, the corresponding isotropy subgroup has order p^r. But, we see that each orbit with exactly $p^{n-r}d$ elements corresponds to exactly one subgroup of order p^r and conversely. This is because in each such orbit $G.A$, there is exactly one group (that gA which contains the identity) since $G = \cup_g gA$, and so, the orbit is the set of left cosets of that group. Suppose the number of such orbits with exactly $p^{n-r}d$ elements is t, then $|S| \equiv tp^{n-r}d \bmod p^{n-r+1}$. Although, one can prove by elementary number theory that $\binom{p^n d}{p^r} \equiv p^{n-r}d \bmod p^{n-r+1}$, one could obtain this already from the above sentence as it is valid for any G of order $p^n d$ and applying this to the corresponding cyclic group, one gets $t = 1$. Hence, $\binom{p^n d}{p^r} \equiv tp^{n-r}d \equiv p^{n-r}d \bmod p^{n-r+1}$. Thus, $t \equiv 1 \bmod p$. This proves Frobenius's result.

Exercise 3.8
(i) Prove Sylow's second theorem by using group actions.
(ii) Prove lemma 1.2 using Sylow theorems.
(iii) Consider the alternating group $A(\mathbf{N})$ defined as the set of bijections of $\mathbf{N}$ which move only finitely many natural numbers and move them as even permutations. Prove that $A(\mathbf{N})$ is simple.
(iv) Prove that any finite group is isomorphic to a subgroup of a finite simple group.
(v) Show that A_n is $(n-2)$-transitive on $\{1, 2, \cdots, n\}$.
(vi) Prove that in the infinite group $SL(2, F)$ for any infinite field F, each element is expressible as a finite product of elements of finite order.

Exercise 3.9
(i) *Prove that any element in A_n is a commutator $xyx^{-1}y^{-1}$ in S_n where x is an n-cycle.*
(ii) *In S_{2n+1}, prove that the cycle $(1\ 2\ \cdots\ 2n{+}1)$ is expressible as $xyx^{-1}y^{-1}$ where x is a $n+1$-cycle.*
(iii) *In any S_n, show that every element is a product of at the most two cycles.*
(iv) *Let F be a finite field. Prove that $Sym(F)$ is generated by the permutations $\sigma : x \mapsto x^{-1}$ for $x \neq 0$; $\sigma(0) = 0$ and $\tau_{a,b} : x \mapsto ax + b$ for $a, b \in F$.*

Examples 3.10
1. The group of rotations of a cube which leave it invariant are

(I) 90 degree (clockwise or anti-clockwise) rotations about the axes joining the centres of the opposite faces - there are 6 such;
(II) 180 degree rotations about each of the above axes - there are 3 such;

(III) 120 degree (clockwise or anti-clockwise) rotations about the axes joining the opposite vertices - there are 8 such;
(IV) 180 degree rotations about the axes joining the midpoints of the opposite edges - there are 6 such and;
(V) the identity.

2. The cyclic group C_n can be regarded as the group of permutations of the vertices of a regular n-gon. That is, it is the subgroup of S_n generated by an n-cycle $(1, 2, \cdots, n)$.

3. For $n > 2$, the dihedral group D_n is defined as the group of rotations of the regular n-gon given by n rotations about an n-fold axis perpendicular to the plane of the n-gon and reflections about the n two-fold axes in the plane of the n-gon like the spokes of a wheel, where the angle between consecutive spokes is $\frac{2\pi}{n}$ or $\frac{\pi}{n}$ according as n is odd or even. It has order $2n$.
It can be regarded as a subgroup of S_n as follows. The n rotations corresponding to the powers of $\sigma = (1, 2, \cdots, n)$ and the group D_n is the subgroup

$$\{Id, \sigma, \cdots, \sigma^{n-1}, \tau, \tau\sigma, \cdots, \tau\sigma^{n-1}\}$$

where $\tau = (2, n)(3, n-1) \cdots$ So, the dihedral group D_6 is the symmetry group of the hexagon. One can represent it as the subgroup of S_6 generated by (16)(25)(34) and (123456).

Group actions can also be used to get information on the number a_n of subgroups of a given index n in a group G in the following manner.

Proposition 3.11
Let G be any group.
(i) Denote by t_n, the number of transitive actions of G on $\{1, 2, \cdots, n\}$. Then $a_n = t_n/(n-1)!$.
(ii) If $h_n = |\text{Hom }(G, S_n)|$, then one has the relation

$$a_n = h_n/(n-1)! - \sum_{k=1}^{n-1} \frac{h_{n-k}}{(n-k)!} a_k.$$

(iii) For the free group F_r, one has Hall's formula

$$a_n(F_r) = n(n!)^{r-1} - \sum_{k=1}^{n-1} (n-k)!^{r-1} a_k(F_r).$$

(iv) A d-generated group G has at most $n(n!)^{d-1}$ subgroups of any index n.
(v) The number of subgroups of index n in $\mathbb{Z}^2$ is $\sigma(n)$, the sum of the divisors

of n.
(vi) An application of (ii) to $\mathbb{Z}^2$ yields the well-known partition identity

$$np(n) = \sum_{i=1}^{n-1} \sigma(i)p(n-i) + \sigma(n).$$

Proof
(i) If $\rho : G \to S_n$ is any transitive representation, then the subset $H := \{g \in G : \rho(g)(1) = 1\}$ is a subgroup of G, of index n. Conversely, if $H \leq G$ is a subgroup of index n in G, the set G/H of left cosets has n elements and G acts transitively on it. There are $(n-1)!$ ways to identify G/H with the set $\{1, 2, \cdots, n\}$ where the coset H is identified with 1. Thus, $a_n = t_n/(n-1)!$.
(ii) For each $1 \leq k \leq n$, there are $\begin{pmatrix} n-1 \\ k-1 \end{pmatrix}$ ways to choose the orbit of 1, t_k transitive actions on it, and h_{n-k} actions on its complement. This proves the relation

$$h_n = \sum_{k=1}^{n-1} \begin{pmatrix} n-1 \\ k-1 \end{pmatrix} t_k h_{n-k} + t_n.$$

Rewriting the relation in terms of the a_n, one has

$$a_n = h_n/(n-1)! - \sum_{k=1}^{n-1} \frac{h_{n-k}}{(n-k)!} a_k.$$

(iii) is immediate from (ii).
(iv) If G is d-generated, then $h_n \leq (n!)^d$. Hence $a_n \leq \frac{h_n}{(n-1)!} \leq n(n!)^{d-1}$.
(v) Now, any subgroup H of $\mathbb{Z}^2$ is of the form $g\mathbb{Z}^2$ for some $g \in M(2, \mathbb{Z})$. It is of finite index if $g \in GL(2, \mathbf{Q})$. Note that $H = g\mathbb{Z}^2 = gx\mathbb{Z}^2$ for any $x \in GL(2, \mathbb{Z})$. We claim that x can be chosen so that gx is of the form $\begin{pmatrix} a & 0 \\ c & d \end{pmatrix}$ where $a > 0,\ 0 \leq c < d$. Now, if $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, then $\exists\, u, v \in \mathbb{Z}$ such that $au + bv = (a, b)$. Then, $x = \begin{pmatrix} u & -b/(a,b) \\ v & a/(a,b) \end{pmatrix} \in GL(2, \mathbb{Z})$ and gx is of the form $\begin{pmatrix} a_1 & 0 \\ c_1 & d_1 \end{pmatrix}$. By multiplying by matrices like $\begin{pmatrix} \pm 1 & 0 \\ 0 & \pm 1 \end{pmatrix}$, we may assume that gx is of the form $\begin{pmatrix} a_1 & 0 \\ c_1 & d_1 \end{pmatrix}$ with $a_1, d_1 > 0$. Now, multiplying on the right by a matrix of the form $\begin{pmatrix} 1 & 0 \\ l & 1 \end{pmatrix}$, we get $\begin{pmatrix} a_1 & 0 \\ c_1 + d_1 l & d_1 \end{pmatrix}$ where we may assume that $0 \leq c_1 + d_1 l < d_1$. Therefore, we have shown that any

subgroup H of finite index is of the form $g\mathbb{Z}^2$ where $g = \begin{pmatrix} a & 0 \\ c & d \end{pmatrix} \in M(2, \mathbb{Z})$, with $a, d > 0$ and $0 \leq c < d$. Clearly, the index of H is ad. We claim finally that $g\mathbb{Z}^2 = h\mathbb{Z}^2$ for another matrix $h = \begin{pmatrix} a_1 & 0 \\ c_1 & d_1 \end{pmatrix} \in M(2, \mathbb{Z})$, with $a_1, d_1 > 0$ and $0 \leq c_1 < d_1$ if, and only if, $g = h$. To see this, suppose $g^{-1}h \in GL(2, \mathbb{Z})$. This gives $a_1 = a, d_1 = d$ and $d|(c_1 - c)$. As $0 \leq c_1, c < d$, we get $c_1 = c$ as well. Therefore, the number of subgroups of index n is the number of matrices of the form $\begin{pmatrix} n/d & 0 \\ c & d \end{pmatrix}$, with $d|n$ and $0 \leq c < d$. Thus, for each divisor $d|n$, there are exactly d subgroups of index. Therefore, $a_n = \sum_{d|n} d = \sigma(n)$.

