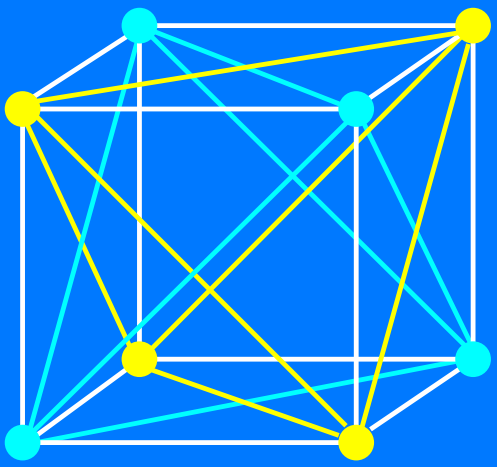
Lectures on Groups, Rings, and Fields



Alexander Givental

Lectures on
Groups, Rings, and Fields

by

Alexander Givental

Sumizdat

Published by Sumizdat
5426 Hillside Avenue, El Cerrito, California 94530, USA
http://www.sumizdat.org

ISBN 978-0-9779852-3-4

In honor of Emil and Emmy

Preface

This book is intended as a primer in abstract algebra.

My own first encounter with the subject happened in the 8th grade, at the Moscow school No. 2, one of several specialized schools in the Soviet Union with advanced math and physics curriculum. It was a topics course based on opening chapters of van der Waerden's "Modern Algebra". After two years of Euclidean geometry, where problems were charming but foundations shaky, the prospect of deriving everything from a few axioms looked very appealing to a juvenile aspiring mathematician. There was one problem though: in the maze of possible implications from the few axioms, I could not find my way to solving a single homework exercise. One day I complained to our teacher, Valery Senderov, and he gave me a guiding thread. It turned out that the gist of the theory was not in forming neat chains of implications from the axioms, but rather in answering meaningful questions about numerous interesting objects populating the realm of mathematics.

The next year, in an effort to solidify my first (which also turned out to be my last) formal exposure to the subject, I served some library time reading thoroughly several chapters from van der Waerden's 1930 classics. Only much later did I realize that his book was based on lecture courses taught by Emil Artin and Emmy Noether in 1924–28. Namely, when my son was about to take his abstract algebra course at Haverford College, I copied for him some chapters from van der Waerden's text — and finally read the Introduction.

Some connections looked somewhat symbolic. Emmy Noether's ashes rest at the grounds of Haverford's sister college Bryn Mawr, where she held her last professorship — after the purge of Jewish faculty from the University of Göttingen. The textbook selected by my son's professor was "Algebra" by Michael Artin; as a child he was brought by his parents from Germany to the US — for much the same reason: Emil Artin's wife had Jewish roots.

M. Artin's "Algebra" intersperses chapters on groups, rings, and fields with others on linear algebra, which in my view makes it inconvenient for a one-semester course (in either subject). However, the professor at Haverford, Elizabeth Milićević, managed to assemble the material around a theme usually left out of the first course — Galois theory. The approach actually made sense: the abstract idea of symmetry (groups) and the theory of factorization (rings) ultimately reconcile in a study of the symmetries of fields.

Next semester I happened to teach an honors course in abstract algebra to a quite capable group of UC Berkeley students. By the spring break we had completed successfully the most of the required material on groups and rings, and so in the remaining part of the semester I decided to realize the same approach. Over the break, I wrote self-contained notes on Galois theory of finite field extensions up to Gauss' theorem on straightedge-and-compass constructions of regular polygons and Abel's theorem on unsolvability of polynomial equations in radicals.

The present book is a reconstruction of that course from the notes.

Alexander Givental
Department of Mathematics
University of California Berkeley
January 2022

Contents

Prologue **1**
Lecture 1. Sets and functions 1
Lecture 2. Integers 7

Groups **17**
Lecture 3. Groups and isomorphisms 17
Lecture 4. Homomorphisms 25
Lecture 5. Cosets 31
Lecture 6. Rotations of the cube 37
Lecture 7. Symmetric and alternating groups 45
Lecture 8. Group actions 51
Lecture 9. Sylow's theorems 57
Lecture 10. Finitely generated abelian groups 61
Lecture 11. Finite abelian groups 67

Rings **73**
Lecture 12. Rings and modules 73
Lecture 13. Homomorphisms and ideals . . 81
Lecture 14. Principal Ideal Domains 87
Lecture 15. Unique Factorization Domains . 93
Lecture 16. Primes in some Euclidean rings 99
Lecture 17. Symmetric functions 105

Fields **111**
Lecture 18. Field extensions 111
Lecture 19. Geometric constructions 115
Lecture 20. Algebraic closures 119
Lecture 21. Separable extensions 123
Lecture 22. The Galois theory 127
Lecture 23. Regular polygons 131
Lecture 24. Cyclic extensions 137
Lecture 25. Unsolvability by radicals 141
Lecture 26. Cubics and quartics 147
Appendix I: The AC, ZT, and ZL 151
Appendix II: The FTA 153

Index **155**

Lectures on

Groups, Rings, and Fields

Prologue

Lecture 1. Sets and functions

The Queen of Sciences has one notable distinction from its loyal subjects: it is self-contained. To validate a mathematical result, there is no need to rely on somebody else's conclusions or costly experiments. Instead, mathematicians are supposed to provide a complete exposition of their theory starting from precise definitions of very basic terms, and the students are expected to verify every detail of the theory at the tip of their own pens. The first difficulty one encounters in following this agenda comes when, attempting to rigorously define basic notions in terms of those previously defined, one runs out of what was "previously defined". This is exactly our present position, for we are about to build a mathematical theory which serves as a common ground for many further developments. Thus, we have to start from "ground zero", which includes introducing some notions so basic that their meaning can only be explained informally, and inferred from our common sense and past experience. In fact all we need in the role of such *undefinable notions* are sets and functions; everything else can be then explained and defined quite accurately and formally.

A. Sets. By a *set* one means any collection of any objects. That is, a set X is considered given when for any object x in the Universe it is known whether this object is an *element* of the set (written $x \in X$) or not ($x \notin X$). One cannot say more formally what sets are, nor what elements are, but assumes it is always well-defined whether a given element and a given set are in this Hamletian relationship: of the former to be or not to be an element of the latter.

By starting from "ground zero" we do not mean to suggest that the reader's life to this point was a total waste. In the contrary, we assume that everyone is well-familiar with *natural, integer, rational, real* and

even *complex numbers*, and use the following standard notations for the sets they form (note the declaration formats):

$$\mathbb{N} = \{1, 2, 3, \dots\},\ \mathbb{Z} = \{0, \pm 1, \pm 2, \dots\},\ \mathbb{Q} = \{\frac{m}{n} \mid m \in \mathbb{Z},\ n \in \mathbb{N}\},$$
$$\mathbb{R} = \{\text{all real numbers}\}, \quad \mathbb{C} = \{a + bi \mid a, b \in \mathbb{R}\},$$
$$[\text{the set's name}] = \{\ [\text{objects}]\ [\ |\ =\ \text{“such that”}]\ [\text{conditions}]\ \}.$$

Of course, $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$, where $A \subset B$ (or $B \supset A$) means that set A is a *subset* of set B, i.e. that all elements of A are elements of B as well.

We also assume that the reader is used to such notations as the *union* of sets $A \cup B$, defined as the set of all elements contained in A *or* B (or both), *intersection* $A \cap B$, consisting of all elements common for A and B, the *complement* A^c, consisting of all elements *not* in A, and is familiar with the basic properties of these operations. For further details, together with some elements of formal logic, we can simply refer the reader to courses in Discrete Mathematics where this material has become standard.

B. Functions. By a *function* (or *mapping*, or *map*) from one set to another one means a rule that to each element of the former set associates exactly one element of the latter. This is written as

$$f : X \to Y,\ X \ni x \mapsto f(x) \in Y,$$

where f is the name of the rule, X is the source set called *domain*, Y the target set called *codomain*, and $f(x)$ denotes the *value* of the function at x. A purist would say that functions can be defined in terms of sets by not distinguishing a function from its *graph*:

$$\text{graph}(f) := \{(x, y) \in X \times Y \mid y = f(x)\}.$$

Here $X \times Y$ is the *Cartesian product* of set X and set Y, which by definition consists of all ordered pairs (x, y) where $x \in X$ and $y \in Y$. To be the graph of a function from X to Y, a subset $\Gamma \subset X \times Y$ must for every $x \in X$ contain exactly one pair (x, y) (in which case that unique y is declared to be the value of the function at x). This formally accurate definition of functions in terms of sets has one defect: it makes sense only to those who already know informally what functions are. So, we will distinguish functions, which are rules, from their graphs, which are certain subsets in the Cartesian product of the domain and codomain.

Given two functions $g : X \to Y$ and $f : Y \to Z$, their *composition* $h = f \circ g : X \to Z$ is defined by consecutively applying the two rules: $h(x) := f(g(x))$. Note that this operation is defined only when the domain of f coincides with the codomain of g. In particular, the composition in the opposite order may be ill-defined. Even when it is defined, there is no reason to expect that the result is the same. For instance, $\sin 2x = 2 \sin x \cos x \neq 2 \sin x$. Thus, generally speaking, composition of functions is not commutative. However, composition of functions is always associative. Namely, given three functions $h : W \to X$, $g : X \to Y$, and $f : Y \to Z$, either composition $(f \circ g) \circ h$ and $f \circ (g \circ h)$ is defined the same way: by consecutively applying the three rules: $w \mapsto h(w) \mapsto g(h(w)) \mapsto f(g(h(w)))$. As we will see, the associativity of algebraic operations is ultimately due to the associative property of the composition of mappings. For example, the associativity of addition $a+b$ of real numbers can be explained this way by interpreting the operation as the composition of translations $x \mapsto y = x + a$ and $y \mapsto y + b$ on the number line.

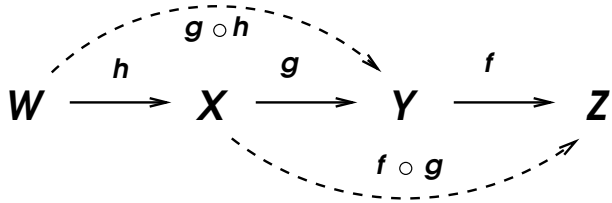


Figure 1: $f \circ (g \circ h) = (f \circ g) \circ h$.

C. Inverses. Given a function $f : X \to Y$, the subset $f(X)$ in Y consisting of all values of the function is called its *range*. When the range is the whole of Y, the function is said to be *onto*, or *surjective*. A function which maps different elements of X to different elements of Y is called *one-to-one*, or *injective*. When both properties hold, the function is called *bijective* or a *one-to-one correspondence* between X and Y. When this happens, the rule $g : Y \to X$ which to each $y \in Y$ assigns that unique $x \in X$ such that $f(x) = y$ is well-defined. The mapping g is called the *inverse* to f (as it undoes whatever f does) and is sometimes denoted by f^{-1}. We have:

$$g \circ f = \mathrm{id}_X, \quad f \circ g = \mathrm{id}_Y,$$

where $\mathrm{id}_A : A \to A : a \mapsto a$ denotes the *identity map*. Conversely, when these compositions of g and f coincide with the respective identity maps, the functions are inverse to each other. Indeed, f is onto,

because every $y \in Y$ is the value of f at $x = g(y)$ (due to $f \circ g = \mathrm{id}_Y$), and whenever $f(x_1) = y = f(x_2)$, we have $x_1 = g(y) = x_2$ (due to $g \circ f = \mathrm{id}_X$), i.e. f is one-to-one, and $f^{-1}(y) = g(y)$. Moreover, these arguments prove the "if" parts of the following

`Proposition 1.` *(i) A map $f : X \to Y$ between non-empty sets is injective if and only if it has a* left inverse*, i.e. there exists $g : Y \to X$ such that $g \circ f = \mathrm{id}_X$.*

(ii) A map $f : X \to Y$ between non-empty sets is surjective if and only if it has a right inverse*, i.e. there exists $g : Y \to X$ such that $f \circ g = \mathrm{id}_Y$.*

To establish the "only if" part of (i), note that an injective map $f : X \to Y$ is a bijection onto its range $f(X) \subset Y$. To define the left inverse $g : Y \to X$, take g to be f^{-1} on the range, and extend g to $y \in Y - f(X)$ (this is the notation for the *complement* of $f(X)$ in Y) by $g(y) := x_0$. Here $x_0 \in X$ is any *pre*-selected element (and that's why we need to assume that X is non-empty). In particular we see that the left inverse g of an injective but not bijective map can be non-unique.

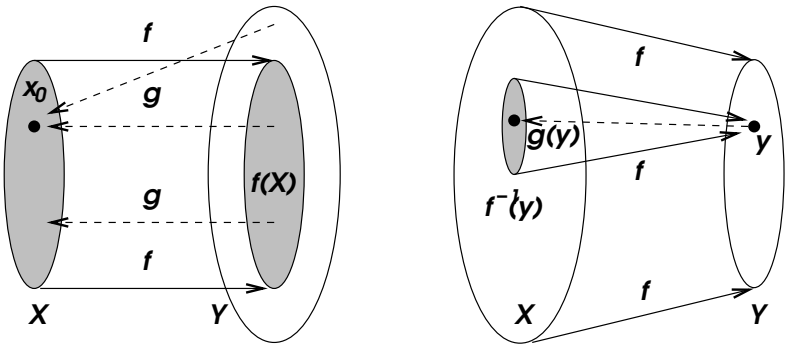


Figure 2: One-sided inverses.

In order to establish the "only if" part of (ii), i.e. to define a right inverse $g : Y \to X$ to an onto map $f : X \to Y$, we pick for each $y \in Y$ one element in the non-empty set $f^{-1}(y) := \{x \in X \mid f(x) = y\}$ (called the *inverse image* of y), and declare it the value of g at y.

D. The Axiom of Choice. In fact the last argument hides a subtle flaw. It seems obvious that one can select an element in each set $f^{-1}(y)$ as long as it is non-empty. However, if you are pressed to point out a *rule*, g, that makes such a selection for every $y \in Y$, you might find yourself at a loss. When the set X is *countable*, that is, can be put into one-to-one correspondence with the set $\mathbb{N}$ of natural numbers, the issue can be resolved as follows. Every non-empty subset in $\mathbb{N}$ contains its smallest element. So, selecting in a non-empty subset of X that element which under the identification of X with $\mathbb{N}$ corresponds to the smallest element in the subset defines the required rule. However, the set of real numbers is known to be *uncountable*, i.e. it cannot be put in one-to-one correspondence with $\mathbb{N}$. So, this method wouldn't work for $X = \mathbb{R}$. Actually in the course of development of mathematical logic in the 20th century it was established that it is impossible to describe for every set a definite selection rule (Kurt Gödel, 1938), and yet assuming the existence of such a rule would not lead to any contradiction (Paul Cohen, 1963). Thus it is safe to accept the existence of the rule as an axiom:

`Axiom of Choice.` *For every set, there exists a function that to every non-empty subset of it associates an element from that subset.*

Starting from this point, mathematical results can be divided into two types: those which do not rely on the Axiom of Choice, and those which do. In algebra, it is usual to accept this axiom, because several important general constructions cannot be executed without it. We will encounter some of such constructions later in this book, and will discuss some other, more convenient (and seemingly less plausible, but actually equivalent) reincarnations of this axiom when the need becomes pressing.

E. Partitions. Let $f : X \to Y$ be an onto function. Then X can be partitioned into non-overlapping subsets $f^{-1}(y)$ according to the values $y \in Y$ of f. More formally, call two elements $x_1, x_2 \in X$ f-equivalent (and write $x_1 \sim_f x_2$) whenever $f(x_1) = f(x_2)$. The notion of f-equivalence has the following three obvious properties:

(i) It is *reflexive*: $x \sim_f x$ for every $x \in X$.

(ii) It is *symmetric*: If $x_1 \sim_f x_2$, then $x_2 \sim_f x_1$.

(iii) It is *transitive*: If $x_1 \sim_f x_2$ and $x_2 \sim_f x_3$, then $x_1 \sim_f x_3$.

In fact $\sim_f$ is an example of a *binary relation*, a notion popular in computer science. By a binary relation on a set X one means any subset $R \subset X \times X$ in the Cartesian square of X. When an ordered pair $(x_1, x_2) \in X \times X$ lies in R, one says that "x_1 and x_2 are in the

relation R", and writes $x_1 R x_2$. As examples, consider the subsets on the plane $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R} = \{(x_1, x_2) \mid x_1, x_2 \in \mathbb{R}\}$ given by: (a) $x_1 = x_2$, (b) $x_1 \neq x_2$, (c) $x_1 \leq x_2$, (d) $x_1 > x_2$, defining on $\mathbb{R}$ the relations of (a) "equal to", (b) "unequal to", (c) "less than or equal to", and (d) "strictly greater than" respectively.

A binary relation on a set X satisfying the three properties (i), (ii), (iii) is called an *equivalence relation*. The following proposition shows that (conversely to the above observation) any equivalence relation defines a partition of X.

`Proposition 2.` *Let $\sim$ be a reflexive, symmetric, and transitive binary relation on a set X. For every $x \in X$, define the* equivalence class *$\bar{x}$ as the subset in X of all elements $x' \in X$ which are in the relation $\sim$ with x: $\bar{x} := \{x' \in X \mid x' \sim x\}$. Then the equivalence classes form a* partition *of X.*

`Proof.` By a partition of X we mean of course a representation of X as the union $X = \cup_\alpha X_\alpha$ of its subsets X_α which are pairwise disjoint, i.e. if $X_\alpha \cap X_\beta \neq \emptyset$ (this is the notation for the *empty set*), then $X_\alpha = X_\beta$.

The reflexive property guarantees that $x \in \bar{x}$ and implies that the union of the equivalence classes equals X. Suppose now that $a \in \bar{x} \cap \bar{y}$, i.e. $a \sim x$ and $a \sim y$. Then $x \sim a$ (symmetricity), and hence any $z \in \bar{x}$ satisfies $z \sim x \sim a \sim y$, i.e. (using transitivity twice) $z \in \bar{y}$. This shows that $\bar{x} \subset \bar{y}$. Since likewise $\bar{y} \subset \bar{x}$, we conclude that $\bar{x} = \bar{y}$ whenever the intersection $\bar{x} \cap \bar{y}$ is non-empty.

As the reader can see, this proposition is *tautological*, i.e. follows directly from the definitions[1]. Nevertheless it is useful as the basis for the following construction.

Given an equivalence relation $\sim$ on X, introduce a new set, sometimes denoted by $X/\sim$, and called the *quotient set*, whose elements are the equivalence classes. Define the *projection map* $\pi : X \to X/\sim$ by associating to an element $x \in X$ its equivalence class: $\pi(x) := \bar{x}$. Then the equivalence relation $\sim$ coincides with $\sim_\pi$ introduced at the start of this section.

This construction of the projection of a set to the quotient set defined by an equivalence relation will serve us as a standard tool beginning with the next lecture.

[1] Dictionaries define "tautology" as something identical to itself.

Lecture 2. Integers

A. Modular arithmetic. Here is an important illustration to the construction of the quotient set.

Let n be a positive integer. Two integers $a, b \in \mathbb{Z}$ are called *congruent modulo* n (notation: $a \equiv b \mod n$) if n divides $a - b$ (notation: $n \mid a - b$), i.e. if there exists $k \in \mathbb{Z}$ such that $a - b = kn$.

This is an equivalence relation: (i) $n \mid (a - a)$; (ii) if $n|(a - b)$, then $n|(b - a)$; (iii) if $n|(a - b)$ and $n|(b - c)$, i.e. $a - b = kn$ and $b - c = ln$ for some k, l, then $a - c = (a - b) + (b - c) = (k + l)n$, and hence $n|(a - c)$.

Therefore the set $\mathbb{Z}$ of all integers is partitioned into the *congruence classes* modulo n. Of course, each congruence class consists of those integers which have the same remainder upon division by n. The "skip-counting" exercises in Kindergarten consist in listing the elements of such congruence classes. Modulo 2, there are two classes: even and odd integers. Modulo 3, there are three classes: $\ldots, -3, 0, 3, 6, \ldots$, $\ldots, -2, 1, 4, 7, \ldots$, $\ldots, -1, 2, 5, 8, \ldots$. The set of congruence classes modulo n is usually denoted by $\mathbb{Z}/n\mathbb{Z}$, but we will often use the shorter notation: $\mathbb{Z}_n$.

`Proposition.` *There is a unique way to equip the set $\mathbb{Z}_n$ of congruence classes modulo n with the operations of addition and multiplication so that the natural projection $\pi : \mathbb{Z} \to \mathbb{Z}_n$ maps the sums and products of integers into respectively the sums and products of their congruence classes: $\pi(a + b) = \pi(a) + \pi(b)$, $\pi(a \cdot b) = \pi(a) \cdot \pi(b)$.*

`Proof.` Note that the operations "+" and "·" on the left of the equality sign are those with integers, and on the right of it are the operations with the congruence classes in $\mathbb{Z}_n$. In fact the *uniqueness* of the latter operations becomes obvious once we read the equalities from right to left. What they tell us is that the sum (product) of two equivalence classes must be defined by choosing *representatives* a and b of these classes in $\mathbb{Z}$, computing the sum $a + b$ (product $a \cdot b$) of the integers, and declaring the congruence class $\pi(a + b)$ (resp. $\pi(a \cdot b)$) of the result to be the sum (product) of the two congruence classes.

What still remains non-obvious is the *existence* of such operations. Namely, the construction of the sum and product of two congruence classes involved some choices: the choices of the representatives of the classes. Therefore we need to show that the results of the operations do not depend on the choices.

Indeed, replacing a and b by integers $a + kn$ and $b + ln$, which differ from them by multiples of n results in the sum and the product

which differ from $a+b$ and ab by some multiples of n:

$$(a+kn)+(b+ln)=a+b+(k+l)n,$$
$$(a+kn)(b+ln)=ab+(al+bk+kln)n.$$

Therefore the congruence classes of the results don't change.

One may wonder what was the point of concealing under an abstract construction of the quotient set the simple and familiar fact that one can add and multiply remainders modulo n, as in $2+2=4\equiv 1 \mod 3$ and $2\cdot 2=4\equiv 1 \mod 3$? Later we will not only see that the abstract point of view has many advantages, but moreover, that the ability to think in terms of equivalence classes (rather than in terms of remainders) is an important prerequisite to understanding algebra. The remainders $0,1,\dots,n-1$ modulo n form a *subset* in $\mathbb{Z}$ (and indeed can be operated upon under the rule of "overflow": whenever the sum or the product exceeds n, it is to be replaced by the remainder. However, in the modular arithmetic, $\mathbb{Z}_n$ is not a subset of $\mathbb{Z}$, but the quotient set. In particular, in $\mathbb{Z}_3$, we have: $\bar{2}+\bar{2}=\bar{4}=\bar{2}\cdot\bar{2}$, i.e. 4 is just as good a representative of its congruence class as its remainder 1. (Note that $\bar{a}$ denotes the congruence class of a, and we use the equality $\bar{a}=\bar{b}$ in $\mathbb{Z}_n$ instead of $a\equiv b \mod n$ in $\mathbb{Z}$.) To illustrate an advantage of the abstract point of view, consider the following example.

`Example.` A standard topic in introductory CS texts is the binary representation of negative numbers. In order to operate with integers ranging between -2^n and $+2^n$, e.g. (choosing for simplicity $n=4$) from -1111_2 to $+1111_2$, one is taught to use $n+1$-digit positive binary encoding. Namely, a positive n-digit binary is augmented by the leftmost 0 (e.g. 1010_2 becomes 01010_2) while a negative $n+1$-digit binary (with 0 in the leftmost place) is replaced with the *complementary* positive binary code plus 1 (e.g. -01010_2 becomes $10101_2+1_2=10110_2$). Thus, the leftmost digit indicates the sign: 0 means positive, and 1 negative. Then, the theory explains, to operate with (say, add) the binaries, one performs the operations with the positive $n+1$-digit binary codes (when of course in the case of overflow, the $n+2$-nd digit is lost). Finally, one demonstrates that the result of the addition represents correctly, through the reverse encoding, the result of the operation with signed n-digit binaries (at least when the normal operations with them would not have resulted in overflow). The demonstration is usually done case-by-case assuming the 1st operand positive or negative, and the 2nd operand positive or negative (so totally 4 cases, of which one is obvious though).

The above (unfortunately typical) exposition demonstrates in fact a lack in understanding of the modular arithmetic. Operating with $n+1$-digit binaries, so that the "overflow" $n+2$-nd digit is lost, is nothing but the arithmetic in $\mathbb{Z}_N$ with $N = 2^{n+1}$. The encoding rule replaces the representatives of congruence classes taken from the range $-2^n \leq a < 2^n$ with those in the range $0 \leq b < 2^{n+1}$. For negative a, this results in $b = a + 2^{n+1} = (2^{n+1} - 1) - |a| + 1$ (which indeed is obtained from the code for $|a|$ by adding 1 to the complementary code). There is no need, however, to verify the correctness of operations by a case-by-case analysis, for our Proposition shows that the operations in $\mathbb{Z}_N$ are well-defined, i.e. do not depend on the choice of representatives of the congruence classes.

B. The Euclidean algorithm. Given two integers m and n, one denotes by $G.C.D.(m,n)$ their *greatest common divisor*. One way to define it (so that it would clearly exist) is to take the greatest of all positive common divisors of m and n. However, the term has another, stronger meaning, and refers to a positive common divisor d of m and n which is divisible by every other their common divisor. Why does such a divisor exist?

The answer was given 23 centuries ago by Euclid in the fifth of his thirteen books of *The Elements*. The key observation is that, given a pair of integers m and $n > 0$, the *set* of their common divisors doesn't change, when this pair is replaced with n and r, where r is the *remainder* upon division of m by n:

$$m = qn + r, \quad 0 \leq r < n.$$

Therefore, if one continues this way, and replaces n and r by r and r_1 where $0 \leq r_1 < r$ is the remainder upon division of n by r, and then by r_1 and r_2, where r_2 is the remainder upon division of r by r_1, and so on, then the set of common divisors of each next pair remains the same as that of m and n. The steps of such *Euclidean algorithm* are possible as long as both integers remain non-zero. However, since $n > r > r_1 > r_2 > \cdots \geq 0$, there can only be finitely many non-zero remainders, and after at most n steps the pair will become $(d, 0)$. Here $d > 0$ is the last non-zero remainder, and the claim is that $d = G.C.D.(m,n)$. Indeed, the common divisors of m and n are the same as the common divisors of d and 0, which are just the divisors of d. Thus, d is a common divisor of m and n, and their every other common divisor is a divisor of d.

`Example:` $m = 2737$, $n = 1001$. We have: $2737 = 2 \cdot 1001 + 735$, $1001 = 1 \cdot 735 + 266$, $735 = 2 \cdot 266 + 203$, $266 = 1 \cdot 203 + 63$, $203 = 3 \cdot 63 + 14$, $63 = 4 \cdot 14 + 7$, $14 = 2 \cdot 7 + 0$. Thus, $G.C.D.(2737, 1001) = 7$.

The Euclidean algorithm has many remarkable features, and will reoccur many times throughout the book. One consequence of it we need to point out right now is that it allows us to represent the greatest common divisor d as a linear combination of m and n.

`Proposition.` *The greatest common divisor $G.C.D.(m, n)$ is the smallest positive integer representable as the linear combination $km + ln$ of m and n with some integer coefficients, k and l.*

Indeed, $d = km + ln$ is divisible by every common factor of m and n, and so, if positive, cannot be smaller than $G.C.D.(m, n)$. On the other hand, the remainder $r = m - qn$ is a linear combination of m and n, as well as every subsequent remainder: $r_1 = n - q_1 r = n - q_1(m - qn)$, and so on, up to the last non-zero one.

E.g., working backward through the above example, we find:

$$\begin{aligned} 7 &= 63 - 4 \cdot 14 = 63 - 4 \cdot (203 - 3 \cdot 63) \\ &= 13 \cdot 63 - 4 \cdot 203 = 13 \cdot (266 - 1 \cdot 203) - 4 \cdot 203 \\ &= 13 \cdot 266 - 17 \cdot 203 = 13 \cdot 266 - 17 \cdot (735 - 2 \cdot 266) \\ &= 47 \cdot 266 - 17 \cdot 735 = 47 \cdot (1001 - 1 \cdot 735) - 17 \cdot 735 \\ &= 47 \cdot 1001 - 64 \cdot 735 = 47 \cdot 1001 - 64 \cdot (2737 - 2 \cdot 1001) \\ &= 175 \cdot 1001 - 64 \cdot 2737. \end{aligned}$$

`Corollary 1.` *When integers m and n are coprime, there exist integers k and l such that $km + ln = 1$.*

`Proof:` Indeed, by definition, *coprime integers* have $G.C.D. = 1$.

`Corollary 2.` *A congruence class $\bar{m}$ in $\mathbb{Z}_n$ is multiplicatively invertible if and only if m is coprime to n.*

`Proof.` If $km \equiv 1 \mod n$, then $km - 1 = ln$ for some l, and hence $G.C.D.(m, n) = 1$. Conversely, if $km + ln = 1$, then $km \equiv 1 \mod n$, and hence k is inverse to $\bar{m}$ in $\mathbb{Z}_n$.

C. Prime factorization. An integer $p > 1$ which is divisible only by 1 and by itself is called ... No, it is not called "prime", it is called *irreducible*, meaning that it cannot be factored in any non-trivial way, or equivalently cannot be represented as the product

$p = ab$ of smaller integers: $|a|, |b| < p$. By definition, a number p is called *prime* if $p|ab$ implies $p|a$ or $p|b$.

Being prime is potentially a stronger property than being irreducible. Indeed, if p is prime, and $p = ab$, then one of a and b must be divisible by p (say, $b = pc$), and hence $ac = 1$, making $a = \pm 1$, and the factorization of p trivial.

Note that this argument is essentially tautological: we use only the definitions of "prime" and "irreducible". In fact the converse statement: "an irreducible integer is prime" is also true, but it cannot be obtained by merely manipulating with definitions. Indeed, such a manipulation would lead to a contradiction when applied to the following example of an irreducible number which is not prime.

`Example.` Instead of integers, consider the system $R \subset \mathbb{C}$ of complex numbers of the form $z = a + b\sqrt{-3}$, where $a, b \in \mathbb{Z}$. In R, we have

$$2 \cdot 2 = (1 + \sqrt{-3})(1 - \sqrt{-3}).$$

This shows that 2 is not prime in R, since it divides the product (equal to 4), but clearly does not divide any of $1 \pm \sqrt{-3}$. We claim that nonetheless 2 is irreducible in R. To show this, introduce the *norm*

$$N(a + b\sqrt{-3}) := (a + b\sqrt{-3})(a - b\sqrt{-3}) = a^2 + 3b^2.$$

It a non-negative integer taking values $0, 1, 3, 4, \dots$, but not 2. Indeed, $N = 0$ means $a = b = 0$; $N = 1$ is true only for $a + b\sqrt{-3} = \pm 1$; when $|b| \geq 1$ we have $N \geq 3$, and if $|a| \geq 2$, then $N \geq 4$. If 2 could be factored in R in any nontrivial way: $2 = zw$, i.e. with none of z, w equal ± 1, then $4 = N(2) = (z\bar{z})(w\bar{w}) = N(z)N(w)$ would be a non-trivial factorization of 4 into the values of the norm, which is impossible.

In fact in this example, the number 4 is factored in two different ways into a product of irreducibles. One could argue that in $\mathbb{Z}$ irreducible numbers are prime because the factorization into irreducibles is unique. Indeed, if an irreducible p divides ab, then it must appear in the factorization of at least one of a, b. However, this argument takes the Fundamental Theorem of Arithmetic for granted. Since we are currently at the "ground zero", we should prove it first.

`The Fundamental Theorem of Arithmetic.` *Every integer $n > 1$ can be factored into the product of irreducible positive integers, and such a factorization is unique up to reordering of the factors.*

`Proof.` The existence part is straightforward. We take an integer $n > 1$; if it is irreducible, we are done. Otherwise we factor it as $n = ab$, where $1 < a, b < n$, and proceed the same way with each a and b. Since there are only finitely many integers between 1 and n, the process would stop after finitely many steps, and yield a factorization $n = p_1 \cdots p_N$, where each p_i is irreducible.

Now let us try to establish uniqueness. For this, we should start with another factorization

$$p_1 \cdots p_N = q_1 \cdots q_M,$$

where all q_j are irreducible too, and show that $M = N$, and that the list of p_i differs from the list of q_j only by ordering. Thus, we start with p_1 and try to find it among q_j. If $M = 1$, then $p_1 | q_1$, i.e. $p_1 = q_1$ (since both are irreducible) and we are done. However, when $M > 1$, we get stuck, exactly because we don't know yet whether $p_1 | q_1 q_2$ implies $p_1 | q_1$ or $p_1 | q_2$, i.e. whether an irreducible p_1 is necessarily prime. So, we need the following lemma, which can be derived from the Euclidean algorithm.

`Lemma.` *In $\mathbb{Z}$, every irreducible is prime.*

`Proof.` Suppose $p | ab$ and $p \nmid a$. If p is irreducible, we have $G.C.D.(p, a) = 1$, and hence $1 = kp + la$ for some k and l. Then $b = kpb + lab$, where ab is divisible by p, and hence $p | b$.

Now we can complete the proof of uniqueness of prime factorization. Indeed, if $p_1 | (q_1 \cdots q_{M-1}) q_M$, then either $p_1 | q_M$ and hence $p_1 = q_M$, or $p_1 | q_1 \cdots q_{M-1}$. In the second case we similarly conclude that either $p_1 = q_{M-1}$, or $p_1 | q_1 \cdots q_{M-2}$, and so on. Thus, one way or another, we find that p_1 must coincide with one of q_j. Renaming the qs so that p_1 comes first, and canceling p_1 on both sides, we obtain equal factorizations: $p_2 \ldots p_N = q_2 \cdots q_M$. Proceeding the same way with p_2, then p_3, and so on, we find after N steps, that $1 = q_{M-N+1} \cdots q_M$, i.e. $M = N$ and the list of ps coincides with that of qs after renumbering.

D. Euler's φ-function. Euler denoted by $\varphi(n)$ the number of remainders modulo n coprime to n. This is therefore the number of multiplicatively invertible congruence classes in $\mathbb{Z}_n$. The set of such congruence classes is denoted $\mathbb{Z}_n^\times$. Of course, the product of invertible classes in invertible (and $b^{-1}a^{-1}$ is the inverse to ab). We will later learn that $\mathbb{Z}_n$ is a ring, and $\mathbb{Z}_n^\times$ is a group, called the *group of units* of the ring $\mathbb{Z}_n$. Anyhow, $\varphi(n) = |\mathbb{Z}_n^\times|$. This function has the following multiplicative property.

`Proposition.` *$\varphi(mn) = \varphi(m)\varphi(n)$ provided that m and n are coprime.*

`Proof.` Consider the map $\rho : \mathbb{Z}_{mn} \to \mathbb{Z}_m \times \mathbb{Z}_n$ which to the congruence class $\bar{a}$ of an integer a modulo mn assigns the pair of its congruence classes modulo m and modulo n respectively. When $G.C.D(m,n) = 1$, the map is injective, Indeed, if $a \equiv b \mod m$ and $a \equiv b \mod n$, then $a-b$ is divisible by m and by n. It follows from the prime factorization that $a-b$ is divisible by mn, i.e. $a \equiv b \mod mn$. Since both $\mathbb{Z}_{mn}$ and $\mathbb{Z}_m \times \mathbb{Z}_n$ have mn elements, we conclude that ρ is a bijection.

This fact is known as the *Chinese Remainder Theorem.*

Clearly, $\rho(a+b) = \rho(a) + \rho(b)$ and $\rho(ab) = \rho(a)\rho(b)$. In other words, the operations with congruence classes modulo mn agree with the componentwise operations with pairs of congruence classes modulo m and n: $(\alpha', \beta') + (\alpha'', \beta'') = (\alpha' + \alpha'', \beta' + \beta'')$ and $(\alpha', \beta')(\alpha'', \beta'') = (\alpha'\alpha'', \beta'\beta'')$. In particular, since $\rho(\bar{1}) = (\bar{1}, \bar{1})$, a class $a \in \mathbb{Z}_{mm}$ is multiplicatively invertible whenever $(\alpha, \beta) := \rho(a)$ is multiplicatively invertible in $\mathbb{Z}_m \times \mathbb{Z}_n$, i.e. when α is invertible in $\mathbb{Z}_m$ and β is invertible in $\mathbb{Z}_n$. Thus, ρ identifies $\mathbb{Z}_{mn}^\times$ with $\mathbb{Z}_m^\times \times \mathbb{Z}_n^\times$. The cardinalities of these sets are $\varphi(mn)$ and $\varphi(m)\varphi(n)$ by the definition of Euler's function.

`Corollary` (Euler's formula). *For $n = p_1^{r_1} \cdots p_k^{r_k}$ where p_i are distinct primes and $r_i > 0$, we have*

$$\varphi(n) = n\left(1 - \frac{1}{p_1}\right) \cdots \left(1 - \frac{1}{p_k}\right) = \prod_{i=1}^{k} (p_i - 1) p_i^{r_i - 1}.$$

`Proof.` Applying this proposition $k-1$ times we find that $\varphi(n) = \varphi(p_1^{r_1})\varphi(p_2^{r_2} \cdots p_k^{r_k}) = \cdots = \varphi(p_1^{r_1}) \cdots \varphi(p_k^{r_k})$. It remains to notice that out of p^r remainders modulo the r-th power of a prime p, there are p^{r-1} not coprime to p^r: they are $0, p, 2p, \ldots$. Thus, $\varphi(p^r) = p^r - p^{r-1}$.

E. Induction. In the recent arguments we had to use several times the awkward phrase "and so on", meaning that the same reasoning is to be repeated over and over again. In fact this is a hidden way of applying a powerful logical tool called *mathematical induction.*

Suppose we have a sequence $P_1, P_2, \ldots, P_n, \ldots$ of *propositions* i.e. statements which are either true or false. The *principle of mathematical induction* says that in order to confirm all P_n, it suffices to

derive each P_n from the *induction hypothesis* that P_k are true for all $k < n$. Note that the set of $k < n$ is empty for $n = 1$ and non-empty for $n > 1$. So, in practice this amounts to proving two statements:

1° (*base of induction*): P_1 is true;

2° (*step of induction*): For all $n > 1$, P_{n-1} implies P_n.

For example, let P_n be: "If a prime p divides a product $a_0 a_1 \cdots a_n$, then p divides at least one of the factors". The base of induction holds because p is prime, i.e. $p|a_0 a_1$ implies $p|a_0$ or $p|a_1$ by definition. Now, if $p|a_0 \cdots a_{n-1} a_n$, then either $p|a_n$, or $p|a_0 \cdots a_{n-1}$ and therefore (by the induction hypothesis) p divides at least one of $a_0, \ldots, a_{n-1}$. Either way, p divides one of $a_0, \ldots, a_{n-1}, a_n$. According to the principle of mathematical induction all P_n are true.

The principle itself is established the following way. We want to show that the set of n for which P_n is false is empty. If non-empty, we take n_0 to be the smallest integer in this set, and arrive at a contradiction. Indeed, since P_k are true for all $k < n_0$, the induction hypothesis holds for P_{n_0}, and hence P_{n_0} must be true.

We see that the principle of mathematical induction is based on the property of every non-empty set of natural numbers to have a smallest element with respect to the usual ordering of $\mathbb{N} = \{1, 2, \ldots\}$. This property is called *well-ordering*.

By definition, a *partial ordering* on a set X is a binary relation (let us denote it $\preceq$) which is reflexive ($x \preceq x$ for all $x \in X$), transitive ($x \preceq y$ and $y \preceq z$ imply $x \preceq z$), and *anti-symmetric*: $x \preceq y$ and $y \preceq x$ imply $x = y$.

For example, any collection of subsets of any set is partially ordered by inclusion "$\subset$". Another example that will be important to us is the ordering of the set $\mathbb{N}$ of natural numbers by *divisibility*: $m \preceq n$ whenever $m|n$.

These orderings are "partial" in the sense that some pairs of elements can be uncomparable. A partial ordering such that any two elements $x, y \in X$ are comparable, i.e. either $x \preceq y$ or $y \preceq x$, is called *linear ordering* or *total ordering*. For instance, the usual order on the number line is linear. However, many non-empty subsets in $\mathbb{R}$, e.g. the whole of $\mathbb{R}$, or the set $\mathbb{R}_{>0}$ of strictly positive real numbers, have no smallest element. A linearly ordered set $(X, \preceq)$ in which every non-empty subset contains a smallest element is called *well-ordered*.

Given a collection of propositions P_α indexed by elements $\alpha \in (X, \preceq)$ of a well-ordered set, the principle of *transfinite induction*

holds true: if for every $\alpha \in X$, P_α is true whenever P_β are true for all $\beta \prec \alpha$ (i.e. $\beta \preceq \alpha$ but $\beta \neq \alpha$), then P_α are true for all $\alpha \in X$.

`Zermelo's theorem`. *Every set can be well-ordered.*

For example, every *countable set*, i.e. a set that can be put into one-to-one correspondence with $\mathbb{N}$, once such a correspondence is established, becomes well-ordered by the usual order on $\mathbb{N}$. It turns out, however, that for uncountable sets, the well-ordering property, despite being called "theorem", cannot be proved, as it is equivalent to the Axiom of Choice. In particular, uncountable sets (e.g. $\mathbb{R}$) don't have any natural, explicitly describable well-ordering. We refer the reader to Appendix I at the end of the book, where the equivalence of the Axiom of Choice, Zermelo's Theorem, and one more formulation, called *Zorn's Lemma* is established.