(vi) We claim that $h_n(\mathbb{Z}^2) = n!p(n)$. The reason is as follows. x can be arbitrarily chosen in S_n, and y chosen in its centraliser $C_{S_n}(x)$, so that $h_n = \sum_x |C_{S_n}(x)| = |S_n| \sum 1/|[x]| = |S_n||[x]|1/|[x]| = |S_n|p(n) = n!p(n)$. This yields us the identity

$$np(n) = \sum_{i=1}^{n-1} \sigma(i)p(n-i) + \sigma(n).$$

Exercise 3.12

Let G be a group such that the set S of all torsion elements of G is a finite set. Then, prove that S is a group.

§4 Nilpotent groups

Lemma 4.1

(i) A group G is nilpotent if, and only if, there exists a series

$$G = G_0 \supseteq G_1 \supseteq \cdots \supseteq G_n = \{1\}$$

for some n where each G_i is normal in G and G_{i-1}/G_i is contained in the center of G/G_i.

(ii) The center of any nontrivial nilpotent group is nontrivial.

(iii) All p-groups are nilpotent.

(iv) Subgroups and quotient groups of nilpotent groups are nilpotent.

(v) $B(n, \mathbf{Q})$ is not nilpotent for $n \geq 2$ whereas its normal subgroup $U(n, \mathbf{Q})$ and the quotient group $B(n, \mathbf{Q})/U(n, \mathbf{Q})$ are.

(vi) A group G is nilpotent if, and only if, $G/Z(G)$ is nilpotent where $Z(G)$ is the center of G.

Proof

(i) If G is nilpotent, then by our definition, the lower central series

$$G \geq C^1(G) \geq C^2(G) \geq \cdots C^n(G) = \{1\}$$

for some n. Evidently, $C^{i+1}(G) \lhd C^i(G)$ and $C^i(G)/C^{i+1}(G)$ is contained in the center of $G/C^i(G)$ by the very definition.

Conversely, let

$$G = G_0 \supseteq G_1 \supseteq \cdots \supseteq G_n = \{1\}$$

for some n where each G_i is normal in G and G_{i-1}/G_i is contained in the center of G/G_i.

Note that $C^0(G) = G = G_0$. We prove by induction that for any $r \leq n$, $C^r(G) \leq G_r$. If we assume $C^m(G) \leq G_m$ for $m < n$, then $C^{m+1}(G) = [G, C^m(G)] \leq [G, G_m] \leq G_{m+1}$ as G_m/G_{m+1} is contained in the centre of G/G_{m+1}. Therefore, the inductive proof follows, and shows $C^n(G) = \{1\}$.

(ii) The penultimate term $C^{n-1}(G)$ in the lower central series

$$G \geq C^1(G) \geq C^2(G) \geq \cdots C^n(G) = \{1\}$$

is contained in the center.

(iii) Note first that any p-group G has a nontrivial centre Z (this follows by using the conjugation action of the group on itself). Applying induction, we may assume that $\tilde{G} = G/Z$ is nilpotent. Evidently, $C^i(G) \leq \tilde{C}^i$ for all i where $C^i(G/Z) = \tilde{C}_i/Z$. If $C^n(\tilde{G}) = 1$, then we have $C^n(G) \leq Z$; so $C^{n+1}(G) = 1$. Thus, G is nilpotent.

(iv) It is clear that for any $H \leq G$, we have $C^r(H) \leq C^r(G)$ for every r, by induction. Thus, subgroups of nilpotent groups are also nilpotent.

Let $N \lhd G$ and let $C^n(G) = \{1\}$. Consider the series

$$G/N \geq C^1(G)N/N \geq \cdots C^n(G)N/N = \{1\}.$$

Clearly, each $C^i(G)N/N \lhd G/N$ and let us check that any successive factor $\frac{C^i(G)N/N}{C^{i+1}(G)N/N}$ is contained in the centre of $\frac{G/N}{C^{i+1}(G)N/N}$. Now, $[G/N, C^i(G)N/N] = [G, C^i(G)]N/N \leq C^{i+1}(G)N/N$. Thus, the above series shows that G/N is nilpotent by (i).

(v) Call $B = B(n, \mathbf{Q})$ and $U = U(n, \mathbf{Q})$ for simplicity.

An easy matrix computation shows that $[U, U] \leq U$ and that $[U, U]$ consists of matrices which have the entries $(1, 2), (2, 3), \cdots, (n-1, n)$ to be zero.

By induction, one can easily show that $C^r(U)$ consists of matrices whose (i, j)-th entries for $i < j \leq i + r - 1$ are all zero. Hence, $C^n(U) = \{1\}$.

Clearly, $B = TU$ with $T \cap U = \{1\}$ where $T \leq B$ is the subgroup of diagonal matrices. Also, $U = Ker(det : B \to \mathbf{Q}^*)$ is a normal subgroup. Since T is abelian, it is nilpotent.

However, B is not nilpotent since $[B, B] = U$ and $[B, U] = U$ by the same computation as above.

(vi) We already know that quotient groups of nilpotent groups are nilpotent. So, we assume that $G/Z(G)$ is nilpotent and show that G must be nilpotent. Using (i) for $G/Z(G)$, we have normal subgroups G_i of G containing $Z(G)$ such that

$$G/Z(G) = G_0/Z(G) \supseteq G_1/Z(G) \cdots \supseteq G_n/Z(G) = \{1\}$$

where G_{i-1}/G_i is contained in the center of G/G_i. Consider the series

$$G = G_0 \supseteq G_1 \cdots \supseteq G_n = Z(G) \supseteq G_{n+1} = \{1\}.$$

Evidently, G_n/G_{n+1} is contained in (in fact, equal to) the center of G/G_{n+1}. Thus, G is nilpotent.

Here is a general result interesting in its own right and to be used in the next proof.

Prime factorisation in groups 4.2

Let G be any finite group and let $g \in G$. Write $O(g) = p_1^{k_1} \cdots p_r^{k_r}$ where p_i's are distinct primes. Then, prove that every element $g \in G$ can be uniquely expressed as a product $g = g_1 \cdots g_r$ where g_i's commute pairwise and $O(g_i) = p_i^{k_i}$ for $i = 1, \cdots, r$.

Proof.

Let $n_1 = \frac{O(g)}{p_1^{k_1}}$. Then, since $p_1 \not| n_1$, we may write $ap_1^{k_1} + bn_1 = 1$ for some integers a, b.

Take $g_1 = g^{bn_1}$ and $g_2 = g^{ap_1^{k_1}}$.

Of course, $g_1 g_2 = g_2 g_1 = g$. We shall check that g_1, g_2 have orders $p_1^{k_1}$ and n_1 respectively.

Now, $g_1^{p_1^{k_1}} = g^{bn_1 p_1^{k_1}} = g^{bO(g)} = 1$; so $O(g_1) | p_1^{k_1}$.

Moreover, if $g_1^d = 1$, then we get $g^{bn_1 d} = 1$; so $O(g) | bn_1 d$. This gives $p_1^{k_1} | bd$; however $(p_1, b) = 1$ as seen from $ap_1^{k_1} + bn_1 = 1$. Thus, $p_1^{k_1} | d$ which proves that $O(g_1) = p_1^{k_1}$.

Similarly, $g_2^{n_1} = g^{ap_1^{k_1} n_1} = 1$ so that $n_1 | O(g_2)$. If $g_2^l = 1$, then $g^{ap_1^{k_1} l} = 1$ so that $O(g) = p_1^{k_1} n_1 | ap_1^{k_1} l$.

Hence $n_1 | l$ as $(n_1, a) = 1$. This implies that $O(g_2) = n_1$.

Now, we may proceed by induction with g replaced by g_2. Ultimately, we

will have elements $x_i, i \leq r$ which are powers of g and have orders $p_i^{k_i}$ such that $g = x_1 \cdots x_r$.
Finally, we show that such an expression is unique. If $g = y_1 \cdots y_r$ is another expression where y_i has order $p_i^{k_i}$ and commute pairwise, then they commute with g. As x_i's are powers of g, the x_i's and the y_j's all commute pairrwise. So,

$$x_1^{-1} y_1 = x_2 \cdots x_r y_2^{-1} \cdots y_r^{-1}$$

has order dividing $p_1^{k_1}$ as seen from the left side and at the same time has order dividing $p_2^{k_2} \cdots p_r^{k_r}$ as seen from the right side. Thus, $x_1 = y_1$. Proceeding by induction, we get $x_i = y_i$ for all i.

Lemma 4.3
For a finite group G, the following are equivalent:
(i) G is nilpotent,
(ii) every subgroup is subnormal,
(iii) every proper subgroup H is properly contained in $N_G(H)$,
(iv) all maximal subgroups are normal,
(v) all p-Sylow subgroups are normal,
(vi) elements of coprime order commute and,
(vii) G is the direct product of its Sylow subgroups.
Proof.
Let G be nilpotent and $H \leq G$. Suppose

$$G \geq C^1 \geq C^2 \geq \cdots C^n = \{1\}$$

be its lower central series. Since, we have that C^i/C^{i+1} is contained in the centre of G/C^{i+1}, it follows that

$$G = GH \geq C^1 H \geq \cdots C^n H = H$$

is a chain of subgroups where each term is normal in the previous one. Thus, H is subnormal in G; that is, (i) implies (ii).
Now, suppose (ii) holds. Let $H < G$. We have a series

$$H = H_0 \lhd H_1 \lhd \cdots H_r = G.$$

If i is the smallest number for which $H_i > H$, it follows that $H_i \leq N_G(H)$ and $H_i \not\leq H$. This proves (ii) implies (iii).
For a maximal subgroup, $M \neq N_G(M)$ shows that $N_G(M) = G$ i.e., M is normal. Thus (iii) implies (iv).
Assume (iv). If P is a p-Sylow subgroup which is not normal, then $N_G(P) \leq$

M for some maximal subgroup M. Since M is normal, we get for any $g \in G$, gPg^{-1} is again a p-Sylow subgroup of $gMg^{-1} = M$. Thus, by Sylow's second theorem applied to M, there is some $m \in M$ with $gPg^{-1} = mPm^{-1}$ i.e., $m^{-1}g \in N_G(P) \leq M$. Hence $g \in M$, a contradiction, since a maximal subgroup, by definition, is a proper subgroup. Hence (iv) implies (v).

Assume (v) holds; that is, the Sylow subgroups are normal. We shall show first that elements of orders powers of different primes commute. Let $O(x) = p^r$ and $O(y) = q^s$ where $p \neq q$ are primes. If P, Q are the unique p-Sylow subgroup and the unique q-Sylow subgroup, then $x \in P$ and $y \in Q$. As $xyx^{-1}y^{-1} \in P \cap Q = \{1\}$, we have $xy = yx$.

Now, we deal with the general case. We saw in 4.2 that in any finite group, every element can be written as commuting elements of prime power orders dividing the order of the element. Let $g, h \in G$ be elements in our group which have coprime orders m, n. Then, $g = x_1 \cdots x_r, h = y_1 \cdots, y_s$ where x_i's commute among themselves, y_j's commute among themselves and each of them has prime power order. Also $O(x_i)|O(g) = m$ and $O(y_j)|O(h) = n$ which are coprime. Thus, each x_i and each y_j commute. Hence $gh = hg$. Thus, (v) implies (vi).

Suppose now that elements of coprime orders commute. We write

$$|G| = p_1^{k_1} \cdots p_r^{k_r}.$$

Let P_i be any p_i-Sylow subgroup of G ; $1 \leq i \leq r$. Then, since $[P_i, P_j] = \{1\}$ for $i \neq j$, the product $P_1 \cdots P_r$ is a group. As the orders of P_i's are pairwise coprime to each other,

$$|P_1 \cdots P_r| = |P_1| \cdots |P_r| = |G|.$$

Hence $G = P_1 \cdots P_r$ and $P_i \cap \prod_{j \neq i} P_j = \{1\}$, which means that $G \cong P_1 \times \cdots \times P_r$.

So, we have proved that (vi) implies (vii).

Finally, since p-groups are nilpotent, (vii) clearly implies that $P_1 \times \cdots P_r$ is nilpotent; that is, (i) follows.

The *Frattini subgroup* $\Phi(G)$ of a finite group G is defined to be the intersection of all (proper) maximal subgroups.

Lemma 4.4

(i) $\Phi(G)$ is the set of 'nongenerators' of G i.e., those elements which can be dropped from any generating set for G.

(ii) $\Phi(G)$ is a characteristic subgroup.

(iii) $\Phi(G)$ is nilpotent.

(iv) For any p-group G, $\Phi(G) = [G, G] < G^p >$.
(v) For a finite abelian p-group A, $A/\Phi(A)$ is an elementary abelian group of order $p^{d(A)}$ where $d(A)$ is the minimal number of generators needed to generate A.

Proof

(i) Let $G =< S >$ and $s \in S \cap \Phi(G)$. It suffices to show that $s \in< S \setminus \{s\} >$. Now, if $H =< S \setminus \{s\} >< G$, then there is some maximal subgroup M of G containing H. But then $S \setminus \{s\}$ is contained in M as well as $s \in M$. This means that $G =< S >\leq M$, which is a contradiction. Thus, we have shown that elements of $\Phi(G)$ can be dropped from any generating set for G.
Conversely, suppose $g \in G$ be such that whenever $< S \cup \{g\} >= G$, we have $< S >= G$. Let, if possible, M be a maximal subgroup not containing g. Then, $< M \cup \{g\} >= G$. By the hypothesis, this gives $< M >= M = G$, a contradiction. Hence all maximal subgroups contain g and (i) is proved.
(ii) Since $\Phi(G)$ is the intersection of all maximal subgroups of G, and since any automorphism permutes maximal subgroups of G, it follows that $\Phi(G)$ is a characteristic subgroup of G.
(iii) By the previous problem, it suffices to prove that the p-Sylow subgroups of $\Phi(G)$, for any p, are normal in it. By the Frattini argument (problem 15 (ii)), we have $G = \Phi(G)N_G(P)$.
In particular, $N_G(P) \cup \Phi(G)$ generate G which implies that $N_G(P) =< N_G(P) >= G$; that is, $P \lhd G$.
In particular, $P \trianglelefteq \Phi(G)$.
(iv) Let $M_1 \cdots, M_r$ be the (proper) maximal subgroups of G. Consider their images in the elementary abelian group $G/[G, G] < G^p >$. They are maximal subgroups and every maximal subgroup of $G/[G, G] < G^p >$ is the image of one of these. Of course, two different M_i's may have the same image. Note that the intersection of all maximal subgroups in an elementary abelian p-group is trivial. This is so because an elementary abelian p-group is isomorphic to $\mathbb{Z}/p \times \cdots \mathbb{Z}/p$, and in this group, the subproducts with one factor being the trivial groups are maximal subgroups and intersect in the identity. Therefore,

$$\Phi(G) := \bigcap_{i=1}^{r} M_i \leq [G, G] < G^p > .$$

Conversely, note that in p-groups, all maximal subgroups are of index p. This means that in our p-group G, $< G^p >\leq \Phi(G)$. Also, since maximal subgroups M in G are necessarily normal, G/M is an abelian group (of order

p) and so, $[G,G] \leq M$. Thus, $[G,G] \leq \Phi(G)$ as well. This proves that

$$\Phi(G) = [G,G] < G^p > .$$

(v) By (iv), it follows that the Frattini subgroup of an abelian p-group A is simply A^p. Therefore, clearly $A/\Phi(A) = A/A^p$ is an elementary abelian p-group. If A/A^p has order p^r, then it can be generated by r elements $x_1A^p, \cdots, x_rA^p$ say. Therefore, $A =< A^p, x_1, \cdots, x_r >$.
However, A^p is the set of nongenerators and, therefore, $A =< x_1, \cdots, x_r >$ which implies that $r \geq d(A)$. However, it is obvious that for any abelian p-group A, the elementary abelian p-group A/A^p is a direct sum of at the most $d(A)$ copies of the group of order p; so $|A/A^p| = p^r \leq p^{d(A)}$. Therefore, we have $d(A) = r$.

Proposition 4.5
Any subgroup of a finitely generated nilpotent group is also finitely generated. Therefore, a finitely generated nilpotent torsion group is finite.
Proof
Let S be a finite set of generators of a nilpotent group G. Consider the sets $S_0 = S$, $S_{i+1} = \{aba^{-1}b^{-1} : a \in S, b \in S_i\}$ for all $i \geq 0$. Of course, each S_i is a finite set. Since G is nilpotent, there is some n so that $S_r = 1$ for all $r \geq n$. For each $i \geq 0$, consider the subgroup G_i of G generated by the set $\cup_{r \geq i} S_r$. As $S_r = 1$ for $r \geq n$, each G_i is finitely generated. It is clear that $[G, G_i] = G_{i+1}$ since S, S_i, S_{i+1} generate G, G_i and G_{i+1} and $S_{i+1} = \{aba^{-1}b^{-1} : a \in S, b \in S_i\}$. In particular, each G_i is normal in G and the subgroup G_i/G_{i+1} is central in G/G_{i+1}. Now, if H is any subgroup of G, then $(H \cap G_i)/(H \cap G_{i+1})$, being a subgroup of the finitely generated abelian group G_i/G_{i+1}, is finitely generated. Since $H \cap G_n = 1$, this gives inductively that each $H \cap G_i$ is finitely generated. In particular, $H = H \cap G_0 = H \cap G$ is finitely generated.
To prove the second assertion, look at the lower central series of G. As $C^n(G)/C^{n+1}(G)$ is a subgroup of the finitely generated, nilpotent group $G/C^{n+1}(G)$, this subgroup is finitely generated as well. But, this is an abelian, torsion group and is, hence, finite. Thus, G itself is finite.