On several important occasions we will have to rely on the Axiom of Choice and the principle of transfinite induction, disguised as an application of Zorn's Lemma. We postpone the discussion of it until the need occurs.

F. Groups, rings, and fields. The three parts of the theory are quite distinct in their nature.

In mathematics and physics one deals with *sets* equipped with some *structures*. We have seen sets equipped with an ordering, or an equivalence relation, or arithmetic operations, or sets *per se* (i.e. equipped with no structure). All such objects have *symmetries*, i.e. permutations of the set preserving the structure. By definition, *permutations* are invertible functions from a set to itself. They can be composed to form new permutations. Composing symmetries of a given structure, one obtains new symmetries. The algebraic system thus formed is called a *group*: the group of all symmetries of a given structure. Thus, the theory of groups studies symmetries of any objects.

The theory of rings takes its origin in the arithmetic of integers, which can be added and multiplied. Thus, $\mathbb{Z}$, $\mathbb{Z}_n$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ are examples of rings. However, number-valued functions can also be added and multiplied, and consequently the rings of numbers can be studied in parallel with the rings of functions, occurring in geometry. The main focus in our theory of rings will be on the uniqueness of factorization, and the Fundamental Theorem of Arithmetic will serve us as a model.

The rings $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ of rational, real and complex numbers are actually *fields*, meaning that all their non-zero elements are invertible.

This property is important, for instance, in the definition of a vector space in linear algebra, where the scalars need to form a field (in order to enable, e.g., the "row operations"). Integers modulo a prime p also form a field, $\mathbb{Z}_p$, but in fact there are many more. The theory due to Évariste Galois studies *symmetries of fields* (of which the complex conjugation can serve as the first example). Here the factorization theory in the rings of polynomial functions combines with the theory of groups in order to solve the millenniums-old problems about the possibility of constructing regular polygons by straightedge and compass, and representing the roots of polynomial equations by radicals.

EXERCISES

1. Give examples of matrices which: (a) have right inverse but no left inverse, (b) have left inverse but no right inverse.

2. Give examples of binary relations which possess two of the three properties of reflexivity, symmetricity, and transitivity, but not the third one.

3. Using the Euclidean algorithm, (a) find the greatest common divisor d of 1763 and 991, (b) express d as an integer linear combination of 1763 and 991, and (c) if 991 is invertible modulo 1763, find the inverse.

4. Check that the set R of complex numbers of the form $a + b\sqrt{-5}$, where a, b are arbitrary integers, is closed with respect to addition and multiplication, and show that 2 and 3 are irreducible in R but not prime.

5. Compute multiplicative inverses of all elements in: (a) $\mathbb{Z}_8^\times$, (b) $\mathbb{Z}_{10}^\times$.

6. Establish explicitly the bijection $\rho : \mathbb{Z}_6 \to \mathbb{Z}_2 \times \mathbb{Z}_3$.

7. Prove that $\varphi(mn) \neq \varphi(m)\varphi(n)$ whenever m and n are not coprime.

8. For prime p, show that $p|\binom{p}{k} := n!/k!(n-k)!$ when $0 < k < p$. Using the binomial formula and induction on $a = 0, 1, 2, \ldots, p-1$, derive *Fermat's Little Theorem*: $a^p \equiv a \mod p$.

9. Verify the principle of transfinite induction.

10. Show that $\mathbb{Z}_p$ is a field if p is prime.

11. Let $ABCD$ be a rectangle which is not a square, $X = \{A, B, C, D\}$ the set of its vertices, and $Y = \{V, H, S\}$ the set of three objects, namely of the 3 partitions of X into two pairs: vertices connected by vertical edges (say, AB and DC), by horizontal edges (BC and AD), and by diagonals (AC and BD). Study how permutations on X induce those on Y. Show that each of the 6 permutations of Y is induced by 4 permutations of X, and describe geometrically which four induce the identity permutation on Y.

Groups

Lecture 3. Groups and isomorphisms

A. Groups of symmetries. Most expositions of the group theory begin, naturally, with the definition of an abstract group. So, let's for a change begin ours with examples. As we mentioned at the end of the previous lecture, groups usually arise in studying symmetries of various structures.

Let X be a set. An invertible function $\sigma : X \to X$ is called a *permutation* of X. Such permutations can be considered as symmetries of the set X equipped with no structure. The composition $\sigma \circ \lambda$ of two permutations (i.e. λ followed by σ) is a permutation again. Its inverse $(\sigma \circ \lambda)^{-1} = \lambda^{-1} \circ \sigma^{-1}$ (in this order: undoing first σ, and then λ). All such permutations (including the *identity* permutation id_X) form a group, denoted $S(X)$, and called the *group of permutations* on the set X.

When $X = \{1, \dots, n\}$, the group $S(X)$, usually denoted S_n, is called the *symmetric group* on n objects. It consists of $n!$ elements.

It is harder to imagine the group of all permutations of an infinite set, e.g. the set $\mathbb{R}^2$ of points on the plane. Here $S(\mathbb{R}^2)$ consists of all bijective functions $\sigma : \mathbb{R}^2 \to \mathbb{R}^2$ (moving the points around with no regard for any geometric properties we usually associate with the plane).

Now equip the plane $\mathbb{R}^2$ with the *topological structure*, i.e. information on which sequences of points converge and to what limits. A permutation $\sigma : \mathbb{R}^2 \to \mathbb{R}^2$ such that both σ and σ^{-1} are continuous, i.e. such that $\lim p_n = p_\infty$ implies $\lim \sigma(p_n) = \sigma(p_\infty)$ and $\lim \sigma^{-1}(p_n) \to \sigma^{-1}(p_\infty)$, is called a *homeomorphism* of $\mathbb{R}^2$. All homeomorphisms form a *subgroup* $H(\mathbb{R}^2) \subset S(X)$, the group of symmetries of $\mathbb{R}^2$ considered as a topological space.

If, in addition to be homeomorphisms, the permutations are required to be differentiable (say, infinitely differentiable, to be definite), then they become symmetries of the *smooth structure* on $\mathbb{R}^2$, and form the *group of diffeomorphisms* $D(\mathbb{R}^2) \subset H(\mathbb{R}^2)$ — the group of all invertible (and possibly non-linear) changes of coordinates on the plane.

The much smaller supply of changes of coordinates, the *linear inhomogeneous* ones:

$$(x, y) \mapsto (ax + by + x_0,\ cx + dy + y_0),$$

where $ad-bc$ must be non-zero to guarantee invertibility, form a group $A(\mathbb{R}^2)$ of *affine transformations* of the plane, i.e. the symmetries of the affine structure. They are compositions of homogeneous linear transformations (these have $x_0 = y_0 = 0$) with translations $(x, y) \mapsto (x + x_0, y + y_0)$.

The translations do not preserve the origin. If, however, the origin on the affine plane is considered a part of the structure, the plane $\mathbb{R}^2$ becomes a 2-dimensional real vector space. The symmetries of it are only the invertible homogeneous linear transformations:

$$\begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}, \quad ad - bc \neq 0.$$

The composition of such linear transformations results in the multiplication of the corresponding coefficient matrices. Therefore the group can be identified with that of invertible real 2×2-matrices, called the *general linear group* $GL_2(\mathbb{R})$.

Adding to the structure of the vector space the *orientation* of the plane, i.e. a choice of one of the two directions of rotation: clockwise or counter-clockwise, we obtain in the role of symmetry group the group $GL_2^+(\mathbb{R})$ of orientation-preserving linear transformations. In matrix terms, it consists of real 2×2-matrices with positive determinants: the corresponding linear transformations keep the clockwise direction of rotation clockwise, and counter-clockwise counter-clockwise, while the matrices with $ad - bc < 0$ interchange them.

Since the determinant is multiplicative, $\det AB = (\det A)(\det B)$, the matrices with $\det = 1$ form a subgroup in $GL_2^+(\mathbb{R})$, denoted $SL_2(\mathbb{R})$ and called the *special linear group*. It is the symmetry group of the plane equipped with the structures of a vector space, orientation, and *area*: the absolute value $|\det A|$ equals the factor by which the areas of regions on the plane are changed by the transformation $\vec{u} \mapsto A\vec{u}$.

Returning to the structureless $\mathbb{R}^2$ and then equipping it with the Euclidean distance structure, we obtain in the role of symmetries the group $E(\mathbb{R}^2)$ of *rigid motions* of the plane. Though it is not entirely obvious, all rigid motions are affine. They can be described as compositions of translations $(x, y) \mapsto (x + x_0, y + y_0)$ with *orthogonal transformations*, i.e. linear transformations preserving the dot-product structure on $\mathbb{R}^2$. The orthogonal transformations *per se* form the *orthogonal group* of $\mathbb{R}^2$, denoted O_2. Of course, such transformations have $\det = \pm 1$ (because when distances are preserved, the areas are preserved too). It is not too hard to show that orthogonal transformations with determinant $+1$ (thus, orientation-preserving ones) are rotations of the plane about the origin. They form the *special orthogonal group* of the plane denoted SO_2. An orthogonal transformation of $\mathbb{R}^2$ which has determinant -1 (i.e. reverses the orientation) is in fact a reflection about a line passing through the origin.

One can draw on the Euclidean plane any figure and ask about symmetries of it. While a random figure will have no symmetries (i.e. its symmetry group will consist only of the identity permutation of the plane), some symmetric figures will have more. For example, an equilateral triangle ABC has 6 symmetries: three reflections about its symmetry axes, and three rotations (through $0°$, $120°$, and $-120°$) about the triangle's center. In fact these symmetries permute the vertices A, B, C arbitrarily, identifying thereby the group $S(\triangle)$ of Euclidean symmetries of the equilateral triangle with the group S_3 of $3! = 6$ permutations of $\{A, B, C\}$.

Here is our chart of the symmetry groups of the above structures:

$$\begin{array}{ccccccccccc} S(\mathbb{R}^2) & \supset & H(\mathbb{R}^2) & \supset & D(\mathbb{R}^2) & \supset & A(\mathbb{R}^2) & \supset & GL_2(\mathbb{R}) & \supset & GL_2^+(\mathbb{R}) & \supset & SL_2(\mathbb{R}) \\ & & & & & & \cup & & \cup & & \cup & & \cup \\ & & & & & & E(\mathbb{R}^2) & \supset & O_2 & \supset & SO_2 & = & SO_2 \\ & & & & & & & & \cup & & & & \\ & & & & & & & & S(\triangle) & & & & \end{array}$$

B. Abstract groups. In the above examples of groups, the algebraic operation was the same: the composition of functions. In the abstract definition of a group the nature of the operation is not specified. The reader probably can guess what the definition is.

A *group* is defined as a set, G, equipped with a binary operation, i.e. a function $G \times G \to G$, which to every ordered pair (a, b) of elements from G associates another element from G, called their product, and denoted ab, in such a way that the following three requirements (often referred to as "axioms") hold:

(i) $(ab)c = a(bc)$ for all $a, b, c \in G$ (*associativity*);

(ii) there exists a unique element $e \in G$, called the *identity*, such that $ae = a = ea$ for all $a \in G$;

(iii) for every element $a \in G$, there exists a unique element, a', called the *inverse* of a (and often denoted a^{-1}) such that $aa' = e = a'a$.

C. Remarks. (a) A minimalist could complain that some of the above requirements are redundant. For instance, one does not need to assume the uniqueness of the identity, nor that it is two-sided, for it suffices to assume only the existence of at least one e_L, and one e_R such that $e_L a = a$ and $ae_R = a$ for all $a \in G$. Indeed, this implies $e_L = e_L e_R = e_R$, i.e. that every left identity is equal to every right one, meaning that it is unique and two-sided. Likewise, if a'_L and a'_R are a left and a right inverse to a, then $a'_R = ea'_R = (a'_L a)a'_R = a'_L(aa'_R) = a'_L e = a'_L$, i.e. every left inverse is equal to every right inverse, implying that the inverse is unique and two-sided.

(b) In a group, multiplying $ac = bc$ ($ca = cb$) by c^{-1} on the right (resp. on the left) we derive $a = b$, i.e. the *cancellation rule* holds. Likewise, the equations $ax = b$ and $ya = b$ have unique solutions: $x = a^{-1}b$ and $y = ba^{-1}$. Some authors use the latter properties (instead of (ii) and (iii)) in the definition of a group.

(c) There is an attractive point of view on groups, according to which a group carries three operations: one *binary operation* of multiplication $G \times G \to G : (a, b) \mapsto ab$, another *unary operation* of inversion: $G \to G : a \mapsto a^{-1}$, and one more *"nullary" operation*, with no input but one output: $e \in G$. Then the axioms (i),(iii),(ii) are to specify the mutual properties of the operations.

(d) The notation a^n is often used for $a \cdots a$ (n factors) when $n > 0$, for $a^{-1} \cdots a^{-1}$ ($|n|$ factors) when $n < 0$, and for e when $n = 0$. One should check (by case study, as it is done for numbers in the middle school) that $a^m a^n = a^{m+n}$ for all integer m and n regardless of their signs.

(e) As it is clear from our examples of symmetry groups, the group operation is often non-commutative: ab generally speaking differs from ba. When $ab = ba$ for all elements $a, b \in G$, the group is called *commutative*, or *abelian*. In an abelian group, one often calls the

operation "addition" (rather than product), and denotes it by "+". Respectively, the identity element e is replaced with 0, the inverses a^{-1} become $-a$, and the powers a^m become ma, where $m \in \mathbb{Z}$ and $a \in G$.

(f) Perhaps it is useful to have examples of non-groups. The set $\mathbb{N}$ of natural numbers is not a group with respect to the addition operation, because 0 isn't there. Adjoining 0, we obtain $\mathbb{Z}_+$, which is still not a group, because $-a \notin \mathbb{Z}_+$ when $a > 0$. Adjoining the opposites, we obtain $\mathbb{Z}$, the set of all integers, which is an abelian group with respect to the operation of addition. By the way, the operation in this group does not originate from composition of functions (as it was in the previous examples), which perhaps justifies the idea of defining groups abstractly. The same is true about the modular arithmetic: integers $\mod n$ form an abelian group $\mathbb{Z}_n$ with respect to the operation of addition of congruence classes. Relative to multiplication, $\mathbb{Z}$ is not a group, because the division is not always possible. Adjoining fractions, we obtain the set $\mathbb{Q}$ of rational numbers, which still don't form a group with respect to multiplication, because 0 does not have a multiplicative inverse. Removing 0, we obtain the set of non-zero rationals, which do form an abelian group with respect to multiplication. Another non-example: reflections of the plane about the origin do not form a group, because the composition of two reflections is not a reflection: it reverses the orientation twice, and hence defines a rotation of the plane.

D. The axiomatic method. An instance of it was our postulating the axioms (i), (ii) , (iii) of a group and deriving consequences such as e.g. the cancellation properties. According to a famous quote from Bertrand Russell, *The method of 'postulating' what we want has many advantages; they are the same as the advantages of theft over honest toil.* So, what exactly are our theft-like advantages?

We should warn the reader that the above trivial manipulations with letters based directly on the axioms and definitions do *not* constitute the essence of group theory. In the contrary, the same way as in any meaningful mathematical theory, the goal is to obtain deep and non-trivial information about the world out there (i.e. about symmetries of various structures in the case of group theory).

Furthermore, what we called "axioms" are actually not "truths accepted without proof". Every concrete application of their consequences must be preceded by *proving* the axioms in that particular example. The advantage here consists in the shifting the "honest toil" from directly verifying all consequences in each example to verifying

the axioms. Thus, the axiomatic method here serves as a device for unification of multiple concrete examples. Moreover, as we shall see, the derivations of basic properties of algebraic operations are often more convincing when done by rearrangement of abstract letters without any reference to the actual nature of the objects at hand.

Another motivation for the axiomatic approach comes from the problem of group classification: In order to formulate such a problem, one needs to have an abstract description of the objects to be classified, and a notion of equivalence between them.

E. The problem of classification. How many different groups are out there? Well, infinitely many, because the *orders* (i.e. the numbers of elements) of finite groups (e.g. $\mathbb{Z}_n$) can take on arbitrary finite values. So, let's be more specific: How many different *finite groups of a given order* are there? One answer is still "infinitely many", e.g. because one can consider an unlimited number of copies of a given group. On the other hand, all these copies have the same group structure. So, the interpretation of the question should be: How many *really* different groups of order n are there?

Actually this question is even hard to formulate without the notion of *isomorphism*. A group G is called *isomorphic* to a group G' if there exists a bijection $f : G \to G'$ which identifies the operation in G with the operation in G', i.e. for any $x, y \in G$, the image $f(xy)$ of their product xy in G coincides with the product $f(x)f(y)$ of their images $f(x)$ and $f(y)$ in G'.

This notion serves as an equivalence relation, denoted $\cong$. Namely, the identity map $\mathrm{id}_G : G \to G$ is an isomorphism of G with itself; if $f : G \to G'$ is an isomorphism between G and G', then $f^{-1} : G' \to G$ is an isomorphism between G' and G; when $g : G' \to G''$ is another isomorphism, then $g \circ f : G \to G''$ is an isomorphism too. Thus, all groups are divided into *isomorphism classes*. The classification problem of, say, finite groups can be now stated as the task of finding the number (and a list of representatives) of the isomorphism classes of groups of a given finite order n.

We can try to start right now. Clearly, all groups of order 1 are isomorphic, as they are *trivial*, consisting of only the identity element e. A group of order 2 contains also an element $g \neq e$, whose square $g^2 = e$ (for $g^2 = g$ would imply $g = e$ by cancellation). Thus all groups of order 2 are also isomorphic (to $\mathbb{Z}_2$, as well as to the multiplicative group $\mathbb{Z}^\times = \{\pm 1\}$ of invertible integers). Actually a group of order 3 is also unique *up to isomorphism*, but two groups of order 4 can be non-isomorphic (see Exercises).

EXERCISES

12. Show that every group of order 3 is isomorphic to $\mathbb{Z}_3$.

13. Show that if $g^2 = 2$ for all $g \in G$, then G is abelian.

14. Compare the following groups of order 4, and study the operation (you may write the whole table of multiplication, if you think it helps) to find out which of the groups are isomorphic to each other and which are not.

A. The group $\mathbb{Z}_{10}^{\times}$ of multiplicatively invertible congruence classes modulo 10 (it consists of $1, -1, 3, -3$); the operation is multiplication mod 10.

B. $\mathbb{Z}_8^{\times}$ (the same, but modulo 8).

C. $\mathbb{Z}_2^2$, the group of 2-dimensional vectors (a, b), where $a, b \in \mathbb{Z}_2$ (i.e $\bar{0}$ or $\bar{1}$, as in computer science) with the operation of componentwise addition modulo 2.

D. Complex numbers $1, -1, i, -i$ with respect to multiplication of complex numbers.

E. The group of all symmetries of a rectangle (i.e. of all rigid motions of the plane which preserve the rectangle), with the operation of composition.

F. In the group S_4 of all permutations of $\{1, 2, 3, 4\}$, consider the subgroup (denoted K_4 and called the *Klein subgroup*) of all those permutations which preserve each of the three partitions of $\{1, 2, 3, 4\}$ into pairs: $\{1, 2|3, 4\}$, $\{1, 3|2, 4\}$, and $\{1, 4|2, 3\}$. (This description may seem convoluted, but the very existence of this subgroup is a remarkable fact of Nature with serious consequences, and is worth studying).

G. In the group of all rotations of the unit cube $|x|, |y|, |z| \leq 1$ (operation = composition of rotations), consider the subgroup of all those rotations which preserve each of the three coordinate axes. (In other words, each rotation of the cube somehow permutes the three axes, but for some nontrivial rotations this permutation is trivial; these rotations together with the identity form the subgroup.)

15. Find out which of the following groups are isomorphic to each other and which are not:

H. The group D_3 of all rotations of the 3-space preserving the regular triangular prism (by definition, its bases are parallel equilateral triangles, and the side faces are rectangles perpendicular to the bases).

I. $GL_2(\mathbb{Z}_2)$, the group of 2×2-matrices whose entries are integers mod 2, equipped with the operation of matrix multiplication, and invertible with respect to this operation.

J. S_3, the group of permutations of $\{1, 2, 3\}$.

K. The group of two generators a, b satisfying the relations $a^2 = e, b^2 = e, (ab)^3 = e$.

`Remark.` This is a novel way for us to describe a group, so let's discuss it. Consider the alphabet consisting of two letters a and b. One actually needs to add two more letters, a^{-1} and b^{-1}, but in the example K the relations

$a^2 = e$ and $b^2 = e$ show that a and b are their own inverses. Elements of the group are represented by words in this alphabet. The operation is concatenation of words. The empty word represents the identity element. The main convention is that two words represent the same group element if they are obtained from each other by cancellations that use the given relations and the axioms of a group. For example: $(aba)(aba) = abaaba = abba = aa = e$; $(babab)(aba) = b(ababab)a = ba$, for in the example K, not only $a^2 = b^2 = e$, but also $ababab = e$. Multiplying $ababab = e$ by bab on the right, we find also that $aba = bab$.