The existence of fixed-point free automorphisms on finite groups have very interesting implications on the abelianness or nipotency of a group as the following proposition shows.

Proposition 4.6
Let T be an automorphism of prime order p of a finite group G such that $T(x) = x$ if, and only if, $x = 1$. Then, we have :

(i) The function $F(g) := g^{-1}T(g)$ is a bijection on G.
(ii) For any g in G, the product $g\ T(g)\ T^2(g) \dots T^{p-1}(g)$ is the identity.
(iii) $|G|$ is congruent to 1 modulo p.
(iv) For any prime q dividing the order of G, there is a q-Sylow subgroup Q of G such that $T(Q) = Q$.
(v) For any prime q, prove there is a unique q-Sylow subgroup Q of G such that $T(Q) = Q$.
(vi) Let q be a prime. Then, the q-Sylow subgroup Q fixed by T contains any q-subgroup of G fixed by T.
(vii) For $p = 2$, G is abelian.
(viii) For $p = 3$, G is nilpotent.
The analogue of (viii) was proved for any prime p by the Fields medalist J.Thompson in his thesis.

Proof

(i) Now, $x^{-1}T(x) = y^{-1}T(y)$ for some x, y if, and only if T fixes yx^{-1}; that is, when $x = y$. Thus, the map $x \mapsto x^{-1}T(x)$ is 1-1. Being a map of finite sets, this is onto as well.
(ii) From (i), an arbitrary element of g can be written as $g = x^{-1}T(x)$ so that we get

$$gT(g) \cdots T(g)^{p-1} = 1.$$

(vii) Taking $p = 2$, this gives $T(g) = g^{-1}$ for all g. As T is an automorphism, G is abelian.
(iii) Consider the subgroup Z of $\mathrm{Aut}(G)$ generated by T. This is a cyclic group of order p acting on G. Each orbit has cardinality dividing p and, as T fixes only the identity element, each orbit other than that of 1 has p elements. So, $|G|$ has 1 mod p number of elements.
(iv) Look at the action of Z on the set S of all q-Sylow subgroups of G. Again, the orbits which are not fixed points under Z, have order p. So, if there is no fixed point in S, then the number of elements in S would be a multiple of p. But, we know that the number of elements in S (i.e. the number of q-Sylow subgroups) is a divisor of $|G|$. Thus, p would divide $|G|$, a contradiction of (iii).
(v) Note that if $T(Q) = Q$, then $T(N) = N$ where N is the normaliser of Q in G. Now, if T fixes another q-Sylow subgroup gQg^{-1}, then rewriting $T(gQg^{-1} = gQg^{-1}$, we have $g^{-1}T(g)$ belongs to N. But, applying (i) to N itself, any element of N is of the form $x^{-1}T(x)$ for some x in N. So, $g^{-1}T(g) = x^{-1}T(x)$, which gives $g = x$ belongs to N. So, $gQg^{-1} = Q$.
(vi) Let R be any q-group in G which is fixed by T. Let S be a q-group which contains R and is maximal with respect to this property (i.e., S is fixed by

T, contains R and is not contained in a strictly larger q-group which is T-fixed). We claim that $R = Q$. For this, first look at the normaliser $N(S)$ of S in G. Since $T(S) = S$, we have also $T(N(S)) = N(S)$ clearly. by (v) applied to the subgroup M of $N(S)$ which is T-fixed. Now, S is a q-subgroup of $N(S)$, say, yMy^{-1} where y is in $N(S)$. But then $S = y^{-1}Sy$ is contained in M. By maximality property of S, we get $S = M$ i.e., S is a q-Sylow subgroup of $N(S)$. But, any Sylow subgroup of a subgroup in any group is the intersection of the subgroup with a Sylow subgroup of the big group. So, we have a q-Sylow subgroup L of G such that $S = N(S)$ intersection L. In other words, $S = N_L(S)$, the normaliser of S in L. But, in a q-group (indeed, in any nilpotent group), the normaliser of a proper subgroup contains the subgroup properly. So, we have $S = L$. In other words, S is a q-Sylow subgroup of G itself. By uniqueness of the T-fixed q-Sylow subgroup of G itself. Thus, we have shown that R is contained in $S = P$.
(viii) Finally, we prove the result for $p = 3$.
Now $x\ T(x)\ T^2(x) = e$ for any x in G. So, $T(x^{-1}x^{-1} = (x\ T(x))^{-1} = T^2(x)$ from the above. Putting $y = x^{-1}$, we get $T(y)y = T^2(y^{-1}) = (T^2(y))^{-1}$. Thus, both $yT(y)$ and $T(y)y$ are equal to $(T^2(y))^{-1}$. In other words, every element x in G commutes with the element $T(x)$. Similarly, x commutes with $T^2(x)$ also.
Now, to prove the result, let us consider any prime q and the unique q-Sylow subgroup Q of G which is T-fixed. Let g be any element of G which has q-power order. Consider the subgroup Q' is generated by the three elements $g, T(g)$ and $T^2(g)$. Clearly, Q' is T-fixed. Also, since $g, T(g), T^2(g)$ commute among themselves, Q' is a q-group. Therefore, by (vi), Q' is contained in Q. Thus, Q contains all elements of G of q-power order. Hence, Q is the unique q-Sylow subgroup of G. So, it is normal. This proves the result for $p = 3$.

§5 **Solvable groups**

Lemma 5.1
(i) A group G is solvable if, and only if, there exists a series

$$G = G_0 \supseteq G_1 \supseteq \cdots \supseteq G_n = \{1\}$$

for some n where G_i are normal subgroups of G and G_i/G_{i+1} is abelian.
(ii) Subgroups and quotient groups of solvable groups are solvable.
(iii) If N is a solvable, normal subgroup of a group G such that G/N is solvable, then prove that G is solvable as well.

Proof
(i) Evidently, the derived series is a series as asserted and, so any solvable group has such a series.
Conversely, suppose G admits a finite series as above. We prove by induction that $D^r(G) \leq G^r$ for all r. That would prove $D^n(G) = \{1\}$. The assertion holds for $r = 0$. Supposing $D^r(G) \leq G^r$, we have

$$D^{r+1}(G) = [D^r(G), D^r(G)] \leq [G^r, G^r] \leq G^{r+1}$$

as G_r/G_{r+1} is abelian. This proves (i).
(ii) & (iii) Suppose G is solvable. Start with a series for G as in (i). The subgroups $H_r = G_rN/N$ form a series for G/N as in (i), and since $H_n = \{1\}$, it follows that G/N is solvable. Of course, any subgroup K of G must be solvable as the terms of its derived series are contained in the corresponding term of the derived series for G.
Conversely, suppose $N \trianglelefteq G$ and G/N are solvable. Let

$$N = N_0 \supseteq N_1 \supseteq \cdots \supseteq N_r = \{1\}$$

and

$$G/N = G_0/N \supseteq G_1/N \supseteq \cdots \supseteq G_s/N = \{1\}$$

be series for N and G/N as in (i). But then

$$G = G_0 \supseteq \cdots \supseteq G_s = N \supseteq N_1 \cdots N_r = \{1\}$$

is a series for G as in (i). Thus, G is solvable.

Remark (Odd order and Burnside theorems) 5.2 :
One of the most famous, beautiful theorems in finite group theory is the statement that *every group of odd order is solvable.* The proof by Walter Feit and John Thompson occupies a whole volume of the Pacific Journal of Mathematics apart from using many results proved by others earlier.
A theorem of Burnside asserts that any group of order p^rq^s for primes p, q is solvable.

Exercise 5.3
(i) Let G be a finite group such that all its proper subgroups are abelian. Then, G must be solvable.
(ii) Let G be a finite group such that all its proper subgroups are nilpotent. Then, again G must be solvable.

§6 Matrix groups

As we pointed out earlier, the usefulness of a group in a situation depends on its avataar. A group is especially useful when it appears as a subgroup of $\mathrm{GL}(n, \mathbb{C})$.