Lecture 4. Homomorphisms

A. The category of groups. In modern mathematics, it is considered insufficient to describe the *objects* of interest; to introduce a *category* of objects one also needs to specify the *morphisms*, i.e. the kinds of maps between the objects which are considered legitimate.

In the category of groups, the morphisms are *group homomorphisms*. By definition they are functions $f : G \to G'$ respecting the group operations:

$$f(ab) = f(a)f(b) \text{ for all } a, b \in G,$$

i.e. mapping the products in the domain group G into the products of the corresponding elements in the codomain group G'.

In accordance with Remark (c) in section C of Lecture 3, one should also require that a group homomorphism respects the other two operations, i.e. $f(e) = e'$, and $f(a^{-1}) = f(a)^{-1}$ for all $a \in G$. However, these two properties follow from the first one. Indeed, $f(e)e' = f(e) = f(e^2) = f(e)^2$ implying $e' = f(e)$ by cancellation. Consequently, $e' = f(e) = f(aa^{-1}) = f(a)f(a^{-1})$ showing that $f(a^{-1}) = f(a)^{-1}$.

B. Subgroups. The notion of a homomorphism generalizes that of isomorphism, which by definition is a *bijective* homomorphism. A surjective homomorphism is also called an *epimorphism*, and an injective one a *monomorphism*.

We have seen many monomorphisms, when in Section A of the previous lecture we considered a series of examples of one group, G, containing another group, H. By definition, a *subgroup* $H \subset G$ in a group G is a subset which is a group on its own with respect to the same operations as in G. That is, H must contain the identity element of G, together with every element $a \in H$ contain its inverse, and together with any two elements $a, b \in H$ contain their product. The properties (i),(ii),(iii) then hold in H simply because they hold in G.

Here is, however, a trivial but useful observation. In order to show that a (non-empty) subset $H \subset G$ is a subgroup, instead of checking those three properties, it suffices to check only one other: that together with any two $a, b \in H$, their *ratio* ab^{-1} is in H. Indeed, then for $a_0 \in H$, the ratio $a_0 a_0^{-1} = e \in H$. Next, whenever $a \in H$, the ratio $ea^{-1} = a^{-1} \in H$. Finally, whenever $a, b \in H$, since $b^{-1} \in H$ too, we find that the ratio $a(b^{-1})^{-1} = ab \in H$.

Every group monomorphism $H \to G$ can be considered as an embedding of H as a subgroup into G. More generally, given a group homomorphism $f : G \to G'$, one can associate two subgroups with it. One is the *range* $f(G)$, which is a subgroup in G'. Indeed, if $f(a), f(b)$ are in the range, their ratio $f(a)f(b)^{-1} = f(ab^{-1})$ is in the range too. This subgroup is the whole of G' exactly when f is an *epi*morphism.

The other subgroup associated with f, called the *kernel* of f, and denoted $\ker f$, is the inverse image $f^{-1}(e')$ of the identity element in G': $\ker f := \{a \in G \mid f(a) = e'\}$. It is, indeed, a subgroup in G: if $f(a) = e' = f(b)$, then $f(ab^{-1}) = f(a)f(b)^{-1} = e'$. This subgroup consists of only the identity element exactly when f is a *mono*morphism. Indeed, $f(a) = f(b)$ is equivalent to $e' = f(a)f(b)^{-1} = f(ab^{-1})$, i.e. to $ab^{-1} \in \ker f$, which implies $a = b$ whenever $\ker f = \{e\}$.

C. Cyclic subgroups. To every element $g \in G$, there corresponds a homomorphism $\mathbb{Z} \to G : \quad n \mapsto g^n$ of the additive group of integers to G. This is another way to say that $g^m g^n = g^{m+n}$ for all $m, n \in \mathbb{Z}$. The range of this homomorphism is called the *cyclic subgroup* generated by g.

In general, for any subset $S \subset G$, one can introduce the *subgroup generated by* S as the smallest subgroup containing S. It is the intersection of all subgroups containing S. (Checking the fact that intersection of subgroups is a subgroup we leave as an exercise to the reader.) Obviously, the cyclic subgroup of g is the smallest subgroup containing g.

The kernel of the homomorphism $n \mapsto g^n$ is a subgroup in $\mathbb{Z}$. When it is $\{0\}$, the homomorphism is "mono", and hence an isomorphism onto its range. In this case the cyclic group of g is isomorphic to $\mathbb{Z}$; it is an *infinite cyclic group*. Alternatively, when the kernel is non-trivial, there exists the smallest positive d such that $g^d = e$. It follows (by long division) that any n such that $g^n = e$ must be a multiple of d, i.e. the kernel is $d\mathbb{Z} \subset \mathbb{Z}$. In simple words, the powers g^n are distinct for $n = 0, 1, \ldots, d-1$, and then $g^d = g^0 = e$, and the powers begin to repeat cyclically: $g^{d+1} = g$, $g^{d+2} = g^2$, and so on (and the same for negative n: $g^{-1} = g^{d-1}$, and so on). In this case, the cyclic group generated by g is finite, and is isomorphic to the additive group $\mathbb{Z}_d$ of congruence classes modulo d.

D. Real-world homomorphisms. Here are some examples of groups homomorphisms familiar to the readers from their previous lives.

The exponential function satisfies $e^{x+y} = e^x e^y$ and thus defines a group homomorphism from the *additive* group $(\mathbb{R}, +)$ or real numbers to the group $\mathbb{R}^\times$, the *multiplicative* group of non-zero real numbers. The range of it consists of positive reals, and the logarithmic function $\log : \mathbb{R}^\times_{>0} \to (\mathbb{R}, +)$ defines the inverse homomorphism: $\log(xy) = \log x + \log y$. Thus the additive group of all real numbers is isomorphic to the multiplicative group of positive real numbers.

According to the *Euler formula*, $e^{ix} = \cos x + i \sin x$, i.e. function $x \mapsto e^{ix}$ defines a 2π-periodic mapping of the number line onto the circle $U_1 = \{z \in \mathbb{C} \mid |z| = 1\}$ of unit complex numbers. The circle is a group, with the operation interpreted as the multiplication of complex numbers, or identified (by an isomorphism!) with the group SO_2 of rotations of the plane. The periodic map is a group *epimorphism* of $(\mathbb{R}, +)$ onto U_1.

The natural projection $\pi : \mathbb{Z} \to \mathbb{Z}_n$ of integers to the set of congruence classes modulo n is a group homomorphism with respect to the operations of addition (introduced in $\mathbb{Z}_n$ in Lecture 2, and in $\mathbb{Z}$ much earlier). The same is true about the map $\rho : \mathbb{Z}_{mn} \to \mathbb{Z}_m \times \mathbb{Z}_n$ used in the proof of Euler's formula.

As it should be familiar from linear algebra, a map $A : V \mapsto W$ between two vector spaces is called linear if it maps linear combinations of vectors to linear combinations of their images with the same coefficients:

$$A(\lambda\vec{u} + \mu\vec{v}) = \lambda A\vec{u} + \mu A\vec{v}.$$

In fact vector spaces are abelian groups with respect to the addition operation, and a linear map is a group homomorphism: $A(\vec{u} + \vec{v}) = A\vec{u} + A\vec{v}$.

The determinant of 2×2-matrices (and more generally, of $n \times n$-matrices) is multiplicative, $\det(AB) = (\det A)(\det B)$. Consequently, *invertible* $n \times n$-matrices have non-zero determinant, and form a group with respect to matrix multiplication. It is denoted $GL_n(\mathbb{R})$ (when the matrix entries are assumed to come from $\mathbb{R}$), and called the *general linear group*. The determinant function $\det : GL_n(\mathbb{R}) \to \mathbb{R}^\times = GL_1(\mathbb{R})$ is a group homomorphism. Of course, the same is true not only for real matrices.

E. Even and odd permutations. The last example prompts us to recall from linear algebra the definition of $n \times n$-determinants:

$$\det \begin{bmatrix} a_{11} & \dots & a_{1n} \\ & \dots & \\ a_{n1} & \dots & a_{nn} \end{bmatrix} := \sum_{\sigma \in S_n} \epsilon(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)}.$$

It is the sum of $n!$ signed *elementary products*, where $\epsilon(\sigma) = \pm 1$ is the *sign* of permutation $\sigma = \begin{pmatrix} 1 & \dots & n \\ \sigma(1) & \dots & \sigma(n) \end{pmatrix}$. Here is how the sign is defined.

Introduce the following polynomial in n variables $x_1, \dots, x_n$:

$$\Delta_n(x_1, \dots, x_n) := \prod_{i>j} (x_i - x_j).$$

For instance,

$$\Delta_1 = 1, \quad \Delta_2 = x_2 - x_1, \quad \Delta_3 = (x_2 - x_1)(x_3 - x_1)(x_3 - x_2), \quad \dots$$

A permutation $\sigma \in S_n$ acts on any polynomial P in n variables by permuting the variables and producing a new polynomial,

$$(\sigma P)(x_1, \dots, x_n) := P(x_{\sigma(1)}, \dots, x_{\sigma(n)}).$$

For example, for $\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$, we have

$$\sigma\Delta_3 = (x_1 - x_3)(x_2 - x_3)(x_2 - x_1) = (-1)^2 (x_2 - x_1)(x_3 - x_1)(x_3 - x_2).$$

More generally, applying any $\sigma \in S_n$ to Δ_n results in $\pm\Delta_n$. Indeed, permuting the indices $\{1, \dots, n\}$, the permutation σ also permutes the set of $\binom{n}{2}$ *pairs* of indices $\{i, j\}$. Thus, the $\binom{n}{2}$ linear factors in Δ_n are permuted as each $x_i - x_j$ with $i > j$ is transformed into $x_{\sigma(i)} - x_{\sigma(j)}$. When $\sigma(i) > \sigma(j)$, the transformed factor occurs in the product $\sigma\Delta_n$ with the same sign as it does in Δ_n, and when $\sigma(i) < \sigma(j)$, with the opposite sign. Thus,

$$\sigma\Delta_n = (-1)^{l(\sigma)} \Delta_n,$$

where $l(\sigma)$, called the *length of permutation*, equals the number of pairs $i > j$ such that $\sigma(i) < \sigma(j)$. One could say that σ acts by a linear transformation in the space of polynomials of degree n, and Δ_n is an eigenvector of σ with the eigenvalue $\epsilon(\sigma) := (-1)^{l(\sigma)}$.

`Proposition.` $\epsilon : S_n \to \{\pm 1\}$ *is a group homomorphism:*

$$\epsilon(\sigma \circ \sigma') = \epsilon(\sigma)\epsilon(\sigma') \quad \textit{for all } \sigma, \sigma' \in S_n.$$

`Proof.` It is tautological: $\sigma'\Delta_n = \epsilon(\sigma')\Delta_n$; applying σ to this, we get $\epsilon(\sigma')\epsilon(\sigma)\Delta_n$. But by the definition of ϵ, the composition $\sigma \circ \sigma'$ transforms Δ_n into $\epsilon(\sigma \circ \sigma')\Delta_n$. Thus $\epsilon(\sigma \circ \sigma') = \epsilon(\sigma')\epsilon(\sigma)$. It remains to notice that multiplication in the group $\{\pm 1\}$ is commutative.

A permutation σ is said to be *even* if $\epsilon(\sigma) = 1$ and *odd* if $\epsilon(\sigma) = -1$. Thus, even permutations form the kernel of the sign homomorphism ϵ. It is a subgroup in S_n, denoted A_n, and called the *alternating group* on n objects.

EXERCISES

16. Show that if $\ker f \neq \{e\}$, then f is not "mono".

17. Check that if a group homomorphism is bijective then the inverse map is a group homomorphism too.

18. Show that intersection of (two or more) subgroups is a subgroup.

19. Prove that in a cyclic group, all subgroups are cyclic. `Hint:` Check this first for subgroups in $\mathbb{Z}$.

20. How many distinct cyclic subgroups are there in the group of symmetries of a regular triangle?

21. The symmetry group of a regular n-gon ($n \geq 3$) is denoted D_n is called the nth *dihedral group*. Show that $|D_n| = 2n$ and describe all cyclic subgroups in D_n.

22. Consider the map $\exp : \mathbb{C} \to \mathbb{C}^\times$ which to a complex number $z = x + iy$ associates $\exp(z) = e^x e^{iy} = e^x(\cos y + i \sin y)$. Show that this an epimorphism of the additive group of complex numbers onto the multiplicative group of non-zero complex numbers, but not an isomorphism between them, and find its kernel.

23. Compute the length of the *transposition* τ_{ij}, $i < j$, defined as the permutation on $\{1, \ldots, n\}$ which swaps i and j: $\tau_{ij}(i) = j$, $\tau_{ij}(j) = i$, $\tau_{ij}(k) = k$ for all $k \neq i, j$. Derive that transpositions are odd.

24. Compute the kernel and range of the homomorphism $\rho : \mathbb{Z}_{mn} \to \mathbb{Z}_m \times \mathbb{Z}_n$ when m and n are not coprime.

25. Write out the explicit formula for 2×2 and 3×3 determinants following the definition of $n \times n$-determinants as the sums of $n!$ signed elementary products.

26. Show that the action of S_4 on the set of 3 partitions of $\{1, 2, 3, 4\}$ into two pairs defines a homomorphism $\pi : S_4 \to S_3$. Prove that $\ker \pi = K_4$, and that π maps even permutations to even and odd to odd: $\epsilon \circ \pi = \epsilon$.

Lecture 5. Cosets

A. Two equivalence relations. Let $H \subset G$ be a subgroup. We say that $a, b \in G$ are *left-congruent modulo* H, and write $a \ {}_H\!\equiv b$, whenever $b^{-1}a \in H$, and respectively say that they are *right-congruent modulo* H, and write $a \equiv_H b$, whenever $ab^{-1} \in H$. These are equivalence relations. Indeed, since H contains e, and is closed with respect to inversion and multiplication, we have: $a^{-1}a = e \in H$, $b^{-1}a \in H$ implies $a^{-1}b = (b^{-1}a)^{-1} \in H$, and $a^{-1}b, b^{-1}c \in H$ imply $a^{-1}c = (a^{-1}b)(b^{-1}c) \in H$ for any $a, b, c \in G$. That is, ${}_H\!\equiv$ is reflexive, symmetric, and transitive, and one can similarly check this for $\equiv_H$.

Thus, the group G is partitioned into equivalence classes with respect to ${}_H\!\equiv$ ($\equiv_H$), which are called *left H-cosets* (*right H-cosets* respectively). Note that the left H-coset of $a \in G$ consist of all elements of the form ah where $h \in H$, and moreover, the map $H \ni h \mapsto ah$ (which can be interpreted as the *left translation* by a) establishes a 1-to-1 correspondence between the subgroup and the coset. Likewise, the *right translation* by a, $H \ni h \mapsto ha$, establishes a 1-to-1 correspondence between H and the right H-coset of a. Consequently, we will often denote the cosets as aH and Ha respectively. While the left and right H-cosets of e coincide ($eH = H = He$), this is not necessarily true for $a \neq e$.

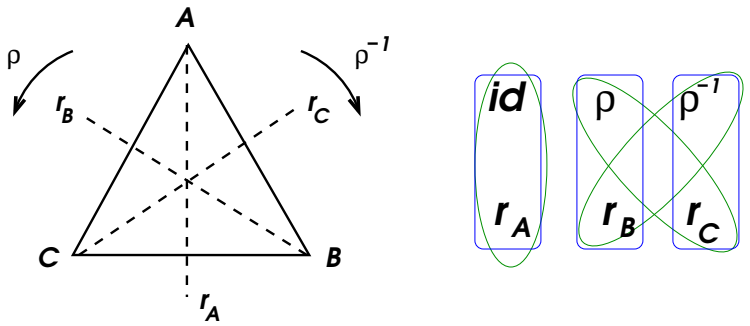


Figure 3: Left (blue) and right (green) H-cosets

`Example.` The symmetry group of a regular triangle ABC consists of three reflections r_A, r_B, r_C in the symmetry lines passing through the vertices (Figure 3), rotations: ρ (through 120° counterclockwise), ρ^{-1} (through 120° clockwise), and the identity id. Let H be the cyclic group of r_A. Then (as it is not hard to conclude from the picture), the left H-cosets are: $H = \{\mathrm{id}, r_A\}$, $r_B H = \{r_B, \rho\}$

(because $r_B r_A(A) = r_B(A) = C$ indicates the counter-clockwise rotation), and $r_C H = \{r_C, \rho^{-1}\}$. Likewise, the right H-cosets are: $H = \{\mathrm{id}, r_A\}$, $Hr_B = \{r_B, \rho^{-1}\}$, and $Hr_C = \{r_C, \rho\}$. Here $r_B H$, $r_C H$, Hr_B, Hr_C are all distinct.

B. Lagrange's theorem. From the fact that all parts of either partition of a group G — into left or right H-cosets — are in bijection with the subgroup H, we obtain:

`Theorem.` *In a finite group, the order of any subgroup divides the order of the whole group: if $|G| < \infty$, then $|H| \mid |G|$.*

We remind that the order of a finite group is the number of elements in it. The order of the cyclic subgroup generated by an element $g \in G$, i.e. the least positive power n such that $g^n = e$, is called the *order of the element.*

`Corollary 1.` *The order of every element in a finite group divides the order of the group.*

`Corollary 2.` *Every finite group of prime order is cyclic.*

Indeed, the cyclic subgroup of $g \neq e$ in a group of prime order p must have order p and hence coincide with the whole group.

`Examples.` Thus, the groups of orders $n < 8$ are necessarily isomorphic to $\mathbb{Z}_n$ for $n = 1, 2, 3, 5, 7$. A non-cyclic group of order 4 consists of e and three more elements a, b, c, which by Corollary 2 have order 2 (i.e. $a^2 = b^2 = c^2 = e$), and satisfy the property that the product of any two of them is equal to the third one (e.g. $ab = c$, for if $ab = e$ or $= a$, or $= b$ would imply $b = a$, $b = e$, $a = e$ respectively). This completely determines the "multiplication table" in the group, showing that up to isomorphism, there is only one non-cyclic group of order 4. The Klein group K_4 fits the bill. For $n = 6$, besides $\mathbb{Z}_6$, we know also the (smallest non-abelian!) group of symmetries of the triangle (which is also isomorphic to the group S_3 of permutations of the vertices). In fact (although we don't know this yet) up to isomorphism, there are no other groups of order 6 other than $\mathbb{Z}_6$ and S_3.

`Corollary 3.` *In a finite group G, $g^{|G|} = e$ for all $g \in G$.*

Indeed, if n is the order of g, then $|G| = nl$ for some l, and $g^n = e$. Thus $g^{nl} = e^l = e$.

`Corollary 4.` (*Fermat's Little Theorem*). *If p is prime and a is not divisible by p, then $a^{p-1} \equiv 1 \mod p$.*

This is Corollary 3 applied to the group $\mathbb{Z}_p^\times$ of non-zero congruence classes of integers mod p with respect to the operation of multiplication. Generalizing this argument, consider the group $\mathbb{Z}_n^\times$ of all invertible congruence classes of integers modulo n, i.e. the classes of integers a coprime with n. The number of such classes is denoted $\varphi(n)$ and is known as *Euler's function* of n. E.g. for a prime p, $\varphi(p) = p - 1$.

`Corollary 5` (Euler's theorem). *For a and n coprime, $a^{\varphi(n)} \equiv 1 \mod n$.*

C. Normal subgroups. A subgroup $H \subset G$ is called *normal* if its left cosets coincide with its right cosets: $aH = Ha$ for all $a \in G$, or in other words, if the equivalence relations $\equiv_H$ and $_H\equiv$ coincide. Equivalently, H is normal if and only if it stays invariant under the transformations of *conjugation* by any elements of G. Namely, the conjugation by $a \in G$ on G is defined as $g \mapsto aga^{-1}$ (and is a bijection since the conjugation by a^{-1} is its inverse). The condition $aH = Ha$ means that for every $h \in H$ there exist $h', h'' \in H$ such that $ah = h'a$ and $ha = ah''$. In other words, $aha^{-1} = h' \in H$ and $h = ah''a^{-1} \in aHa^{-1}$ for all $h \in H$, i.e. $H = aHa^{-1}$. Conversely, $aHa^{-1} = H$ means that for every $h \in H$, both $h' := aha^{-1} \in H$ and $h = ah''a^{-1}$ for some $h'' \in H$. These are respectively equivalent to $ah \in Ha$ and $ha \in aH$, showing that $aH = Ha$.

`Examples.` (1) In an abelian group, all subgroups are normal.

(2) In the symmetry group of a regular triangle, the rotations $\{\mathrm{id}, \rho, \rho^{-1}\}$ form a normal subgroup (check this!)

(3) The kernel $\ker f$ of a group homomorphism $f : G \to G'$ is a normal subgroup in G. Indeed, if $f(h) = e'$ then $f(aha^{-1}) = f(a)f(h)f(a^{-1}) = f(a)e'f(a)^{-1} = e'$ for all $a \in G$.

The last example turns out to be universal due to the following

`Theorem.` *Let $H \subset G$ be a normal subgroup, G/H denote the set of H-cosets, and $\pi : G \to G/H$ be the natural projection to the set of equivalence classes defined by the equivalence relation $\equiv_H$. There exists a unique group structure on G/H such that π is the group homomorphism. Besides, $\pi(G) = G/H$ and $\ker \pi = H$.*

`Proof.` The homomorphism requirement says: $\pi(ab) = \pi(a)\pi(b)$ for all $a, b \in G$. Reading it from right to left, we conclude that the only way to define the product of H-cosets so that the requirement holds is to pick their representatives a and b, multiply them in G, and

then take the H-coset of the product. This confirms the uniqueness claim. To establish the existence, we need to make sure that the result does not depend on the choice of the representatives. So, let $a' = ha$, $h \in H$, be another representative from Ha. Then $a'b = hab \in Hab$. Now let $b' = hb$ be another representative from Hb. Then $ab' = ahb = (aha^{-1})ab \in Hab$ — this time because H is normal implying that $aha^{-1} \in H$ whenever $h \in H$. Thus, the product operation in G/H is well-defined.

The group-theoretic properties of the operation on H-cosets hold true because they do so for the representatives. Indeed, the associativity $(\pi(a)\pi(b))\pi(c) = \pi(a)(\pi(b)\pi(c))$ holds because $(ab)c$ and $a(bc)$ are legitimate representatives of the left-hand-side and right-hand-side respectively. Taking $e' := \pi(e)$, we find that $\pi(a)e' = \pi(a) = e'\pi(a)$ since $ae = a = ea$, i.e. e' does serve as the identity element in G/H. In particular, $\ker \pi = \pi^{-1}(e') = H$ (the H-coset of e). Furthermore, $aa^{-1} = e = a^{-1}a$ implies $\pi(a)\pi(a^{-1}) = e' = \pi(a^{-1})\pi(a)$ showing that $\pi(a^{-1})$ qualifies for the inverse of $\pi(a)$. Finally $\pi(G) = G/H$ because the projection of a set to the set of the equivalence classes of an equivalence relation is always surjective.

The group G/H constructed in the theorem is called the *quotient group* (or sometimes the *factor group*) of the group G by its normal subgroup H. The finite cyclic groups $\mathbb{Z}_n := \mathbb{Z}/n\mathbb{Z}$ are the familiar examples of this construction.

`Corollary` (The Homomorphism Theorem). *Every group homomorphism $f : G \to G'$ decomposes as the projection of the domain group onto its quotient by the kernel of the homomorphism, followed by an isomorphism between the quotient group and the range of the homomorphism, followed by the inclusion of the range into the codomain group: $G \xrightarrow{\pi} G/\ker f \xrightarrow{\cong} f(G) \hookrightarrow G'$.*

`Proof.` Indeed, $f(a) = f(b)$ in the range of f if and only if $ab^{-1} \in \ker f$, i.e. whenever a and b lie in the same congruence class relative to the (normal!) subgroup $\ker f$. This establishes a bijection $i : G/\ker f \to f(G)$ such that $i(\pi(a)) = f(a)$ for all $a \in G$. It is a group homomorphism because $i(\pi(a)\pi(b)) = i(\pi(ab)) = f(ab) = f(a)f(b)$, where the first equality is due to the construction of the product in the quotient group, the second one is the defining property of i, and the last one holds because f is a group homomorphism.

The significance of this (frankly — tautological) result lies in the complete description it provides for all homomorphic images of a given group in intrinsic terms of the group itself: it suffices to examine G/H for all normal subgroups $H \subset G$. For instance, every subgroup in $\mathbb{Z}$ is (normal and), if non-zero, consist of all multiples of the smallest positive element in it (check this!) Consequently, all homomorphic images of $\mathbb{Z}$ are isomorphic to $\mathbb{Z}/d\mathbb{Z}$ for some $d \geq 0$. (In fact we have already established this in our discussion of cyclic subgroups.) Another example: the symmetry group of a regular triangle has order 6. By Lagrange's theorem, a proper subgroup in it must have order 1, 2, or 3, and hence be cyclic. Cyclic subgroups of the reflections (such as $\{\mathrm{id}, r_A\}$) are not normal, leaving the cyclic group of $\rho^{\pm 1}$ the only non-trivial normal subgroup. Thus, every homomorphism of this group to any group G' either maps the whole group to e', or is an embedding (with the kernel $\{\mathrm{id}\}$), or maps the rotation subgroup $H = \{\mathrm{id}, \rho, \rho^{-1}\}$ to e' and all reflections r_A, r_b, r_C (which form the other H-coset) to any element of order 2 in G'.

EXERCISES

27. Show that left and right translations, $L_a : g \mapsto ag$ and $R_a : g \mapsto ga$, by a on a group G are bijections.

28. Prove that the inversion operation $g \mapsto g^{-1}$ on G transforms left H-cosets to right ones, and *vice versa*.

29. Prove that if the number of left H-cosets is finite, then the number of right H-cosets is also finite and is equal to the number of left H-cosets. (This number is called the *index* of the subgroup H in the group G, and is often denoted by $[G : H]$.)

30. Symmetries of a group G form a group denoted $Aut(G)$ and called the *automorphism group* of G. By definition it consists of *automorphisms* of G, i.e. isomorphisms of G with itself, with the operation of composition of such isomorphisms. Prove that $Aut(K_4) \cong S_3$.

31. Prove that $300^{3000} - 1$ is divisible by 1001.

32. Check that the subgroups $\{e\} \subset G$ and $G \subset G$ (called *trivial subgroups*) are normal and describe the corresponding quotient groups.

33. Prove that the conjugation $g \mapsto aga^{-1}$ by $a \in G$ is an automorphism of G. (Such automorphisms are called *interior*.) Show that interior automorphisms form a subgroup in $Aut(G)$, and that it is normal.

34. Show that the subgroups $\{\mathrm{id}, r_A\}$, $\{\mathrm{id}, r_B\}$, $\{\mathrm{id}, r_C\}$ in the symmetry group of the triangle are transformed into each other by conjugations.

35. Prove that any subgroup of index 2 is normal.

36. (a) Find all normal subgroups of the dihedral group D_n. (b) Describe up to isomorphism all homomorphic images of D_n.

37. In a group G, consider the *commutator subgroup* $[G, G]$ which by definition is the smallest subgroup containing all *commutators* $aba^{-1}b^{-1}$ of elements $a, b \in G$. Prove that $[G, G]$ is normal, and that $G/[G, G]$ is abelian. Show that, moreover, every homomorphism from G to an abelian group contains $[G, G]$ in its kernel.

38. Let $K \subset G$ be a normal subgroup, and H be a subgroup in G containing K. Show that: (a) K is normal in H, (b) the natural projection $\pi : G \to G/K$ maps H to a subgroup isomorphic to H/K, and (c) H/K is normal in G/K whenever H is normal in G.

Lecture 6. Rotations of the cube

A. The order. We have introduced a number of abstract group-theoretic concepts, such as subgroups, cosets, conjugations, homomorphisms, kernels, normal subgroups, quotients, etc., and here we are going to explore one non-trivial example in a hope to get an intuitive grip on some of those abstractions.

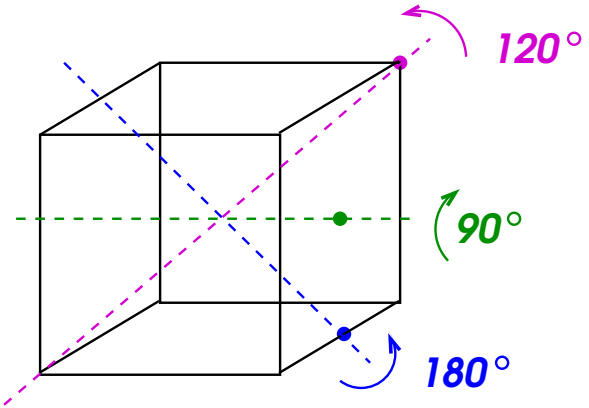


Figure 4: Three types of rotation axes

To begin with, the cube is a very symmetric solid. By merely rotating the cube one can transform any of its 8 vertices into any other vertex, any of its 6 faces into any other face, and any of its 12 edges into any other edge. With respect to the operation of composition of rotations, all rotations of the cube (together with the identity) form a group — the *rotation group of the cube* — which we will denote here by G. It consists of all those symmetries of the Euclidean 3-space which preserve the orientation (i.e. map left gloves into left ones) and, well, preserve the standard unit cube $|x|, |y|, |z| \leq 1$. In Figure 4, the three types of axes of rotation are described. The dashed line shown in green passes through the centers of a pair of opposite faces; the cube can be rotated through 90, 180, or 270 (the same as -90) degrees about such an axis. The blue dashed line passes through the midpoints of a pair of opposite edges and serves as the axis of a 180 degree rotation. The dashed red line passes through a pair of opposite vertices and serves as the axis of rotations through ± 120 degrees. Our first concern is the order $|G|$ of the group.

So far we have apprehended: $4 \times 2 = 8$ rotations through $\pm 120^\circ$ about 4 axes of the red type, 6 rotations thought 180° about the axes of the blue type, and $3 \times 3 = 9$ rotations about each of the 3 axes of the green type through $90^\circ, 180^\circ, 270^\circ$ for each. Together with the identity transformation, this gives $1 + 8 + 6 + 9 = 24$ different elements of G. Have we missed anything? Let us use the following three different ways to check.

Clearly, transformations from G preserving the vertex v marked red in Figure 4 form a subgroup H_v of order 3 — the cyclic subgroup of the 120° rotation. Let a be a transformation from G mapping v to another vertex v'. Then the elements of the left H_v-coset aH_v all map v to v'. Indeed, if $h \in H_v$, then $h(v) = v$, and $ah(v) = a(v) = v'$. Conversely, if $a'(v) = v'$, then $a' = a(a^{-1}a')$ where $a^{-1}a'$ maps v to v' and then back to v, i.e. $a^{-1}a' \in H_v$, and $a' \in aH$. Thus, the number of left H_v-cosets is equal to the number 8 of the vertices of the cube, and $|G| = 8 \times |H_v| = 8 \times 3 = 24$.

Likewise, the subgroup $H_e \subset G$ of transformations preserving the edge whose center is marked blue in Figure 4 has order 2 (it is the cyclic subgroup of the 180° rotation). Thus $|G| = \#$ edges $\times\, |H_e| = 12 \times 2 = 24$.

Yet another check: the subgroup H_f preserving the face whose center is marked green is the cyclic group of the 90° rotation. Thus $|G| = \#$ faces $\times\, |H_f| = 6 \times 4 = 24$.

By the way, we have defined H_v as the subgroup fixing the vertex v, but all vertices of the cube are alike. How is H_v related to the subgroup $H_{v'}$ fixing v'? If $a(v) = v'$ and $h(v) = v$, then the composition $aha^{-1} \in H_{v'}$ since it first maps v' to v by a^{-1}, then keeps v fixed by h, and then maps it back to v'. Thus $H_{v'} = aH_va^{-1}$ is subgroup *conjugated* to H_v, i.e. $H_{v'}$ is obtained from H_v by the *interior automorphism* $G \to G : g \mapsto aga^{-1}$, the conjugation by a. In particular, the subgroup H_v is not normal (for it does not fix all the vertices of the cube — only some of them).

B. S_4**.** In fact we know another group of order 24, the group of all permutations on the set $\{1, 2, 3, 4\}$ of four objects. How can one show that two groups which happen to have the same order are *not* isomorphic? An isomorphism is not just a bijection, but one which respects the group operations. So, if exists, it would transform subgroups to subgroups, cyclic groups to cyclic groups, elements of order m to elements of order m, conjugated elements (or subgroups) to conjugated elements (resp. subgroups), etc. So, let's compare the orders of elements in G with those in S_4.

In G, there are $4\times 2 = 8$ elements of order 3 (the $\pm 120^\circ$ rotations), $3 \times 2 = 6$ elements of order 4 ($\pm 90^\circ$ rotations), $3 + 6 = 9$ elements of order 2 (180° rotations about 3 green axes and 6 blue axes), and (as in any group) 1 element of order 1 — the identity.

A permutation $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$ can be described by the table $\begin{pmatrix} 1 & \ldots & n \\ \sigma(1) & \ldots & \sigma(n) \end{pmatrix}$ of its values (bottom line) on the inputs (top line). However, properties of a particular permutation become more transparent from the so-called *cycle decomposition* of it. Namely, σ maps 1 to $\sigma(1)$, $\sigma(1)$ to $\sigma(\sigma(1))$, and so on, until certain iteration of σ maps 1 back to 1. (It will surely do that because it cannot map two different objects to the same one.) Setting aside the objects involved in the first cycle and starting with a new object one obtains another cycle, and so on, ending with a partition of the whole set $\{1, \ldots, n\}$ into non-overlapping cycles cyclically permuted by σ within each of them. For example $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 3 & 4 & 1 & 5 & 9 & 8 & 6 & 7 & 2 \end{pmatrix}$ decomposes as $1 \mapsto 3 \mapsto 1$, $2 \mapsto 4 \mapsto 5 \mapsto 9 \mapsto 2$, $6 \mapsto 8 \mapsto 7 \mapsto 6$. This result can be recorded as the product $\sigma = (1\ 3)(2\ 4\ 5\ 9)(6\ 8\ 7)$ of cycles $(i_1\ i_2\ \ldots\ i_l)$ which are permutations mapping i_1 to i_2, i_2 to i_3, etc. with the understanding that i_l is mapped back to i_1, and all objects other than i_α stay fixed. The ordering of the cycles in the product (of the three in this example) is irrelevant since non-overlapping cycles commute. The order of σ, as it not hard to check, equals the *least common multiple* of the lengths of the cycles in the cycle decomposition of σ (i.e. 12 in this example).

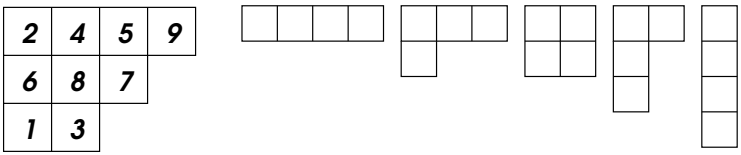


Figure 5: Young diagrams

On the left of Figure 5, a *Young tableau* representing our permutation σ is shown. The format of the tableau (called a *Young diagram*) represents the *partition* of 9 as the sum $4+3+2$ of positive integers, encoded by the numbers of the cells in each row, which are non-increasing ($4 \geq 3 \geq 2$) going from top to bottom. The tableau is obtained by filling in the cells of the diagram by integers $1, 2, \ldots, 9$. Of course, the fillings obtained from each other by cyclically shifting

the entries in a row represent the same permutation. (Also, if rows of the same length were present, reordering them would not change the permutation σ represented by the tableau.)

On the right of Figure 5, the Young diagrams of all 5 partitions of the number 4 are represented — from left to right: $4 = 3 + 1 = 2 + 2 = 2 + 1 + 1 = 1 + 1 + 1 + 1$. Thus, the orders of elements in S_4 (i.e. the least common multiples of the terms of the partitions) are respectively: 4, 3, 2, 2, and 1.

How many permutations are there of each order? The row of 4 cells can be filled in by the numbers $1, 2, 3, 4$ in $4! = 24$ different ways. However, cyclic shifts of the numbers in the row doesn't change the permutation, i.e. the number of permutations of order 4 equals $24/4 = 6$. The next diagram $3 + 1$ (the same as any other in fact) can also be turned into a tableau in 4! different ways, but up to 3 cyclic shifts in the top row this yields $24/3 = 8$ permutations of order 3. The partition $2 + 2$ contributes $24/(2 \cdot 2 \cdot 2) = 3$ to the number of permutations of order 2 (where the last factor 2 in the denominator offsets the overcounting by the 2! reorderings of the two equal rows of the diagram). The partition $2 + 1 + 1$ contributes $24/(2 \cdot 2!) = 6$ more permutations of order 2. The rightmost diagram represents the identity permutation regardless of its filling by the numbers, and does it $24/4! = 1$ times. Thus, the numbers of elements of the orders $4, 3, 2, 1$ in S_4 are the same as in G: $6, 8, 9$ and 1 respectively.

Let us try to refine this information by introducing *conjugacy classes*. Namely, when in a group $b = gag^{-1}$ for some g, let's call b *conjugated* to a ($b \sim a$). It is an equivalence relation (check this!) and thus partitions the group into conjugacy classes.

What happens with the young tableau representing permutation a when it is replaced with permutation $b = gag^{-1}$? Here g is also a permutation on $\{1, \ldots, n\}$, and all it does is *renaming* the n objects, i.e. replacing each name $i \in \{1, \ldots, n\}$ with $g(i)$. Thus, if we rename all entries in the Young tableau of a according to the renaming scheme g, we obtain the Young tableau for b — but the Young diagram, i.e. the partition of n into the lengths of cycles in the cycle decomposition remains the same! The converse is also true: two fillings of the same Young diagram define a renaming scheme g. Thus, S_4 consists of 5 conjugacy classes of sizes $6, 8, 3, 6, 1$ whose elements have the orders $4, 3, 2, 2, 1$ respectively.

What about the rotation group G? The 6 rotations through $\pm 90°$ around the green axes form one conjugacy class. Indeed, to the face marked with the green center in Figure 4, we can associate a particular $90°$ rotation about the shown green axis by applying the left-hand

rule. But the cube has 6 faces, and each one can be rotated into each other, showing that all such rotations are conjugated to each other. The 8 rotations through 120° are similarly related to the 8 vertices of the cube, and hence form a single conjugacy class. The 3 rotations through 180° about the green axes are conjugated by the rotations of the cube permuting the axes, and the same applies to the 6 rotations through 180° about the blue axes, which are conjugated by the rotations interchanging the edges of the cube. Together with the conjugacy class of the identity transformation, we have 5 conjugacy classes of sizes $6, 8, 3, 6, 1$ consisting of elements of the orders $4, 3, 2, 2, 1$ respectively.

Perhaps it is time to abandon the working hypothesis that G is not isomorphic to S_4, and conjecture the opposite.

C. $G \cong S_4$. How does one prove that two groups *are* isomorphic? The group structure is fully encoded in the function $G \times G \to G$ of multiplication, and some authors advocate for the use of "multiplications tables". In our case it will be a square grid of size 24×24 with the names of group elements labeling rows and columns of it and the product (row element) $\cdot$ (column element) written inside each cell of the grid. Comparing two such tables — one for G and one for S_4 — one should not be surprised if they won't look the same: this does not mean that the groups aren't isomorphic. To check whether they are not one "only" needs to try the tables obtained from the one for G by each of the $24! \approx 6 \cdot 10^{23}$ permutations applied via simultaneous renaming of the rows, the columns, and the entries of the grid. If none of the 24! tables thus obtained coincide with the one for S_4, then the groups are not isomorphic; if for one of the permutations the two grids agree, then — *voila!* — an isomorphism is found. The futility of this approach is also underscored by the fact that it suits for comparing *arbitrary* functions $G \times G \to G$, and thereby ignores the group-theoretic nature of the problem.

In reality, if two meaningful groups happen to be isomorphic, there usually is a meaningful reason for this. In our case, S_4 is the group of all symmetries of a set of 4 objects. To have a candidate for an isomorphism $G \to S_4$ it suffices to find 4 objects associated with the cube which are permuted by the cube's symmetries. It should not be too hard to guess now that the set of 4 diagonals qualifies. Thus, we obtain a *group homomorphism* $f : G \to S_4$ which to a rotation g of the cube associates the permutation $f(g)$ of the 4 diagonals which is induced by g. When two rotations are composed, the permutation $f(g_1 g_2)$ their composition induces is obtained by the consecutive application of the permutations $f(g_2)$ followed by $f(g_1)$,

i.e. $f(g_1)f(g_2)$ if the functional notation (where the input of functions stays on the right of the function's name) is used. This is why f is a group homomorphism. To check whether it is an isomorphism, it suffices to verify that $\ker f = \{\mathrm{id}\}$. Indeed, is that's the case, then f is 1-to-1, and hence is onto, because the sets G and S_4 have the same number of elements.