Lemma 6.1
$SU(2) \cong H^1$, the group of unit quaternions; that is, the quaternions $a+bi+cj+dk$ which satisfy $a^2+b^2+c^2+d^2=1$.
Further, there exists a homomorphism from $SU(2)$ to $SO(3)$ whose kernel is $\{\pm I_2\}$.
Proof.
Any element of $SU(2)$ looks like $\begin{pmatrix} a+ib & c+id \\ -c+id & a-ib \end{pmatrix}$ with $a,b,c,d \in \mathbf{R}$ such that $a^2+b^2+c^2+d^2=1$.
This reminds us of Hamilton's quaternions. Therefore, let us define the map

$$\theta : SU(2) \to H \; ; \begin{pmatrix} a+ib & c+id \\ -c+id & a-ib \end{pmatrix} \mapsto a+ib+cj+dk.$$

Note that the image is contained in the group H^1 of 'unit' quaternions; that is, $a+ib+cj+dk$ for which $a^2+b^2+c^2+d^2=1$ or, equivalently,

$$(a+ib+cj+dk)^{-1} = a-bi-cj-dk.$$

It is trivial to check that θ is a homomorphism. Also, clearly it is 1-1 as well as onto H^1. Therefore, $SU(2) \cong H^1$.
Now, as the group of nonzero elements H^* acts on the 3-dimensional real vector space V generated by i,j,k by conjugation, we may think of $SU(2)$ as acting on this space. Thus, we have a homomorphism

$$\rho : SU(2) \to GL(V).$$

To explicitly evaluate it, we use the isomorphism θ to view $SU(2)$ as H^1. For any $g = \begin{pmatrix} a+ib & c+id \\ -c+id & a-ib \end{pmatrix} \in SU(2)$, we may compute qiq^{-1}, qjq^{-1} and qkq^{-1} where $q = \theta(g)a+bi+cj+dk$. We arrive at the following matrix with respect to the ordered basis $\{i,j,k\}$:

$$\rho(g) = \begin{pmatrix} a^2+b^2-c^2-d^2 & 2(bc-ad) & 2(ac+bd) \\ 2(ad+bc) & a^2-b^2+c^2-d^2 & 2(cd-ab) \\ 2(bd-ac) & 2(ab+cd) & a^2-b^2-c^2+d^2 \end{pmatrix}.$$

Now, since $g^{-1} = \overline{g^t}$ is obtained by changing b,c,d to theire negatives, it is clear that $\rho(g)^{-1} = \rho(g)^t$; that is, $\rho(g) \in O(3, \mathbf{R})$.

Note that if $g \in Ker\rho$, then the corresponding quaternion q commutes with i, j, k and, hence, with the whole of H.
In particular, this gives $g \in Z(SU(2)) = \{\pm I\}$.
Of course, it is clear that $-I$ does indeed belong to Ker ρ.
The final assertion left is to show that the determinant of $\rho(g)$ is always 1 is proved by direct (but messy) calculation. A better way would be to use a little bit of topology which shows that $SU(2)$ is connected, ρ on $SU(2)$ and the determinant function on $GL(3, \mathbf{R})$ are continuous, that the image of det $\circ\rho$ is a connected subset of $\mathbf{R}$ contained in $\{\pm 1\}$ and containing I.

Bruhat lemma 6.2
For any field K, $GL(n, K)$ is the disjoint union of double cosets BwB over permutation matrices w. Here $B := B(n, K)$, the group of invertible upper triangular matrices with entries from K.

Minkowski lemma and more 6.3
(i) For any odd prime number p, an $n \times n$ integral matrix $A \neq Id$ with entries congruent to the corresponding entries of the identity matrix modulo p is of infinite order. In other words, the 'principal congruence subgroup' Ker $(GL(n, \mathbb{Z}) \to GL(n, \mathbb{Z}/p)$ is torsion-free for an odd prime p.
(ii) For $p = 2$, any matrix A in Ker $(GL(n, \mathbb{Z}) \to GL(n, \mathbb{Z}/p))$ of finite order, has order 1 or 2 and is expressible as

$$M.diag(1, 1, \cdots, 1, -1, -1, \cdots, -1).M^{-1}$$

for some $M \in GL(n, \mathbb{Z})$.
Proof of (i).
Write $A = I + pB$, with $B \in M(n, \mathbb{Z})$. The characteristic polynomial of A is $\chi_A(t) = $ det $(tI - A) = $ det $(t-1)I - pB)$. Therefore, $\chi_A(\alpha) = 0$ if, and only if, $\chi_B(\frac{\alpha-1}{p}) = 0$. Now, if A has finite order, its eigenvalues are all roots of unity. Since $|\alpha| = 1$, we have

$$|\frac{\alpha - 1}{p}| \leq \frac{|\alpha| + 1}{p} \leq \frac{2}{p} < 1.$$

Thus, all the roots of $\chi_B(t)$ have absolute values < 1. Since $\chi_B(t)$ is an integral polynomial whose top coefficient is 1, this implies that $\chi_B(t) = t^n$; that is, all the eigenvalues of B are zero. In other words, all eigenvalues of A are 1. As A has finite order, it must be the identity matrix.

Proposition 6.4
(i) The matrices $S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ and $X = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ generate $SL(2, \mathbb{Z})$.

(ii) $PSL(2,\mathbb{Z})$ is isomorphic to the free product $\mathbb{Z}/2 * \mathbb{Z}/3$.
(iii) $SL(2,\mathbb{Z}) =< x, y|x^2y^{-3}, x^4 >$.
(iv) The abelianisation of $SL(2,\mathbb{Z})$ is the cyclic group of order 12.

Proof

(i) We shall prove, equivalently, that S and $A := S^{-1}X^{-1}$ generate $SL(2,\mathbb{Z})$. Note that $AS = Y := \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$.

Now, start with any $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2,\mathbb{Z})$. We shall show that left and right mutliplications by powers of X and Y lead to $\pm I$ by the usual Euclidean division algorithm.

For any integer l, we have $X^l g = \begin{pmatrix} a+lc & b+ld \\ c & d \end{pmatrix}$. This shows us that one can divide a by c and replace a by its residue mod c.

Similarly, one can see that by left multiplication by some Y^l, one can reduce c mod a. Repeating these finitely many times, the division algorithm implies that one of a and c becomes zero; the other has to be ± 1 as the determinant is always 1.

So, g becomes $g_1 = \begin{pmatrix} 0 & \pm 1 \\ \mp 1 & d \end{pmatrix}$ or $g_2 = \begin{pmatrix} \pm 1 & b \\ 0 & \pm 1 \end{pmatrix}$.

Now,

$$g_1 X^{\pm d} = \begin{pmatrix} 0 & \pm 1 \\ \mp 1 & 0 \end{pmatrix} = S^{\mp 1}$$

and $g_2 = X^b$ or $-X^{-b}$.

Since $-I = S^2$, the assertion follows.

(ii) Since $S^2 = -I$ also represents the identity element in $PSL(2,\mathbb{Z})$, the image s of S has order 2 in $PSL(2,\mathbb{Z})$.

Also, the image b of $B := SX = \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix}$ in $PSL(2,\mathbb{Z})$ has order 3 as $(SX)^3 = -I$.

We know that the elements s, b generate the whole group; so, we need only show that no matrix

$$SB^{a_1}SB^{a_2}\cdots SB^{a_r}$$

with each a_i either 1 or 2, can be the matrices $I, -I$.

Since $SB = -X$ and $SB^2 = Y$, it follows that any word in the positive powers of SB and SB^2 is a matrix $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ in which a, b, c, d are of the same sign. Therefore, if $b \neq 0$, then the corresponding entry $-b-d$ of SBg and b of SB^2g are non-zero as well. Similarly, if $c \neq 0$, the corresponding entries of SBg and SB^2g are nonzero. Since SB and SB^2 have the property

that either the $(1,2)$-th entry or the $(2,1)$-th entry is non-zero, any word g in their positive powers has this property; hence g can not be the identity matrix.
(iii) follows from (ii) by sending x to S and y to SX.
(iv) Now,

$$SL(2,\mathbb{Z})_{ab} =< x,y|x^2y^{-3},x^4,xyx^{-1}y^{-1} >=< e,f|2e-3f,ef-fe,4e >$$

$$\cong (\mathbb{Z}e \otimes \mathbb{Z}f)/ < 2e-3f,4e > .$$

The invariant factors of the subgroup $< 2e-3f, 4e >$ above are the invariant factors of the matrix $A = \begin{pmatrix} 2 & -3 \\ 4 & 0 \end{pmatrix}$. The latter is computed by computing $h_1(A) = 1 = d_1$ and $h_2(A) = 12 = d_1d_2$. Clearly, $d_1 = 1$ and $d_2 = 12$. Therefore, the abelianisation of $SL(2,\mathbb{Z})$ is the cyclic group of order 12.

§7 Miscellaneous results

Lemma 7.1
The groups $\mathbb{C}^*$ and S^1 are isomorphic.
Proof.
First of all, the polar co-ordinates provide an isomorphism $\mathbf{C}^* \cong \mathbf{R}^{>0} \times S^1 \cong \mathbf{R} \times S^1$. So, it suffices to show that $\mathbf{R} \times S^1 \cong S^1$.
Start with any $\mathbf{Q}$-vector space basis $\mathcal{B}$ of $\mathbf{R}$ which contains 1. Already, we have used Zorn's lemma here when trying to extend the linearly independent set $\{1\}$ to a basis. Then, the set

$$\mathcal{C} := (\mathcal{B} \times \{0\}) \bigcup (\{0\} \times \mathcal{B})$$

is a basis of the $\mathbf{Q}$-vector space $\mathbf{R} \times \mathbf{R}$. Once again, it is a consequence of Zorn's lemma that $\mathcal{B}$ and $\mathcal{C}$ are in bijection. Let $\theta : \mathcal{C} \to \mathcal{B}$ be a bijection which maps $(0,1)$ to 1. Then, there is a unique extension of θ to a $\mathbf{Q}$-vector space isomorphism from $\mathbf{R} \times \mathbf{R}$ onto $\mathbf{R}$. Since $\theta(0,1) = 1$ and is a $\mathbf{Q}$-vector space isomorphism, it gives an isomorphism of $\{0\} \times \mathbb{Z}$ onto $\mathbb{Z}$. Therefore, θ gives

$$(\mathbf{R} \times \mathbf{R})/(\{0\} \times \mathbb{Z}) \cong \mathbf{R}/\mathbb{Z}.$$

Evidently, the left hand side above is isomorphic to $\mathbf{R}/\{0\} \times \mathbf{R}/\mathbb{Z}$. This completes the proof, since $t \mapsto e^{2i\pi t}$ gives an isomorphism from $\mathbf{R}/\mathbb{Z}$ onto S^1.

Ore's covering lemma 7.2
Let $H, K \leq G$ be subgroups of the same finite index. Then, there exist $g_1, \cdots, g_n$ such that

$$G = \sqcup_{i=1}^n g_i H = \sqcup_{i=1}^n K g_i.$$

Even the case $H = K$ is interesting.
Proof.
Writing G as a union of the double cosets KgH, it is necessary and sufficient to write each double coset in the form asserted since each double coset KgH is a union of right cosets of K in G and also a union of left cosets of H in G.
We note the two bijections :

$$KgH/H \to K/(K \cap gHg^{-1});$$

$$kgH \mapsto k(K \cap gHg^{-1})$$

$$K\backslash KgH \to (H \cap g^{-1}Kg)\backslash H;$$

$$Kgh \mapsto (H \cap g^{-1}Kg)h.$$

Here, the notation is explained as follows. If B is a subgroup of a group A and if S is a subset of A which is a union of left cosets of B in A, then one has written S/B to denote the set of left cosets of B contained in S. Similarly, if T is a subset of A which is a union of right cosets of B, then one has written $B\backslash T$ to denote the set of right cosets of B contained in T.
It is easy to verify that the above are bijections. Note that the right sides of these bijections count cosets of subgroups in groups (and not merely sets).
If there is a bijection,

$$K/(K \cap gHg^{-1}) \to (H \cap g^{-1}Kg)\backslash H$$

say

$$k_i(K \cap gHg^{-1}) \mapsto (H \cap g^{-1}Kg)h_i$$

where $i = 1, \cdots, r$, then we see that

$$KgH = \sqcup_{i=1}^r k_i g h_i H = \sqcup_{i=1}^r K k_i g h_i.$$

Therefore, we only have to establish a bijection between $K/(K \cap gHg^{-1})$ and $(H \cap g^{-1}Kg)\backslash H$.
Now, we note that $K/(K \cap gHg^{-1})$ is in bijection with $g^{-1}Kg/(g^{-1}Kg \cap H)$.
Now, since H, K have the same finite index in G, so do the subgroups H

and $g^{-1}Kg$ of G. Thus, it suffices to show that the subgroup $g^{-1}Kg \cap H$ has finite index. But, this is true since both H and $g^{-1}Kg$ do.

Frobenius's theorem 7.3

Let G be a finite group and, for any natural number n, denote by $f_n(G)$ the cardinality of the set $\{x \in G : x^n = 1\}$. Then, :

(i) The number of elements of order n is a multiple of $\phi(n)$. In particular, for a prime p such that $p^k \| |G|$, one has $f_{p^k}(G) - f_{p^r}(G) \equiv 0 \bmod p^r$ for all $r < k$. More particularly, if $f_{p^k}(G) \equiv 0 \bmod p^k$, then $f_{p^r}(G) \equiv 0 \bmod p^r$ for all $r < k$.

(ii) Let $p^k \| n$ and let Q be a set of representatives of conjugacy classes of elements y such that $y^{|G|/p^k} = 1$. Then, $f_n(G) = \sum_{y \in Q}[G : C_G(y)]f_{p^k}(C_G(y))$.

(iii) If $p^k \| |G|$, then $f_{p^k} \equiv 0 \bmod p^k$.

(iv) If $n/|G|$, then $f_n(G) \equiv 0 \bmod n$.

Proof

(i) The relation $x \sim y \Leftrightarrow \langle x \rangle = \langle y \rangle$ is an equivalence relation. The equivalence class of x looks like $\{x^i/(i, O(x)) = 1\}$; it has $\phi(O(x))$ elements. Writing the set of elements of order n as a union of equivalence classes (of elements of order n), the assertion follows. Further, if $p^k/|G|$ and $r < k$, then $f_{p^k}(G) - f_{p^r}(G)$ is the number of elements of orders $p^{r+1}, \ldots, p^k$ in G. As $\phi(p^{r+1}) = p^r(p-1), \phi(p^{r+2}) = p^{r+1}(p-1)$ etc., we have that the number of elements of orders $p^{r+1}, \ldots, p^k$ is certainly a multiple of p^r. The last assertion is obvious.

(ii) By 4.2, we have each $g = xy$ uniquely where $xy = yx$ and $O(x)$ is a power of p and $(O(y), p) = 1$. Thus, $g \mapsto (x, y)$ gives a bijection between G and ordered pairs (x, y) for which $(O(y), p) = 1$ and $x \in C_G(y)$ has p-power order. So, $g^n = 1 \Leftrightarrow x^{p^k} = 1 = y^{\frac{n}{p^k}}$.

Hence $f_n(G) = \sum\limits_{y \in G} f_{p^k}(C_G(y))$ is constant where y varies through its conjugacy class, and as this conjugacy class has $[G : C_G(y)]$ elements, $y^{n/p^k} = 1$ we have $f_n(G) = \sum\limits_{y \in Q}[G : C_G(y)]f_{p^k}(C_G(y))$.

(iii) The proof is by induction on $|G|$. It is vacuously true for $G = \{1\}$. Assume for all proper subgroups H that $f_{p^r}(H) \equiv 0 \bmod p^r$ where $p^r \| |H|$. Then, by (i), $f_{p^l}(H) \equiv 0 \bmod p^l$ if $p^l/|H|$.

Now, apply (ii) with $n = |G|$. We get

$$\begin{aligned} |G| &= f_n(G) = \sum_{y \in Q}[G : C_G(y)]f_{p^k}(C_G(y)) \\ &= \sum_{y \in Q \cap Z(G)} f_{p^k}(C_G(y)) + \sum_{y \in Q \setminus Z(G)} [G : C_G(y)]f_{p^k}(C_G(y)) \end{aligned}$$

$$= \quad |Q \cap Z(G)| f_{p^k}(G) + \sum_{y \in Q \setminus Z(G)} [G : C_G(y)] f_{p^k}(C_G(y)).$$

Look at any of the proper subgroups $C_G(y)$ for $y \in Q\setminus Z(G)$. If $p^{\ell}\||C_G(y)|$, with $\ell \leq k$, then $[G : C_G(y)] \equiv 0 \ mod \ p^{k-\ell}$ and, $p^{\ell}|f_{p^{\ell}}(C_G(y))$ by induction hypothesis.
Note $f_{p^{\ell}}(C_G(y)) = f_{p^k}(C_G(y))$. Therefore each term in the sum above is a multiple of p^k. As $p^k||G|$ we have $p^k||Q \cap Z(G)|f_{p^k}(G)$. But $p \not| |Q \cap Z(G)|$ for, if it did then it would have an element y of order p which would also satisfy $y^{\frac{n}{p^k}} = 1$. This is a contradiction, as $(\frac{n}{p^k}, p) = 1$. Therefore $p^k|f_{p^k}(G)$.
(iv) Let $n||G|$. We write $= p_1^{k_1} \dots p_r^{k_r}$ for distinct primes p_i. Then, it suffices to show that $p_i^{k_i}/f_n(G)$. Now, by (ii), $f_n(G)$ is expressible as a sum of terms of the form $[G : H]f_{p_i^{k_i}}(H)$ for subgroups H. By (iii), each such H (including G itself) satisfies $p_i^r/f_{p_i^r}(H)$ where $p_i^r\||H|$. Therefore $p_i^{k_i} = p_i^{k_i - r}$. p_i^r divides $[G : H]f_{p_i^r}(H) \ \forall \ H$. Thus $p_i^{k_i}$ divides the sum which is $f_n(G)$.