To this end, let g be a rotation of the cube which leaves each of the 4 diagonals in its place, and hence merely preserves or reverses the direction of each of them. We claim that the sequence of four signs $\pm$ thus obtained is in fact $+,+,+,+$, i.e. each vertex of the cube remains fixed, and thus $g = \mathrm{id}$. Indeed, any three of the diagonals can be taken for (non-Cartesian) coordinate axes, and in this coordinate system the linear transformation defined by g will have a diagonal matrix with the corresponding three signs ± 1 on the diagonal. But their product is the determinant of g, i.e. $+1$, since rotations have positive determinants (in fact always $= 1$), the same as the identity transformation. This implies that each of the four signs equals the product of all four, and thus must be $+$, for otherwise the product of each three would be negative.

D. Normal subgroups in S_4. Inspired by the above success, let's try to construct other homomorphisms from G and from S_4 in order to find their normal subgroups.

Besides acting on the set of 4 pairs of opposite vertices (diagonals) of the cube, the group G also acts on the 6 pairs of opposite edges, and on the 3 pairs of opposite faces. The former action defines a homomorphism $G \to S_6$ which in all likelihood is injective (since $|S_6| = 6! = 720 > 24$) and has trivial kernel. However, the latter defines a non-trivial homomorphism $h : G \to S_3$. Its kernel consists of all those rotations which preserve all the three pairs of opposite faces. Besides the identity map, only the three 180° rotations about the green axes qualify. Thus, $\ker h$ has order 4. By the homomorphism theorem, h identifies the quotient group $G/\ker h$ of order $24/4 = 6$ with the entire S_3 (since $|S_3| = 3! = 6$ as well). We conclude that there should exist an epimorphism $\tilde{h} : S_4 \to S_3$ with the kernel of order 4 (and such that $\tilde{h} \circ f = h$). What is it?

Note that the 180° rotation about the green axis in Figure 4 permutes the four diagonals of the cube (which we can label by $1, 2, 3, 4$ following the vertices around the face marked by the green center) as the product of two 2-cycles $(13)(24)$. In the left Figure 6, this permutation corresponds to the central symmetry of the rectangle. The 180° rotations about the lines of centers of the other two pairs of

faces of the cube permute the diagonals as (12)(34) and (14)(23), i.e. as the symmetries of the rectangle about the dashed midlines. Together with the identity permutation, these three form the subgroup in S_4 preserving three partitions of the vertices of the rectangle into two pairs: connected by the long edges, by the short edges, and by the diagonals. The group S_4 of all permutations of the vertices of the rectangle permutes the three partitions of $\{1,2,3,4\}$ into pairs, which therefore defines the homomorphism $S_4 \to S_3$. The kernel of it is the Klein group K_4 consisting of the permutations induced by the symmetries of the rectangle.

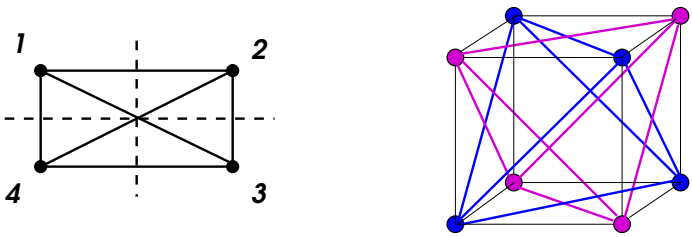


Figure 6: Toward normal subgroups in S_4.

Reversing the logic of this argument, note that we know of the epimorphism $\epsilon : S_4 \to S_2 = \{\pm 1\}$ defined by the parity of permutations. Its kernel, the alternating group A_4, consists of 12 even permutations. We obtain therefore the epimorphism $\epsilon \circ f : G \to S_2$. What is the kernel of it as a subgroup in the rotation group G? In the right Figure 6, two regular tetrahedra inscribed into the cube are shown: one blue and one red, whose edges are the diagonals in the faces of the cube, and whose vertices split those of the cube, each picking one from every pair of the opposite ones. Rotations of the cube act on the set of the two tetrahedra, thereby defining an epimorphism $G \to S_2$. The rotations preserving each tetrahedron form a normal subgroup in G, consisting of 12 rotations of each tetrahedron (and inducing even permutations of each tetrahedron's 4 vertices).

In the next lecture, we will prove that with one notable exception each symmetric group S_n contains only one non-trivial normal subgroup — the alternating group A_n, the exception being the Klein subgroup $K_4 \subset S_4$.

EXERCISES

39. Use linear algebra to prove that every orientation-preserving orthogonal transformation of the Euclidean 3-space (i.e. every element of the group we denoted SO_3 in Lecture 3) is a rotation through some angle about some axis passing through the origin.

40. Show that none of cyclic subgroups in G (the rotation group of the cube) is normal.

41. Prove that the order of a permutation on a finite set is equal to the least common multiple of the lengths of the cycles in the cycle decomposition of the permutation.

42. Check that conjugacy is an equivalence relation.

43. Prove that orthogonal transformations of the Euclidean 3 space have determinant ± 1, and that rotations have the same determinant $+1$ as the identity transformation.

44. Check that the action of G on the set of 6 pairs of opposite edges of the cube defines an embedding of G into S_6.

45. Show that the group of Euclidean symmetries of a regular tetrahedron in 3-space can be identified with the group S_4 of arbitrary permutations of the tetrahedron's vertices, and that even permutations correspond to orientation-preserving transformations.

46. Define a homomorphism $S_4 \to S_3$ by associating to a symmetry of a regular tetrahedron the permutation it induces on the set of 3 pairs of opposite edges of the tetrahedron, and show that the kernel of this homomorphism is the Klein subgroup.

47. Identify all non-cyclic subgroups in S_4, and find out whether different non-cyclic subgroups of the same order are (a) isomorphic and (b) conjugated to each other.

Lecture 7. Symmetric and alternating groups

A. Conjugacy classes in S_n. Since a normal subgroup is one invariant under conjugations, it must be a disjoint union of whole conjugacy classes of the group. So, let us collect here what we know about conjugacy classes in S_n.

Each conjugacy class corresponds to a partition $n = k_1 + \cdots + k_r$ of n into positive summands, which we may assume non-decreasing, and visualize the partition by a Young diagram with r rows of lengths $k_1 \geq \cdots \geq k_r > 0$. Permutations from a given conjugacy class are obtained by turning the Young diagram into a Young tableau by filling somehow the n cells of the diagram with the numbers $1, \ldots, n$: the corresponding permutation shifts cyclically the elements in each row of the tableau. We leave it to the reader as an exercise to show that the number of elements in the conjugacy class represented by a partition with l_1 ones, l_2 twos, l_3 threes, etc. is equal to

$$\frac{n!}{\prod_s s^{l_s} l_s!}, \quad \text{where } n = l_1 + 2l_2 + 3l_3 + \cdots$$

For example, the conjugacy class of a cycle $(1\ 2\ \ldots\ k)$ corresponds to the partition $n = k+1+\cdots+1$ ($n-k$ ones) and contains $n!/k\ (n-k)!$ permutations — $n!$ ways of filling the tableau (which has the shape of the letter Γ) divided by k cyclic shifts of the entries in the top row and by $(n-k)!$ permutations of the remaining entries written in the "leg" of the diagram.

Note that conjugated permutations have the same parity:

$$\epsilon(\lambda\sigma\lambda^{-1}) = \epsilon(\lambda)\epsilon(\sigma)\epsilon(\lambda)^{-1} = \epsilon(\sigma).$$

In other words, the parity of a permutation of n objects does not depend on how the objects are called, but is an intrinsic property of the permutation's structure. How does the parity of σ express in terms of the partition of n, i.e. in terms of the cycle structure of σ? The transposition (12) is odd. Indeed, the sign of the polynomial $\Delta_n = \prod_{i<j}(x_j - x_i)$ is reversed by interchanging x_1 with x_2 because this turns the factor $x_2 - x_1$ into $x_1 - x_2$ and preserves the rest. Therefore all transpositions $(i\ j)$ are odd. Furthermore, a cycle $(1\ 2\ \ldots\ k)$ can be written as the composition $(1\ k)\ldots(1\ 3)(1\ 2)$ of $k-1$ transpositions and is therefore even whenever k is odd. We conclude that *a permutation σ consisting of cycles of lengths $k_1, \ldots, k_r$ is even if and only if the number of its cycles of even lengths is even.*

`Proposition.` *Every even permutation can be represented as a product of* 3*-cycles* $(i\ j\ k)$.

`Proof.` Note that a 3-cycle $(i\ j\ k)$ (it is assumed that i, j, k are distinct) is even, and its inverse is a 3-cycle too. We prove the proposition by induction on the number of objects fixed by a given even permutation σ. Simply speaking, we will compose σ with a 3-cycle such that the number of objects fixed by the permutation increases. When it becomes n, the permutation turns into id, and hence σ becomes the product of the 3-cycle's inverses.

Suppose the cycle decomposition of σ contains a cycle of length $k > 2$, e.g. $(1\ 2\ \ldots\ k)$. Then $(2\ 1\ 3)(1\ 2\ \ldots\ k) = (1)(2)(3\ 4\ \ldots\ k)$ acquires two more fixed points (1 and 2) when $k > 3$, and three more when $k = 3$. Thus, the problem is reduced to the case when all non-trivial cycles of σ have lengths 2. Since σ is even, the number of such cycles must be even, and they can be split into (non-overlapping) pairs, such as e.g. $(1\ 2)(3\ 4)$. But $(2\ 1\ 3)(1\ 2)(3\ 4) = (1)(2\ 3\ 4)$ which also has an extra fixed point. This finishes the proof.

`Corollary.` *The alternating group* A_n *is generated by* 3*-cycles, and for* $n \geq 5$ *by the conjugacy class of* $(1\ 2)(3\ 4)$.

`Proof.` The first statement is a rephrasing of the proposition, and the second follows from $(i\ j\ k) = (i\ j)(j\ k) = (i\ j)(l\ m)\ (l\ m)(j\ k)$.

B. Normal subgroups in S_n. Recall that any group G has two *trivial subgroups*, G and $\{e\}$, which are both normal. We know that A_n is a non-trivial normal subgroup in S_n for $n \geq 3$. It is the only normal subgroup in S_3 (where all non-trivial subgroups must be cyclic and all have been looked at under the disguise of the symmetries of a regular triangle) while in S_4 there is also the Klein normal subgroup K_4. From the proof of the following theorem it will be also clear that S_4 has no other normal subgroups.

`Theorem.` *For* $n \geq 5$*, the only non-trivial normal subgroup in* S_n *is the alternating group* A_n.

`Proof.` We will show that for $n \geq 5$, a non-trivial normal subgroup $H \subset S_n$ contains A_n; since the order $|A_n| = |S_n|/2$ is the largest proper divisor of $|S_n|$, the equality $H = A_n$ would follow.

Pick $\sigma \in H$ other than id. It cannot be a transposition, because then H contains all transpositions, and they generate the whole of S_n. If it is a 3-cycle, then H contains all 3-cycles, and hence contains A_n. Otherwise either σ is a cycle of length ≥ 4 or it contains at least two distinct cycles of lengths ≥ 2. In either case there exist $i \neq j$

such that $\sigma(i)$ and $\sigma(j)$ are distinct from both i and j (and of course from each other). Then $\sigma^{-1}(i)$ and $\sigma^{-1}(j)$ are also distinct from both i and j.

Consider now the expression $\sigma\tau\sigma^{-1}\tau^{-1}$ (known as the *commutator* of σ and τ and lying in H for any τ since H is normal), where we take $\tau = (i\ j) = \tau^{-1}$. We claim that the commutator is the product $(\sigma(i)\ \sigma(j))(i\ j)$ of two disjoint 2-cycles.

Indeed, when $k \neq i, j, \sigma(i), \sigma(j)$, we have $\tau^{-1}(k) = k$, $\sigma^{-1}(k) \neq i, j$ (for otherwise $k = \sigma(i)$ or $\sigma(j)$), hence $\tau(\sigma^{-1}(k)) = \sigma^{-1}(k)$, which is mapped by σ back to k. Thus, the commutator fixes all objects except $i, j, \sigma(i), \sigma(j)$, which are indeed mapped by the commutator as described. Namely,

$$i \overset{(i\ j)}{\longmapsto} j \overset{\sigma^{-1}}{\longmapsto} \sigma^{-1}(j) \overset{(i\ j)}{\longmapsto} \sigma^{-1}(j) \overset{\sigma}{\longmapsto} j,$$

$$\sigma(i) \overset{(i\ j)}{\longmapsto} \sigma(i) \overset{\sigma^{-1}}{\longmapsto} i \overset{(i\ j)}{\longmapsto} j \overset{\sigma}{\longmapsto} \sigma(j),$$

and similarly for j and $\sigma(j)$.

Thus, H contains a whole conjugacy class which according to the above corollary generates the entire A_n when $n \geq 5$.

`Remark.` Note that for $n = 4$ we can only conclude that together with this commutator, H contains K_4. Therefore H projects to a proper normal subgroup H/K_4 in the quotient group $S_4/K_4 = S_3$. When $H/K_4 = A_3$, we have $H = A_4$. When $H/K_4 = \{\mathrm{id}\}$ we have $H = K_4$. Thus, the only non-trivial normal subgroups in S_4 are A_4 and K_4.

C. The simplicity theorem. A group is called *simple* if it has no non-trivial normal subgroups (and hence all homomorphisms of it are either constant or injective). The program of classification up to isomorphism of all finite simple groups was completed by 2004 thanks to heroic several-decades-long efforts of many mathematicians, and is considered the crowning achievement of the theory of groups. The answer is rather complicated, as it involves several infinite series (of simple groups of the orders growing indefinitely within each series) plus 27 "sporadic groups" of which the largest one has the order comparable, people say, to the number of particles in the Universe. On the other end: it turns out that the smallest non-abelian simple group (for simple abelian groups, see Exercises) is the alternating group A_5, which has order 60. It begins one of the infinite series of finite simple groups.

`Theorem.` *The alternating groups A_n are simple for $n \geq 5$.*

`Remark.` This result does not follow directly from our description of normal subgroups in S_n because a normal subgroup $H \subset A_n$ is not *a priori* guaranteed to remain normal in S_n. Most of the usual proofs of this theorem consist in showing that H contains 3-cycles, and hence all of them, since for $n \geq 5$ all 3-cycles turn out to be conjugated not only in S_n but also in A_n. Instead, we show below that for all n a normal subgroup in A_n must remain normal in S_n.

`Proof.` Let us argue *ad absurdum*: assume that H is a normal subgroup in A_n which is not normal in S_n, and arrive at a contradiction. Thus, $\lambda H \lambda^{-1} = H$ for all even permutations λ, while for some odd permutation λ_0, $H' := \lambda_0 H \lambda_0^{-1} \neq H$ is another normal subgroup in A_n. Then in fact, since compositions of odd permutations are even, the same subgroup H' results from conjugating H by every odd λ, and moreover, conjugating H' by an odd λ returns H. The intersection $H \cap H'$ must be invariant under conjugations by all elements of S_n (the odd ones interchange H and H' and the even ones preserve each), and is therefore a normal subgroup in S_n. But there is no such subgroup in S_n — unless $H \cap H' = \{\mathrm{id}\}$ (since even for $n = 4$, K_4, which has prime index 3 in A_4, is contained in only one larger subgroup, A_4). This verdict already looks suspicious, because this means that for every $\sigma \in H$, its conjugacy class in S_n must be split evenly between H and H' into two conjugacy classes of A_n — one of σ (in H) and one of $\lambda_0 \sigma \lambda_0^{-1}$ (in H'). So, our current goal is to figure out how conjugacy classes of S_n decompose into conjugacy classes of A_n.

Consider two permutations $\sigma, \sigma' \in A_n$ with the same cycle structure represented by a partition $n = k_1 + \cdots + k_r$. They can be described by two fillings (Young tableaux) of the same Young diagram. The comparison of the tableaux reveals a renaming scheme g of the objects $1, \ldots, n$, which conjugates σ into σ'. If g is itself an even permutation then $\sigma' = g\sigma g^{-1}$ in A_n. Suppose then that g is odd. Take, however, one of the rows in the Young diagram and cyclically shift the entries of one of the two tableaux (say, the one for σ'). This does not change the permutation σ' but changes the renaming scheme g by composing it with the cyclic shift. If the length of the row is even, then the cyclic shift is odd, and the new renaming scheme conjugating σ into σ' becomes even. So, assume that the Young diagram has no rows of even length, but has two rows of equal odd length. Swapping the rows of the Young tableau for σ' (e.g. replacing (1 2 3)(4 5 6) with (4 5 6)(1 2 3)) we find a new renaming scheme obtained from g by composing it with the product ((1 4)(2 5)(3 6) in our example) of an odd number of transpositions. This also changes the parity of the

renaming scheme without changing the permutations σ and σ'. We conclude that *a conjugacy class in S_n of even permutations remains a single conjugacy class in A_n unless the terms k_α of the corresponding partition $n = k_1 + \cdots + k_r$ are all odd and distinct.* The converse is also true, but we won't use it and leave it as an exercise to the reader.

Thus, every non-identity permutation σ in the normal subgroup $H \subset A_n$ which is not normal in S_n must have its cycle decomposition consisting of cycles of odd distinct lengths. But this is already *very* suspicious. For, if σ has two such cycles of length $k_\alpha > k_\beta > 1$, then $\sigma^{k_\beta} \neq \mathrm{id}$ also lies in H but has at least $k_\beta > 1$ cycles of length 1 — a contradiction! Thus, σ must have only one cycle of odd length $k > 1$. Moreover, k must be prime, for if it factors non-trivially as ml, then $\sigma^m \in H$ has m cycles of length l — a contradiction! The only remaining possibilities therefore are: (i) $n = p$ an odd prime, and σ has one cycle of length p, or (ii) $n = p+1$, and σ has one cycle of length p and one of length 1.

In case (i), the conjugacy class of σ in S_n consists of $p!/p = (p-1)!$ elements which are split evenly between H and H', and so the subgroup H has order $1+(p-1)!/2$ (where 1 stands for the identity element). In case (ii), the conjugacy class consists of $(p+1)!/p = (p+1)\ (p-1)!$ elements, resulting in $|H| = 1 + (p+1)\ (p-1)!/2$. In either case, $|H|$ must be divisible by the order p of σ, which leads to a contradiction. Namely, according to *Wilson's theorem*, $(p-1)! \equiv -1 \mod p$ implying $2|H| \equiv 1 \neq 0 \mod p$.

D. Wilson's theorem. This is a famous fact of elementary number theory:

`Theorem.` *For any odd prime p, $(p-1)! \equiv -1 \mod p$.*

`Proof.` In any finite abelian group G, the product $\prod_{x\in G} x$ of all group elements is equal to the product $\prod_{x:\ x^2=\mathrm{id}} x$ of elements equal to their inverses. This is because all other element come in pairs $x \neq x^{-1}$ which cancel in the product. Applying this to the multiplicative group $\mathbb{Z}_p^\times$, we find $(p-1)! \equiv (+1)\cdot(-1) \equiv -1 \mod p$. It is important here that p is prime and consequently all non-zero congruence classes in $\mathbb{Z}_p$ lie in the group $\mathbb{Z}_p^\times$. Therefore $x^2 - 1 = (x-1)(x+1) = 0$ in $\mathbb{Z}_p$ only when $x = 1$ or $x = -1$.

`Remark.` Note that modulo 8, all the four elements $\pm 1, \pm 3$ of $\mathbb{Z}_8^\times$ are roots of $x^2 - 1$. For example, $(3-1)(3+1) = 2 \cdot 4 \equiv 0 \mod 8$. Here the product of non-zero classes $\bar{2}$ and $\bar{4}$ is zero, making them *zero divisors*, an event impossible modulo a prime.

EXERCISES

48. Find the smallest n for which the symmetric group S_n contains a cyclic subgroup of order 60.

49. Show that the number of elements in the conjugacy class of S_n represented by a partition with l_1 ones, l_2 twos, l_3 threes, etc. ($n = 1l_1 + 2l_2 + 3l_3 + \cdots$) is equal to $n!/\prod_s s^{l_s} l_s!$.

50. Prove that S_n is (a) generated by transpositions, (b) generated by $n-1$ transpositions $(1\ 2), (2\ 3), \ldots, (n-1\ n)$.

51. Prove that a non-trivial simple finite abelian group is isomorphic to one of $\mathbb{Z}_p$ with prime p.

52. Show that 3-cycles do form a single conjugacy class in A_n for $n > 4$, but do not for $n = 3$ and 4.

53. Given a permutation σ with l_1 cycles of length 1, l_2 cycles of length 2, etc., describe permutations commuting with σ, and show that their total number is equal to $1^{l_1} l_1! 2^{l_2} l_2! 3^{l_3} l_3! \cdots$

54. Prove that a conjugacy class in S_n of an even permutation remains a single conjugacy class in A_n if and only if there exists an odd permutation commuting with it.

55. Prove that a conjugacy class of even permutations whose cycle lengths k_α ($n = k_1 + \cdots + k_r$) are all odd and distinct splits into two equal size conjugacy classes of A_n.