Examples of presentations 7.4

1. $\mathbb{Z} = \langle x \mid \phi \rangle$; $\mathbb{Z}_n = \langle x \mid x^n \rangle$.
2. $\mathbb{Z} \oplus \mathbb{Z} = \langle x, \ y \mid [x, \ y] \rangle$.
3. $\mathbb{Z}^n = \langle x_1, \ \dots \ , \ x_n \mid \{[x_i, \ x_j] \ : \ 1 \leq i < j \leq n\} \rangle$
Note that n is the minimal number of elements needed to generate $\mathbb{Z}^n$. Moreover, there does not exist a presentation with less than $n(n-1)/2$ relations.
4. Every group has a presentation. For example,

$$G = \langle G \mid x.y.(xy)^{-1} \ : \ x, \ y \in G \rangle.$$

5. $\langle x, \ y \mid x^2y^3, \ x^3y^4 \rangle$ is a presentation for the trivial group.
6. The symmetric group $\mathcal{S}_3$ of degree 3 has presentation

$$\mathcal{S}_3 = \langle r, \ s \mid r^3, \ s^2, \ srsr \rangle.$$

7. Let $D_n, \ n > 1$ be the symmetry group of the regular n-gon P_n. This group is generated by the rotation r with angle $2\pi/n$ and a reflection s in the line through the centre and one of the vertices.

$$D_n = \langle r, \ s \mid r^n, \ s^2, \ (sr)^2 \rangle.$$

8. Let D_∞ be the infinite dihedral group consisting of the motions of $\mathbb{R}$ which map the integers to integers, i.e. the transformations $\mathbb{R} \to \mathbb{R}, \ x \mapsto \pm x + k$, with $k \in \mathbb{Z}$.

$$D_\infty = \langle s,\ t \mid s^2,\ stst\rangle.$$

9. $\mathbb{Q} = \langle \{x_n\ :\ n \geq 1\} \mid \{x_n = x_{nk}^k\ :\ n,\ k \geq 1\}\rangle$
10. Let $n \geq 2$, $X = \{x_1, \dots, x_{n-1}\}$ and $R = \{x_i^2, (x_i x_{i+1})^3 for\ 1 \leq i \leq n-2,\ [x_i,\ x_j]\ for\ |j-i| > 1\}$. Then, the symmetric group $\mathcal{S}_n = \langle X \mid R\rangle$.
Proof: Consider $\varphi : X \longrightarrow \mathcal{S}_n,\ x_i \mapsto \sigma_i = (i,\ i+1),\ 1 \leq i \leq n-1$ and observe that $\mathcal{S}_n$ is a homomorphic image of $G = \langle X \mid R\rangle$. Consider $H = \langle x_2,\ \dots\ ,\ x_n\rangle$ and its cosets

$$K_1 = H,\ K_2 = K_1x_1,\ K_3 = K_2x_2, \dots, K_n = K_{n-1}x_{n-1}$$

and verify that for each i $(1 \leq i \leq n)$, j $(1 \leq j \leq n-1)$ we have $K_ix_j = K_l$ for some l $(1 \leq l \leq n)$. Thus H has index at most n and by induction H has order at most $(n-1)!$.
11. The group

$$G = \langle x,\ y \mid x^2,\ y^3,\ (xy)^5\rangle$$

is the alternating group A_5.
Proof: The elements $\sigma = (12)(34),\ \tau = (135)$ of A_5 satisfy $\sigma^2 = 1,\ \tau^3 = 1,\ (\sigma\tau)^5 = 1$, and hence generate a subgroup of order 30 or 60; since A_5 has no subgroup of order 30 (being simple), therefore $\sigma,\ \tau$ generate A_5.
12. Consider the functions α and β on the set $\mathbb{C} \cup \{\infty\}$ defined by the rules

$$\alpha(x) = x + 2, \quad \beta(x) = \frac{x}{2x+1}.$$

here the symbol ∞ is subject to such formulae as $\frac{1}{0} = \infty$ and $\frac{\infty}{\infty} = 1$. Then α and β are bijections since they have inverses

$$\alpha^{-1}(x) = x - 2, \quad \beta^{-1}(x) = \frac{x}{1-2x}.$$

Thus α and β generate a group F of permutations of $\mathbb{C} \cup \{\infty\}$.
F is free on the set $\{\alpha,\ \beta\}$.
Proof: Note that every non-zero power of α maps the interior of the circle $|z| = 1$ to the exterior of the unit circle and a non-zero power of β maps the exterior of the unit circle to the interior with 0 removed.
13. The group with generators $a_1,\ a_2,\ a_3$ and relations

$$a_1^{-1}a_2a_1 = a_2^2;$$

$$a_2^{-1}a_3a_2 = a_3^2,$$
$$a_3^{-1}a_1a_3 = a_1^2$$

is the identity group.

14. The group G generated by $a_1,\ a_2,\ a_3,\ a_4$ subject to the defining relations

$$a_1^{-1}a_2a_1 = a_2^2;$$
$$a_2^{-1}a_3a_2 = a_3^2,$$
$$a_3^{-1}a_1a_3 = a_4^2,$$
$$a_4^{-1}a_1a_4 = a_1^2$$

has no proper normal subgroup of finite index. Since every finitely generated group has at least one maximal normal subgroup, it follows that there exists a finitely generated infinite simple group, namely the quotient group G/N where N is any maximal normal subgroup of G.

15. Let p be an odd prime. The subgroup G_p of the symmetric group on $\mathbf{R}$ generated by $x \mapsto x+1$ and $x \mapsto x^p$ is free of rank 2. This was proved by S.White in J.Algebra 118 (1988) but a more transparent proof by S.D.Cohen & A.M.W.Glass appears in Journal of the London Math. Society 55 (1997).

Exercise (Josephus permutation) 7.4

The story goes that Flavius Josephus and 39 of his comrades were surrounded when holding a revolt against the Romans during the 1st century A.D. Rather than become slaves, they decided to kill themselves. They arranged themselves along a circle. Starting somewhere, they went around the circle and the 7th person was killed. This continued with each 7th among the surviving ones being killed at each step. Apparently, Josephus was a clever mathematician and arranged himself in such a position that he would be the last survivor. The story goes that he did not kill himself but came and joined the Romans. The problem is to find out Josephus's position. Try also to solve the problem in general (i.e., with 7 and 39 replaced by other numbers). I offer a prize for the first solution.

Exercise (Sam Lloyd generalized) 7.5

The famous puzzle of Sam Lloyd, for a solution of which he offered a thousand dollars in 1879 goes as follows. Look at the picture here of a 4×4 square on which 15 coins have been placed leaving out the last square empty.

1	2	3	4
5	6	7	8
9	10	11	12
13	14	15	

The idea is to slide the coins utilising the empty square and to find out what kind of arrangements are possible. Sam Lloyd offered 1000 dollars to anyone who could get the arrangement with 14 & 15 switched above. He knew his money would be safe with him. Analyse how to decide whether a given position of coins in 15 of the 16 places is 'good' (i.e., can be brought back to the beginning position by the sliding process. Do this for a square of general size.

Exercise (Polya's enumeration of isomers) 7.6
Suppose D is a set of m objects to be coloured using a range R of k colours. Let G be the group of symmetries of D. Then, prove that the number of colour patterns $= \frac{1}{|G|} z(G; k, k, \ldots, k)$ where $z(G; s_1, s_2, \ldots, s_n)$, the 'cycle index' of G is the polynomial expression given by

$$z(G; s_1, s_2, \ldots, s_n) = \frac{1}{|G|} \sum_{g \in G} s_1^{\lambda_1(g)} s_2^{\lambda_2(g)} \ldots s_n^{\lambda_n(g)}.$$

Here $\lambda_i(g)$ denotes the number of i-cycles in G.

§8 Combinatorial group theory

Connections with topology 8.1

The notions of free groups, free products, and of free products with amalgamation come naturally from topology. For instance, the fundamental group of the union of two path-connected topological spaces joined at a single point is isomorphic to the free product of the individual fundamental groups.
The Seifert-van Kampen theorem asserts that if $X = V \cup W$ is a union of path-connected spaces with $V \cap W$ non-empty and path-connected, and if the homomorphisms $\pi_1(V \cap W) \to \pi_1(V)$ and $\pi_1(V \cap W) \to \pi_1(W)$ induced by inclusions, are injective, then $\pi_1(X)$ is isomorphic to the free product of $\pi_1(V)$ and $\pi_1(W)$ amalgamated along $\pi_1(V \cap W)$.
All these notions have found many group-theoretical applications. Recall that:
If $G_i, i \in I$ are groups, then a group G along with injective homomorphisms $\phi_i : G_i \to G$ is said to be their free product if, for every group H and homomorphisms $\theta_i : G_i \to H$, there is a unique homomorphism $\phi : G \to H$ so that $\phi \circ \phi_i = \theta_i$ for all $i \in I$.
In other words, G has the universal repelling property with respect to homomorphisms from G_i's to groups.

To construct G, one starts with a presentation $< X_i|R_i >$ of each G_i and takes $< X|R >$ as a presentation of G where X is the disjoint union of the X_i's and R is the union of the R_i's. The homomorphisms $\phi : G_i \to G$ are, therefore, simply inclusions. The uniqueness of such a free product G up to isomorphism follows from the uniqueness property of ϕ above.

One writes $G = *_{i\in I}G_i$. If I is a finite set, say, $I = \{1, 2, \cdots, n\}$, then it is customary to write $G = G_1 * G_2 * \cdots * G_n$.

For example,

The free group of rank r is the free product $\mathbb{Z} * \cdots * \mathbb{Z}$ *of* r *copies of* $\mathbb{Z}$.

More generally, a free group $F(X)$ *on a set* X *is the free product* $*_{x\in X} < x >$.

The group $PSL(2, \mathbb{Z})$ *is the free product of a cyclic group of order* 2 *and a cyclic group of order* 3.