56. Provide another ending to the proof of the simplicity theorem by exploiting the fact that $|H|$ must divide the order of the alternating group, $p!/2$ in case (i) and $(p+1)!/2$ in case (ii).

57. Prove that for $n > 1$, the commutator subgroup $[S_n, S_n]$ coincides with A_n, and for $n > 4$, $[A_n, A_n] = A_n$. Identify the commutator subgroups in A_3 and A_4.

Lecture 8. Group actions

A. Two definitions. Our informal introduction to groups was focused on subgroups $G \subset S(X)$ of the group of all permutations of a set (and preserving a certain structure on it). Here is how that situation generalizes to the abstract setting.

One says that a group G *acts* on a set X if a homomorphism $G \to S(X)$ (not necessarily injective) of G to the group of all permutations on the set X is given.

Equivalently, an action of G on X is defined by a map $G \times X \to X$ (whose value on a pair $(g, x) \in G \times X$ is denoted $gx \in X$) such that $ex = x$ for all $x \in X$, and for all $g_1, g_2 \in G$ and $x \in X$ the following "associativity property" holds: $(g_1 g_2)x = g_1(g_2 x)$.

The last condition means that the maps $g_i : X \to X$ defined by $x \mapsto g_i x$, when composed in the order "g_2 first, g_1 last" yield the map $x \mapsto (g_1 g_2)x$ corresponding to the product $g_1 g_2$ in G. In particular, since the map $x \mapsto ex$ defined by the unit element is $\mathrm{id} \in S(X)$, the maps $X \to X$ defined by g and g^{-1} turn out to be inverse to each other, implying that each map $x \mapsto gx$ is a permutation on X. Thus, the action in the second sense indeed defines a homomorphism $G \to S(X)$. Conversely, given such a homomorphism, π, the action map $(g, x) \to gx$ is defined by applying the permutation $\pi(g)$ to x. This shows that the two definitions are equivalent indeed.

Not unlike Moliére's character who didn't know he was always speaking *prose*, we have been actually using group actions all the time without calling them so. Rotations of the cube act on the sets of the cube's vertices, faces, edges, rotation axes (altogether or of each color separately). Any group acts on itself by conjugations, left translations, right translations.

In fact there is a subtlety in the last statement. Strictly speaking the notion of action we have just introduced is that of a *left action*. A *right action* of G on X is defined by a function $G \times X \to X : (g, x) \mapsto xg$ satisfying $x(g_1 g_2) = (xg_1)g_2$ for all $g_1, g_2 \in G$ and $x \in X$. The difference from the left action is not just that group elements are written on the right set elements, but that the maps defined by g_1 and g_2 are composed in the opposite order: g_1 first, g_2 last. In particular, left translations $L_g : x \mapsto gx$ on a group G form a left action of G on itself: $L_{g_1 g_2}(x) = L_{g_1}(L_{g_2}(x))$, while right translations $R_g : x \mapsto xg$ form a right action: $R_{g_1 g_2}(x) = R_{g_2}(R_{g_1}(x))$. Respectively, R is not a homomorphism from G to $S(G)$ but an *anti-homomorphism*: it changes the order of multiplication. So, which actions are better, right or left? Here is a way out of this rather perplexing situation.

With each group G one can associate the *opposite group* G^{op} which coincides with G as a set, but is equipped with the new product $\circ$: $g_1 \circ g_2 := g_2 g_1$. The identity map from G to G^{op} is an anti-isomorphism, but the inversion map $g \mapsto g^{-1}$ defines an isomorphism between G and G^{op}: $(g_1 g_2)^{-1} = g_2^{-1} g_1^{-1} = g_1^{-1} \circ g_2^{-1}$. Consequently, given a right action of G on a set X, one turns it into a left action by composing the homomorphism $G^{op} \to S(X)$ (defining the right action of G) with the inversion isomorphism $G \xrightarrow{\cong} G^{op}$, i.e. simply speaking, by making g act as g^{-1} was supposed to: $(g, x) \mapsto x g^{-1}$. Conversely, the composition $G^{op} \xrightarrow{\cong} G \to S(X)$ turns a left action of G on X into the corresponding right action. The rule of thumb here is that one should stick once and for all to only one type of actions — either only left, or only right — and turn all "wrong" type of actions, should they occur, into the correct type by the inversion construction. In what follows we will stick to left actions.

`Cayley's theorem.` *Every group is isomorphic to a subgroup of a permutation group.*

`Proof.` The homomorphism of G to the group $S(G)$ of permutations on itself defined by left translations: $G \ni g \mapsto L_g \in S(G)$ (where $L_g(x) := gx$) is injective, since $g \neq e$ defines a non-identity permutation ($L_g(e) = g \neq e$) and hence doesn't lie in $\ker L$.

This tautological result has no practical significance, but it shows that every group operation can be interpreted as the composition of mappings.

B. Orbits. An action of a group G on a set X partitions the set into equivalence class called *orbits* of the action: two elements lie in the same orbit if they can be transformed into each other by the action of the group elements. More formally: $x \equiv y$ if there exists $g \in G$ such that $y = gx$. It is immediate to check that this is an equivalence relation, and that the orbit of x is Gx: the image of $G \times \{x\} \subset G \times X$ under the map $G \times X \to X$ defining the action. For example, the action of the rotation group G of the cube on the set of all its axes of rotation has three orbits: the 4 red axes, 3 green axes, and 6 blue axes (Figure 4). Sometimes the whole set X is one orbit, in which case the action is called *transitive*. For instance, the rotation group G of the cube acts transitively on the set of its 8 vertices (as well as on the sets of its 6 faces, and of its 12 edges). Obviously each orbit is a smallest subset in X invariant under the action (that is, no proper non-empty subset of it is invariant), and the action, when restricted to the orbit, becomes transitive. Our current goal is to understand the connection

between the structure of a group and that of its transitive action, or equivalently — its action on a particular orbit. The earlier analysis of the way the rotation group of the cube acts on the set of 8 vertices should serve us as an illuminating example.

Let $x_0 \in X$ be any point. The elements of G preserving this point (i.e. G-symmetries of the additional structure on X introduced by picking this point) form a subgroup $H := \{g \in G \mid gx_0 = x_0\}$ called the *stabilizer* of x_0 under the given action, and often denoted G_{x_0}. Let x_1 be another point in the same orbit, i.e. $x_1 = g_1 x_0$ for some $g_1 \in G$. Then the whole left H-coset of g_1 maps x_0 to the same x_1: $(g_1 h)x_0 = g_1(hx_0) = g_1 x_0 = x_1$ whenever $h \in H$. Conversely, when $g_1' x_0 = x_1$, we have $g_1' = g_1 h$, where the composition $h = g_1^{-1} g_1'$ maps x_0 to x_1 and then back to x_0, i.e. lies in H. We conclude that elements of the orbit Gx_0 are in bijection with the set of left H-cosets.

We will denote now the set of left H-cosets by G/H even when H is not normal. Note that left translations by G on itself induce a left action of G on G/H. Indeed, when $g_2 = gg_1$, L_g transforms $g_1 H$ to $gg_1 H = g_2 H$. (With right translations $R_{g^{-1}}$, the same would be true for right H-cosets.) We claim that under the above bijection between G/H and the orbit Gx_0, the action of G on the orbit is identified with the action by left translations on G/H. Indeed, the coset $g_2 H$ obtained by L_g from the coset $g_1 H$ corresponds to $x_2 := g_2 x_0 = (gg_1)x_0 = g(g_1 x_0) = gx_1$ which is therefore obtained by the action of g on $x_1 = g_1 x_0$ corresponding to the coset $g_1 H$.

The fact that a map $f : Y \to X$ between two sets carrying G-actions respects the two actions, i.e. that $f(gy) = g(f(y))$ for all $y \in Y$ and all $g \in G$, is expressed by saying that "f commutes with the actions", or that f is *G-equivariant*. We have established the following analogue of the homomorphism theorem; it describes transitive actions in intrinsic terms of the group.

`The orbit-stabilizer theorem.` *The map $g \mapsto gx_0$ from a group G to the orbit Gx_0 of $x_0 \in X$ under a left G-action factors through the projection $\pi : G \to G/G_{x_0}$ to the set of left cosets of the stabilizer subgroup of x_0 and defines an equivariant bijection between the set of left cosets (with the action induced by left translations on the group) and the orbit: $G \xrightarrow{\pi} G/G_{x_0} \xrightarrow{\cong} Gx_0 \subset X$.*

In particular, when the group is finite, $|Gx_0| = |G|/|G_{x_0}|$, and hence *the number of elements in any orbit of a finite group divides the order of the group.*

What if the element we first pick in the orbit is not x_0 but $x_1 = g_1 x_0$? The stabilizer $G_{x_1} = \{g \in G \mid gx_1 = x_1\}$ is in fact a subgroup conjugated to G_{x_0} by means of g_1: $G_{x_1} = g_1 G_{x_0} g_1^{-1}$.

Indeed, mapping x_1 to x_0 by g_1^{-1}, then applying any transformation fixing x_0, and then mapping it back to x_1 by g_1, we obtain a transformation fixing x_1, i.e. $g_1 G_{x_0} g_1^{-1} \subset G_{x_1}$. By the same token, since $x_0 = g_1^{-1} x_1$, we find that $g_1^{-1} G_{x_1} g_1 \subset G_{x_0}$ and hence $G_{x_1} \subset g_1 G_{x_0} g_1^{-1}$.

C. Cauchy's counting principle. It is an orbit-counting formula, which is also known as Cauchy–Frobenius' formula, Burnside's counting theorem, Burnside's lemma, or "the lemma which is not Burnside's".

`Theorem.` *The number of orbits of a finite group action on a finite set is equal to the average number of fixed points of the group's elements:*

$$|X/G| = \frac{1}{|G|} \sum_{g \in G} |X^g|,$$

where $X^g = \{x \in X \mid gx = x\}$ denotes the set of points in X fixed by a given $g \in G$, and X/G denotes the set of G-orbits in X.

`Example.` The symmetric group S_n acts on the set of n objects with one orbit, and the formula suggests that the expected number of fixed points of a random permutation equals 1. This sounds right, since each of the n objects is fixed by $(n-1)!$ permutations, i.e. with the frequency $(n-1)!/n! = 1/n$.

`Proof.` The numerator $\sum_g |X^g|$ represents the total number of pairs $(g, x) \in G \times X$ such that $gx = x$. It is additive with respect to X: if the set is a disjoint union of X_1 and X_2, then $|X^g| = |X_1^g| + |X_2^g|$. Thus, it suffices to prove the formula separately for each orbit. So, let's assume that G acts on X transitively. Then, according to the orbit-stabilizer theorem, $|X| = |G/G_x| = |G|/|G_x|$ for each $x \in X$. The condition $gx = x$ means that $g \in G_x$. Therefore the number of pairs (g, x) such that $gx = x$ is equal to $|X| \cdot |G_x| = |G|$. Divided by $|G|$, it becomes equal indeed to the number 1 of the orbits in the transitive case.

As a typical application of Cauchy's counting principle, let's solve the following

`Problem.` A necklace is formed by 17 identically shaped symmetric beads of black or white color strung on a circular string. How many different necklaces can be formed this way?

`Solution.` Denote by X the set of all $2^{17} = 131\,072$ different colorings of 17 beads into 2 colors. Two colorings represent indistinguishable necklaces if one can be obtained from the other by cyclic shifts or by combining such shifts with flipping the necklace over. The group of these operations is therefore the *dihedral group* D_{17} of symmetries of a regular 17-gon. It consists of 17 rotations through the angles $2\pi k/17$, $k = 0, \ldots, 16$, and 17 reflections about the symmetry lines of the 17-gon (Figure 7). The group D_{17} acts on the set X of all colorings, the indistinguishable colorings form one orbit, and our problem consists in counting such orbits.

One can try to enumerate the orbits directly, but the orbits don't have the same size, and the process would seem rather unsystematic and tiresome if not outright hopeless. Here is when Cauchy's counting principle comes handy, for we only need to count the number of g-symmetric colorings for each of the 34 group elements $g \in D_{17}$.

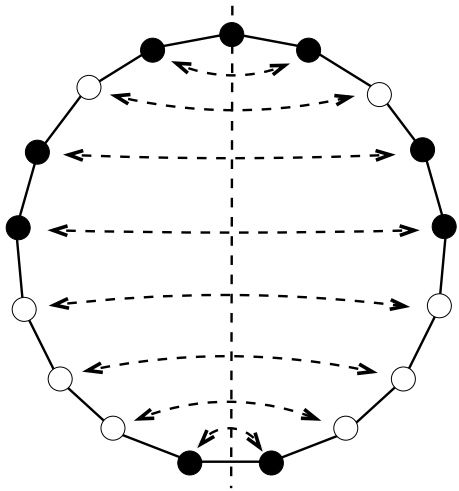


Figure 7: A reflection of the 17-gon.

Of course, all 2^{17} colorings are symmetric with respect to the identity transformation. Since 17 is a prime number, a coloring cannot be invariant under a non-trivial rotation of the 17-gon unless it is constant: all beads black or all beads white. For a coloring to

be invariant under one of the 17 reflections (Figure 7), the beads in each of the 8 symmetric pairs must have the same color (while the remaining 17-th bead can have any color), i.e. number of fixed points of each reflection is 2^9. According to the orbit-counting formula,

$$|X/D_{17}| = \frac{2^{17} + 16 \cdot 2 + 17 \cdot 2^9}{34} = \frac{(2^{17} - 2) + 17 \cdot (2 + 2^9)}{34}$$
$$= \frac{2^{16} - 1}{17} + 1 + 2^8.$$

By the way, $2^{16} - 1 = (2^8 + 1)(2^4 + 1)(2^4 - 1) = 257 \cdot 17 \cdot 15$ is indeed divisible by 17 (which also follows from Fermat's Little Theorem). Thus the number of different necklaces $|X/D_{17}| = 257 \cdot 16 = 4\,112$.

EXERCISES

58. Give an example showing that the requirement $ex = x$ for all $x \in X$ in the second definition of group actions does not follow from the "associativity" axiom.

59. Show that right actions of G on X are represented by homomorphisms $G^{op} \to S(X)$.

60. Show that two left actions of G on itself: one defined by left translations L_g, the other by right translations $R_{g^{-1}}$, commute.

61. By definition, the *center* $Z(G)$ of a group G consists of all elements commuting with all elements of the group. Show that $Z(G)$ is an abelian normal subgroup. Give an example of a group with $Z(G) = \{e\}$. What does $Z(G) = G$ mean for G?

62. Let $g \in G$ be a non-central element. Show that its *centralizer* $Z(g) := \{x \in G \mid xg = gx\}$ is a proper subgroup in G which contains $Z(G)$ as a proper subgroup (i.e. $G \supsetneq Z(g) \supsetneq Z(G)$).

63. For a finite group G, prove the *class formula*: $|G| = |Z(G)| + \sum_\alpha |C_\alpha|$, where C_α are distinct conjugacy classes of non-central elements.

64. Prove that a group of order p^2, where p is prime, is abelian. `Hint:` use the class formula to show that the order of the center is divisible by p.

65. Show that the number of conjugacy classes in a finite group is equal to the average order of the centralizers of the group's elements.

66. A 3-coloring of the cube is obtained by painting each of its faces into one of the three colors R, G, B. Two 3-colorings are called equivalent if they can be matched by rotating the cubes. Find the number of equivalence classes of such 3-colorings. (Try both approaches, direct and via Cauchy's principle.)

Lecture 9. Sylow's theorems

Let p be a prime integer. By a *p-subgroup* of a group G (or simply *p-group*) one means (sub)group of order p^r. Due to Lagrange's theorem, in a finite group of order $|G| = p_1^{r_1} \cdots p^{r_k}$ where p_i are distinct primes, non-trivial p-subgroups can exists only if p is one of p_i. A *Sylow p-subgroup* of a finite group G is defined as a subgroup of maximal order p^r dividing $|G|$, i.e. $r = r_i$ when $p = p_i$. The parts of the following (quite non-trivial) result are known as *Sylow's 1st, 2nd, and 3rd theorems*.

`Theorem.` *(1) A finite group G contains p-subgroups of every order p^r dividing $|G|$. In particular Sylow p-subgroups of G exist.*

(2) Every p-subgroup of G is conjugated to a subgroup in a given Sylow p-subgroup. In particular, all Sylow p-subgroups are conjugated in G, and are therefore isomorphic to each other.

(3) The number of Sylow p-subgroups in G is a divisor of $|G|$ which is congruent to 1 modulo p.

`Example.` The group G of rotations of the cube has order $|G| = 3 \cdot 2^3$. It contains $4 (\equiv 1 \mod 3)$ conjugated cyclic Sylow 3-subgroups generated by rotations about one of the cube's 4 diagonals. It must contain 3 conjugated Sylow 2-subgroups of order 8, since their number is an odd divisor of 24, and isn't 1 (because $G \cong S_4$ doesn't have normal subgroups of order 8). To identify these Sylow 2-subgroups, think of the cube as a square prism (e.g. by squeezing the cube between one of its 3 pairs of opposite faces). Rotational symmetries of a square prism are the same as those of a square in space (which can be rotated in its plane around its center, or flipped — through rotation in space — around one of its 4 symmetry axes). Thus, rotations of the prism form a dihedral group D_4 of order 8. Each of these 3 Sylow 2-subgroups contains two 90° rotations which generate one (of 3 in G) cyclic subgroups of order 4. All other non-identity elements of D_4 have order 2. The flips about the square's diagonals generate a subgroup in D_4 *isomorphic* to K_4 (one of 3 conjugated ones in G), while the flips about the square's midlines generate yet another K_4 — *the* Klein subgroup of $G \cong S_4$ (which is normal and hence common to all the three D_4). Finally, order 2 subgroups in G are generated by 180° rotations of two types: 6 about the lines passing through the centers of opposite edges (in each D_4 they are represented by the flips about the square's diagonals), and 3 about the lines passing through the centers of opposite faces (they correspond to the remaining 3 order 2 elements in each D_4).

Many different proofs of Sylow's theorems are known, but most seem to be based on clever choices of sets on which G acts.

To motivate our choice, note that the partition of a group G into left cosets of a subgroup S stays invariant under left translations (that's why G acts on G/S: a left translate of a coset is a coset). But the converse is also true: any partition of G invariant under left translations is the partition into left cosets of a subgroup. Namely, take S to be the part of the partition containing e. For any $h \in S$, the left translate $h^{-1}S$ by h^{-1} intersects S (since $h^{-1}h = e$) and hence must coincide with S. This shows that ratios $h^{-1}h'$ of elements from S lie in S, i.e. that S is a subgroup. All other parts of the partition, being left translates of S, are left S-cosets indeed.

Thus, given a group G of order $|G| = p^r m$ where p is prime, we introduce the set $\mathcal{P}$ of all partitions of G into m disjoint subsets of cardinality p^r. The action of G on itself by left translations induces an action of G on the set $\mathcal{P}$ of such partitions. A fixed point of this action, i.e. $x \in \mathcal{P}$ such that $gx = x$ for all $g \in G$, represents the partition of G into left cosets of subgroups S of order p^r. To prove Sylow's 1st theorem, we establish the existence of such fixed points by induction on m.

When $m = 1$ (i.e. G is a p-group) there is only one partition, which is the fixed point, so the base of induction holds. Our induction hypothesis says that all groups of order $p^r m'$ with $m' < m$ contain p-subgroups of order p^r.

According to the orbit-stabilizer theorem, an orbit $\mathcal{O}$ of the G-action on $\mathcal{P}$ has cardinality $|\mathcal{O}| = p^k n$ for some $k \leq r$ and some $n|m$, with the stabilizer G_x of each partition $x \in \mathcal{O}$ having order $p^{r-k}(m/n)$ respectively. One of the following 3 possibilities holds: (i) $|\mathcal{O}| = 1$, i.e. $\mathcal{O}$ is a fixed point, (ii) $|G_x| = p^r m'$ with $m' < m$, in which case G contains subgroups of order p^r because $G_x \subset G$ contains them by the induction hypothesis, or (iii) p divides $|\mathcal{O}|$. To show that the last possibility cannot hold for all orbits, we will prove that $|\mathcal{P}|$ is not divisible by p.

First, let's find the cardinality of $\mathcal{P}$:

$$|\mathcal{P}| = \frac{(p^r m)!}{(p^r!)^m m!}.$$

Indeed, thinking of a partition of G as represented by a Young tableau of m rows of length p^r filled with elements of G, we find that two tableaux represent the same partition if and only if they are obtained by $(p^r)!$ permutations in each row, $m!$ permutations of the rows, or any combination of these transformations.