Recall that if A is a group, $G_i, i \in I$ is a family of groups and $\alpha_i : A \to G_i$ $(i \in I)$ are injective homomorphisms, then a group G is said to be the free product of G_i's amalgamated along A, if there are homomorphisms $\phi_i : G_i \to G$ satisfying $\phi_i \circ \alpha_i = \phi_j \circ \alpha_j$ for all $i, j \in I$ such that the following universal property holds: for every group H and homomorphisms $\theta_i : G_i \to H$ with $\theta_i \circ \alpha_i = \theta_j \circ \alpha_j$ for all $i, j \in I$, there is a unique homomorphism $\theta : G \to H$ with $\phi \circ \phi_i = \theta_i$.

One denotes G by $*_AG_i$ if there is no confusion as to what the maps α_i are. Sometimes, the maps α_i are taken to be not necessarily injective and still the above definition can be carried out. Note that, if α_i are trivial, then $*_AG_i = *G_i$, the free product.

The construction is as follows. If $G_i =< X_i|R_i >$, $i \in I$, then

$$G :=< \sqcup_{i\in I}X_i| \cup R_i \cup \cup_{i,j}R_{ij} >$$

where $R_{ij} = \{\alpha_i(a)\alpha_j(a)^{-1}; a \in A\} >$.

The uniqueness of G up to isomorphism is clear once again by the uniqueness of θ.

For instance, $SL(2, \mathbb{Z}) = \mathbb{Z}/4 *_{\mathbb{Z}/2} \mathbb{Z}/6$.

The fundamental group of the Möbius strip is isomorphic to $\mathbb{Z} *_{2\mathbb{Z}} \mathbb{Z}$. A free product with amalgamation could be the trivial group even if the groups $\alpha_i(A)$ are not.

For example, let $\alpha_1 : \mathbb{Z} \to PSL(2, \mathbb{Q})$ be an injective homomorphism and let $\alpha_2 : \mathbb{Z} \to \mathbb{Z}/2$ be the natural homomorphism. Then, $G_1 *_{\mathbb{Z}} G_2 = \{1\}$.

Finally, we recall the notion of HNN extensions named after G.Higman, B.H.Neumann & H.Neumann. The construction is akin to adjoining elements to fields to get field extensions.

Let $G =< X|R >$ *be a group and let* A *be a subgroup. For an injective homomorphism* $\phi : A \to G$*, the HNN extension of* G *with respect to* ϕ *is the*

group $G^* =< X \cup \{t\} | R \cup \{tat^{-1}\phi(a)^{-1}\} >$.
It is a fact that G^* is independent of the presentation of G chosen and that G embeds naturally into G^*. It can be shown that, given two elements a, b of equal order in a group G, this construction enables one to embed the group G into a bigger group in which a, b are conjugate. The HNN construction also finds a natural topological interpretation.
For, suppose V and W are open, path-connected subspaces of a path-connected space X and suppose that there is a homeomorphism between V and W inducing isomorphic embeddings of $\pi_1(V)$ and $\pi_1(W)$ in $\pi_1(X)$. One constructs a space Y by attaching the handle $V \times [0,1]$ to X, identifying $V \times \{0\}$ with V and $V \times \{1\}$ with W. Then, the fundamental group $\pi_1(Y)$ of Y is the HNN extension of $\pi_1(V)$ relative to the isomorphism between its subgroups $\pi_1(V)$ and $\pi_1(W)$.

Nielsen-Schreier Theorem 8.2
If W is a subgroup of a free group F, then W is a free group. Moreover, if W has finite index m in F, the rank of W is precisely $nm + 1 - m$, where n is the rank of F (which may be infinite).
Proof. Let F be a free group on a set X; let W be an arbitrary subgroup of F. Label the right cosets of W in F by means of an index set I containing the symbol 1, with the convention that $W_1 = W$. Choose a right transversal to W in F , the representative of the coset W_i being written as $\overline{W}_i$, with the stipulation that $\overline{W} = 1$.

If $u \in F$, the elements $\overline{W}_i u$ and $\overline{W_i u}$ belong to the same right coset of W, so that
$$\overline{W}_i u (\overline{W_i u})^{-1} \in W.$$

For each $i \in I$ and $x \in X$, choose a symbol y_{ix}, and let $\hat{F}$ be the free group on the set of all y_{ix}. Consider the homomorphism
$$\tau : \hat{F} \longrightarrow W, \quad y_{ix} \longmapsto \overline{W}_i x (\overline{W_i x})^{-1}.$$

Claim: τ is surjective.
Coset maps: For $i \in I$ and $u \in F$ associate an element $u^{W_i} \in F^{-1}$ as follows:
$$F \longrightarrow \hat{F} \quad u \longmapsto u^{W_i}$$
as follows;
$$1^{W_i} = 1, \quad x^{W_i} = y_{ix}, \quad \textit{and } (x^{-1^{W_i}}) = (x^{W_i x^{-1}})^{-1} \ (x \in X).$$

Generally, if $u = vy$ in reduced form with $y \in X \cup X^{-1}$, define u^{W_i} by induction on the length of u by means of the equation

$$u^{W_i} = v^{W_i} y^{W_i v}.$$

If u and v belong to F, then

$$(uv)^{W_i} = u^{W_i} v^{W_i u} \textit{and } (u^{-1})^{W_i} = (u^{W_i u^{-1}})^{-1}.$$

If $u \in F$ and $i \in I$, then $\tau(u^{W_i}) = \overline{W}_i u(\overline{W_i u})^{-1}$.
Let

$$\psi : W \longrightarrow \hat{F}$$

be the restriction of the coset map $u \longmapsto u^W$ to W. Observe that ψ is a homomorphism and

$$\psi\tau = 1.$$

Hence ψ is injective and τ is surjective. Write

$$\chi = \tau\psi,$$

an endomorphism of $\hat{F}$.
The group W has a presentation $\tau : \hat{F} \longrightarrow W$ with generators y_{ix} and defining relators $y_{ix}^{-1} y_{ix}^{\chi}$ $(i \in I,\ x \in X)$.
The elements u^W, where u runs over the set of nontrivial elements in the transversal, form a set of defining relators for the presentation $\tau : \hat{F} \longrightarrow W$.
A transversal for W is called a *Schreier transversal* if it has the following property:

Schreier property: *Every initial subword of a representative is a representative.*

There exists a right Schreier transversal to W in F.
The Nielsen-Schreier theorem can now be proved.
Reidemeister-Schreier Theorem 8.3
Let G be a group. Suppose that $\varphi : F \longrightarrow G$ is a presentation of G with generators X and relators R. Let W be the preimage of H under φ. Then with the above notation:

(i) *$\tau\varphi : \hat{F} \longrightarrow H$ is a presentation of H with generators y_{ix} and defining relators r^{W_i}, u^W, $i \in I$, $r \in R$, and u is a non-trivial element of a* *transversal to W in F.*

(ii) *If $[G : H]$ is finite and G is n-generator group, then H can be generated by $mn + 1 - m$ elements.*

Corollary 8.4

If G is finitely presented and H is a subgroup of finite index in G, then H is finitely presented.

Example 8.5

Consider the free group $F = \langle x,\ y \mid \phi\rangle$. The homomorphism $\varphi : F \longrightarrow \mathbb{Z}_n = \{0, 1, 2, \ldots, n-1\},\ n \geq 2$ defined by $x,\ y \mapsto 1$ and let U be the kernel of φ. As coset representatives we take $1, y, y^2, \ldots, y^{n-1}$; clearly this system satisfies the Schreier property. Then the Reidemeister-Schreier generators for U are the non-trivial elements of the following set:

$$\{y^i.x.(\overline{y^i x})^{-1},\ y^i.y(\overline{y^{i+1}})^{-1} \mid i = 0, 1, 2, \ldots, n-1\}$$

We obtain the non-trivial elements

$$y^n,\ x_i = y^i.x.y^{-(i+1)}$$

for $i = 0, \ldots, n-2$ and $x_{n-1} = y^{n-1}.x$. The rank of this group is $n + 1$. Note that this shows that: *The free group of rank 2 contains a free group of rank n as a subgroup of index $n - 1$.*

Exercise 8.6

1. Show that the modular group $PSL(2,\ \mathbb{Z})$ contains a free subgroup of rank 2 and index 6.

2. If F is a free group of rank two, then prove that the commutator subgroup of F is an infinitely generated free group.

3. Prove that the group

   $$\Gamma = \{\begin{pmatrix} a & b \\ 0 & 1 \end{pmatrix} \in GL(2,\ \mathbb{R}) \mid a = 2^n \ and\ b = \frac{p}{2^q}\ for\ some\ n,\ p,\ q \in \mathbb{Z}\}$$

   is generated by the two matrices $\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}$ and $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, while

   $$\Gamma_0 = \{\begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} \in GL(2,\ \mathbb{R}) \mid b = \frac{p}{2^q}\ for\ some\ p, q \in \mathbb{Z}\}$$

   is not finitely generated.