Next, we claim that the multiplicity $o_p(n!)$ of the factor p in the prime factorization of $n!$ is

$$o_p(n!) = \lfloor \frac{n}{p} \rfloor + \lfloor \frac{n}{p^2} \rfloor + \lfloor \frac{n}{p^3} \rfloor + \cdots = \sum_{k=1}^{\infty} \lfloor \frac{n}{p^k} \rfloor,$$

where $\lfloor x \rfloor$ is the *floor function*, or *integer part* of x defined as the largest integer not exceeding x. Indeed, write the factors $1, 2, \ldots, n$ of $n!$ in a row in this order, and mark those which are divisible by p: $p, 2p, 3p, \ldots$. They occur every p positions of the row, so the total number of such positions not exceeding n equals $\lfloor n/p \rfloor$. But every p^2 positions the factor is divisible by an extra p, and this happens $\lfloor n/p^2 \rfloor$ times, and so on. Of course, the sum is actually finite, because the fractions n/p^k eventually become < 1.

Now let us apply this formula to the factorials in the numerator and denominator of our formula for $|\mathcal{P}|$:

$$o_p((p^r m)!) = p^{r-1}m + p^{r-2}m + \cdots + m + \sum_{k=1}^{\infty} \lfloor \frac{m}{p^k} \rfloor \ ,$$

$$o_p(p^r!) = p^{r-1} + p^{r-2} + \cdots + 1 \ , \quad o_p(m!) = \sum_{k=1}^{\infty} \lfloor \frac{m}{p^k} \rfloor \ .$$

We see that the multiplicity of p in the prime factorization of the numerator and denominator coincide, and consequently the ratio (which is in fact an integer) is not divisible by p at all — as promised.

To prove Sylow's 2nd theorem, consider the set G/S of left cosets of a given Sylow p-subgroup. It carries the transitive action of G induced by left translations, but we will restrict this action to elements of a p-subgroup $H \subset G$. This action of H has no reason to be transitive, and moreover, we claim that it has fixed points. Indeed, the cardinalities of all H-orbits other than fixed points, being non-trivial divisors of $|H|$, are divisible by p, but $|G/S|$ is not. So, let the coset $g_0 S$ be a fixed point of H. This means that for every $h \in H$ there is $s \in S$ such that $hg_0 e = g_0 s$, i.e. that $g_0^{-1} h g_0 \in S$. Thus, g_0^{-1} conjugates H into a subgroup of S.

To prove Sylow's 3rd theorem, consider the set (which we will denote by X) of all Sylow p-subgroups of G. The action of G on itself by conjugations induces an action of G on X, which is transitive by Sylow's 2nd theorem, implying that $|X|$ divides $|G|$. We will restrict this action to elements of a particular Sylow subgroup S.

The subgroup S considered as a point of X is a fixed point (since it is invariant under its own conjugations). We claim that it is the only fixed point. Since orbits of S other than fixed points have cardinalities divisible by p, this would imply that $|X| \equiv 1 \mod p$ as required.

To this end, consider any Sylow p-subgroup H, and introduce the *normalizer* subgroup $N(H)$. By definition, the normalizer (of any subset in G) consists of all group elements which leave the subset invariant under conjugation: $N(H) := \{g \in G \mid gHg^{-1} = H\}$. In other words, this is the stabilizer (and therefore a subgroup in G) of H considered as a point of X. Obviously $N(H)$ contains H as a normal subgroup. So, we can consider the projection $\pi : N(H) \to N(H)/H$ to the quotient group. The latter has order not divisible by p. Indeed, the order $|N(H)|$ is a divisor of $|G| = p^r m$, where m is coprime to p, but $|H| = p^r$, and so $|N(H)/H| = |N(H)|/|H|$ must be a divisor of m.

Suppose now that H is a fixed point of the action of S on X, i.e. conjugations by elements of S leave H invariant. Then $S \subset N(H)$. The range $\pi(S)$ of the projection homomorphism $\pi|_S : S \to \pi(S) \subset N(H)/H$ must have an order dividing the order p^r of S on the one hand, and the coprime to p order of the quotient group $N(H)/H$ on the other. Therefore, $|\pi(S)| = 1$, i.e. $S \subset \ker \pi = H$. Thus, indeed $H = S$ is the only Sylow p-subgroup fixed under conjugations by S.

EXERCISES

67. Construct a non-abelian group of order 8 not isomorphic to D_4.

68. Identify all Sylow subgroups in the alternating group A_4 and in the dihedral group D_6.

69. Prove that a non-abelian group of order $2p$ where p is an odd prime is isomorphic to the dihedral group D_p.

70. Show that Sylow subgroups in a group of order 15 are normal, and derive from this that the group is cyclic.

71. The same for any group of order pq where $p < q$ are primes such that $q \not\equiv 1 \mod p$.

72. For $|G| = p^2 q$ where p, q are distinct primes, show that G has a non-trivial normal subgroup.

73. Construct a Sylow p-subgroup in the symmetric group S_{p^r} inductively on r. `Hint:` Place p^r objects into the Young diagram of p rows of length p^{r-1} each and combine permutations from p copies of a Sylow p-subgroup of $S_{p^{r-1}}$, acting on the objects inside each row separately, with the group $\cong \mathbb{Z}_p$ cyclically permuting the rows.

Lecture 10. Finitely generated abelian groups

A. Direct sums and products. Given two groups, G' and G'', form their Cartesian product $G' \times G'' = \{(g', g'') \mid g' \in G', g'' \in G''\}$, and equip it with a groups structure using the component-wise operations: $(g'_1, g''_1)(g'_2, g''_2) := (g'_1 g'_2, g''_1 g''_2)$, $(g', g'')^{-1} := ((g')^{-1}, (g'')^{-1})$, $e := (e', e'')$. The resulting group is denoted $G' \times G''$ and is called the *direct product* of G' and G'', or $G' \oplus G''$ in which case it is called the *direct sum* of the two groups. Clearly, the projections

$$\pi' : G' \times G'' \to G',\ (g', g'') \mapsto g', \quad \pi'' : G' \times G'' \to G'',\ (g', g'') \mapsto g'',$$

are group homomorphisms with the kernels

$$\ker \pi' = \{(e', g'') \mid g'' \in G''\}, \quad \text{and} \quad \ker \pi'' = \{(g', e'') \mid g' \in G'\}$$

canonically identified with G'' and G' respectively. The groups $G' = \ker \pi''$ and $G'' = \ker \pi'$ considered as normal subgroups in $G' \times G''$ intersect trivially (i.e. at (e', e'') only) and commute with each other (i.e. $(g', e'')(e', g'') = (e', g'')(g', e'')$). They are *complementary* in the sense that each (g', g'') is uniquely written as the product $(g', e'')(e', g'')$ of an element from G' followed by an element from G''.

Conversely, if a group $\mathcal{G}$ contains two commuting complementary (in the above sense) normal subgroups $G', G'' \subset \mathcal{G}$, then $\mathcal{G}$ is canonically identified with $G' \times G''$. Namely, the compositions $G' \hookrightarrow \mathcal{G} \xrightarrow{\pi'} \mathcal{G}/G''$ and $G'' \hookrightarrow \mathcal{G} \xrightarrow{\pi''} \mathcal{G}/G'$, where π' and π'' are the projections to the corresponding quotient groups, identify these quotient groups with G' and G'' respectively. Indeed, writing $g \in \mathcal{G}$ as $g'g''$ with $g' \in G'$ and $g'' \in G''$ we find that the cosets gG'' and $g'G''$ coincide, showing that $\pi'|_{G'}$ is surjective; on the other hand, $g' \in \ker \pi'|_{G'}$ only if $g' \in G' \cap G'' = \{e\}$, implying that $\pi'|_{G'}$ is injective (and similarly for $\pi''|_{G''}$). Therefore the homomorphism $\mathcal{G} \to G' \times G''$ defined by $g \mapsto (\pi'(g), \pi''(g))$ and mapping $g = g'g''$ to (g', g'') (in terms of the previous identifications) is bijective. In this case one says that $\mathcal{G}$ is represented internally as the direct product of its subgroups.

`Example.` Integers modulo mn can be further reduced modulo m and modulo n. This defines two homomorphisms $\pi' : \mathbb{Z}_{mn} \to \mathbb{Z}_m$ and $\pi'' : \mathbb{Z}_{mn} \to \mathbb{Z}_n$, and therefore a homomorphism $(\pi', \pi'') : \mathbb{Z}_{mn} \to \mathbb{Z}_m \oplus \mathbb{Z}_n$. The kernel of it consists of the congruence classes $\mod mn$ of those integers $x \in \mathbb{Z}$ which are divisible by both m and n, i.e. divisible by the least common multiple of m and n. The kernel is

trivial if and only if the least common multiple is mn, i.e. when m and n are coprime. In this case, our homomorphism is injective, and hence surjective, since the orders of its domain and codomain coincide. We conclude that $G.C.D.(m,n) = 1$ implies $\mathbb{Z}_{mn} \cong \mathbb{Z}_m \oplus \mathbb{Z}_n$.

In the example, we used the notation $\oplus$ instead of $\times$ only because the operation in the groups $\mathbb{Z}_n$ is induced by addition of integers, with $\bar{0}$ as the unit element. However, when one generalizes the notion of direct product from the case of two to the case of many factors G_α, it becomes really different from the notion of direct sum whenever the collection of the factors is infinite. By definition the *direct product* $\prod_\alpha G_\alpha$, where α runs a certain index set I, consists of arbitrary collections $(g_\alpha)_{\alpha\in I}$ of elements $g_\alpha \in G_\alpha$, while the *direct sum* $\bigoplus_\alpha G_\alpha$ consists of such collections where, however, *all but finitely many* $g_\alpha = e_\alpha$. Thus, $\bigoplus_\alpha G_\alpha$ is a subgroup in $\prod_\alpha G_\alpha$ (since the operations in both groups are component-wise: $(g_\alpha)(\tilde{g}_\alpha) := (g_\alpha \tilde{g}_\alpha)$), both contain each G_{α_0} as a normal subgroup (consisting of collections (g_α) such that $g_\alpha = e_\alpha$ unless $\alpha = \alpha_0$), and both map onto each G_{α_0} (by $\pi_\alpha : (g_\alpha) \mapsto g_{\alpha_0}$). However, the operations of direct product and direct sum have different "categorical" properties.

A collection $f_\alpha : G \to G_\alpha$ of group homomorphisms from some group G to each G_α defines a homomorphism $f = (f_\alpha) : G \to \prod_\alpha G_\alpha$ to the direct product: $f(g) := (f_\alpha(g))_{\alpha\in I}$, such that its composition with each projection π_α returns the original homomorphism: $\pi_\alpha \circ f = f_\alpha$ for all $\alpha \in I$. Conversely, a homomorphism $f : G \to \prod_\alpha G_\alpha$ defines a collection $f_\alpha := \pi_\alpha \circ f : G \to G_\alpha$ of homomorphism from G to each G_α. One expresses this situation abstractly by saying that in the category of groups, the product group together with the projections $\pi_\alpha : \prod_\alpha G_\alpha \to G_\alpha$ is a "universally repelling object" for collections $(\to G_\alpha)_{\alpha\in I}$ of homomorphism to the factors.

In contrast, the direct sum plays the role of a "universally attracting object" for collections $(G_\alpha \to)_{\alpha\in I}$ of homomorphisms *from* the summands to *abelian* groups. Namely, a collection of homomorphisms $f_\alpha : G_\alpha \to A$ to some abelian group A defines a single homomorphism $f : \bigoplus_\alpha G_\alpha \to A$ such that its composition $f \circ i_\alpha$ with each canonical embeddings $i_\alpha : G_\alpha \hookrightarrow \bigoplus_\alpha G_\alpha$ coincides with f_α. The homomorphism f is defined by $f(g_\alpha) := \sum_\alpha f_\alpha(g_\alpha)$, where the summation (in the abelian group A whose operation is expressed by "+") makes sense because all but finitely many summands $f_\alpha(e_\alpha) = 0$.

In what follows, only finite direct sums (or products) will occur.

B. Free abelian groups. A group is called *finitely generated* if it is generated by finitely many elements, or in other words, if some finite subset of it is not contained in any proper subgroup. The additive group $\mathbb{Q}$ of rational numbers is not finitely generated because any finite collection of fractions has a common denominator, d, and is therefore contained in the proper infinite cyclic subgroup $\frac{1}{d}\mathbb{Z}$. Our goal is to classify up to isomorphism all finitely generated abelian groups. In particular, we will see that they all are isomorphic to the direct sums of finitely many (finite or infinite) cyclic groups.

For an abelian group A, will use "+" for the operation, and so what used to be the unit element is now $\mathbf{0}$, the inverse of $\mathbf{a}$ becomes $-\mathbf{a}$, and the kth power of $\mathbf{a}$ is now denoted $k\mathbf{a}$, where we also borrow the notation from linear algebra by using bold-faced letters for elements of the group. Thus, A being generated by $\mathbf{a}_1, \dots, \mathbf{a}_n$ means that every $\mathbf{x} \in A$ can be written (possibly non-uniquely) as a "linear combination" $\mathbf{x} = k_1\mathbf{a}_1 + \dots + k_n\mathbf{a}_n$ with some coefficients $k_i \in \mathbb{Z}$.

Let A be an abelian group, and $F(A) \subset A$ denote the subset of all its elements of finite order:

$$F(A) := \{\mathbf{a} \in A \mid \text{there exists } k > 0 \text{ such that } k\mathbf{a} = \mathbf{0}\}.$$

It is a subgroup (known as the *torsion subgroup* of A): if $k\mathbf{a} = \mathbf{0}$ and $l\mathbf{b} = \mathbf{0}$, then $kl(\mathbf{a} - \mathbf{b}) = \mathbf{0}$. Then the quotient group $A/F(A)$ is *free* in the sense that it has no non-zero elements of finite order. Indeed, if $k\bar{\mathbf{a}} = \bar{\mathbf{0}}$ in $A/F(A)$ for some $k > 0$ (here we denote by $\bar{\mathbf{a}}$ the coset $\mathbf{a} + F$ considered as en element of the quotient group) then $k\mathbf{a} \in F(A)$ has some finite order $l > 0$, and hence $kl\mathbf{a} = \mathbf{0}$ in A, i.e. $\mathbf{a} \in F(A)$ and $\bar{\mathbf{a}} = \bar{\mathbf{0}}$.

When A is generated by finitely many $\mathbf{a}_i \in A$ the quotient group is still generated by their cosets $\bar{\mathbf{a}}_i$ and therefore is still finitely generated.

`Theorem.` *A finitely generated free abelian group is isomorphic to $\mathbb{Z}^n$ for some $n \geq 0$ — the direct sum $\mathbb{Z} \oplus \dots \oplus \mathbb{Z}$ of finitely many copies of the infinite cyclic group $\mathbb{Z}$.*

`Proof.` The group $\mathbb{Z}^n$, which can be considered as the lattice of integer points in $\mathbb{R}^n$, is generated by the basis vectors $\mathbf{e}_1 = (1, 0, \dots, 0)$, $\mathbf{e}_2 = (0, 1, 0, \dots, 0), \dots, \mathbf{e}_n = (0, \dots, 0, 1)$, and moreover, each element $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{Z}^n$ is uniquely written as their linear combination: $\mathbf{x} = x_1\mathbf{e}_1 + \dots + x_n\mathbf{e}_n$ with integer coefficients. We need to show that any finitely generated free abelian group also has such a basis.

Let $\mathbf{a}_1, \ldots, \mathbf{a}_n$ be a finite collection of elements generating the group with *as few elements as possible*. We claim that they form a required basis, i.e. for every element $\mathbf{x}$ of the group, the coefficients $(x_1, \ldots, x_n)$ of the integer linear combination $\mathbf{x} = x_1\mathbf{a}_1 + \cdots + x_n\mathbf{a}_n$ are uniquely determined by $\mathbf{x}$. This indeed identifies the group with $\mathbb{Z}^n$: if $\mathbf{y} = y_1\mathbf{a}_1 + \cdots + y_n\mathbf{a}_n$, then $\mathbf{x} + \mathbf{y} = (x_1 + y_1)\mathbf{a}_1 + \cdots + (x_n + y_n)\mathbf{a}_n$, i.e. the sum $\mathbf{x} + \mathbf{y}$ of two vectors is expressed by the component-wise addition of the strings of their coordinates in $\mathbb{Z}^n$: $(x_1, \ldots, x_n) + (y_1, \ldots, y_n) = (x_1 + y_1, \ldots, x_n + y_n)$.

To justify our claim, assume that the same $\mathbf{x}$ can be expressed in two different ways as a linear combination of $\mathbf{a}_1, \ldots, \mathbf{a}_n$, and subtracting one from the other obtain $\mathbf{0} = k_1\mathbf{a}_1 + \cdots + k_n\mathbf{a}_n$, where at least one of k_i (for instance k_1) is non-zero. In linear algebra over $\mathbb{R}$ (or even over $\mathbb{Q}$) this would already lead to a contradiction, because we could divide the relation by k_1, express $\mathbf{a}_1$ as a linear combination of $\mathbf{a}_2, \ldots, \mathbf{a}_n$, and thereby furnish a collection of generators with fewer than n elements. However, to remain within our abelian group, we must avoid fractions. So, we use the division of integers with remainders, and apply a version of the Euclidean algorithm.

Namely, without loss of generality we may assume that among *non-zero* coefficients k_i in the relation $k_1\mathbf{a}_1 + \cdots + k_n\mathbf{a}_n = \mathbf{0}$, a smallest in the absolute value is k_1, and even may assume that $k_1 > 0$. Writing $k_2 = q_2k_1 + r_2, \ldots, k_n = q_nk_1 + r_n$ where the remainders r_i satisfy $0 \le r_i < k_1$, we form a new set of n generators by replacing $\mathbf{a}_1$ with $\mathbf{a}'_1 := \mathbf{a}_1 + q_2\mathbf{a}_2 + \cdots + q_n\mathbf{a}_n$, and keeping the rest of $\mathbf{a}_i$. The relation between the generators now reads: $k_1\mathbf{a}'_1 + r_2\mathbf{a}_2 + \cdots + r_n\mathbf{a}_n = \mathbf{0}$. Since our abelian group is free of non-zero elements of finite order, not all $r_i = 0$; otherwise $k\mathbf{a}'_1 = \mathbf{0}$ implying $\mathbf{a}'_1 = \mathbf{0}$ and making $\mathbf{a}_2, \ldots, \mathbf{a}_n$ a set of $n - 1$ generators in conflict with our choice of n. Thus, we can pick a smallest of all non-zero r_i (let it be r_2) and iterate the whole process, noting however that $0 < r_2 < k_1$. Since in each next iteration the smallest positive coefficient is smaller than the previous one, we eventually arrive at a relation $l_1\mathbf{b}_1 + \cdots + l_n\mathbf{b}_n = \mathbf{0}$ (between new generators) with $l_1 = 1$. This would allow us to eliminate $\mathbf{b}_1$ (as in linear algebra) from the set of generators, leading again to a contradiction with our choice of n.

`Remark.` The number n, being characterized as the smallest number of generators of $\mathbb{Z}^n$, does not depend on the choice of the generators, and is called the *rank* of the lattice. The rank $\operatorname{rk} A$ of a finitely generated abelian group A is defined as the rank of $A/F(A)$.

`Corollary.` *If A is finitely generated, then $A \cong F(A) \oplus \mathbb{Z}^{\operatorname{rk} A}$.*

`Proof.` In our discussion of cyclic groups, we have noticed that a homomorphism from $\mathbb{Z}$ to any group is determined by the image of the generator $1 \in \mathbb{Z}$, and moreover, since no multiple 1 is 0 in $\mathbb{Z}$, the image can be any element of the codomain group. Now from to the universality property of direct sums we conclude that a homomorphism from $\mathbb{Z}^n$ to any *abelian* group is determined by the images of the generators $f(\mathbf{e}_i)$ which could be arbitrary. Simply speaking, the rule $f(x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n) := x_1 f(\mathbf{e}_1) + \cdots + x_n f(\mathbf{e}_n)$ defines the homomorphism uniquely and doesn't lead to contradictions because there are no non-trivial relation between the generators $\mathbf{e}_i$ in $\mathbb{Z}^n$.

Following this lead, we pick in A some elements $\mathbf{a}_1, \ldots, \mathbf{a}_n$ such that their cosets $\bar{\mathbf{a}}_i$ in $A/F(A)$ form a basis in $A/F(A) \cong \mathbb{Z}^n$ and define $f : A/F(A) \to A$ by $f(\bar{\mathbf{a}}_i) := \mathbf{a}_i$. This homomorphism is a right inverse to the projection $\pi : A \to A/F(A)$, i.e. $\pi \circ f$ is the identity on $A/F(A)$. Indeed, by our very choice this is true on the generators of $A/F(A)$. Therefore f is injective, i.e. an isomorphism onto its range, which is therefore isomorphic to $\mathbb{Z}^n$ and hence contains no elements of finite order: $F(A) \cap f(A/F(A))) = \{\mathbf{0}\}$. Besides, for every $\mathbf{a} \in A$, $\mathbf{a} - f(\pi(\mathbf{a})) \in \ker \pi = F(A)$, and hence $\mathbf{a}$ can be decomposed into the sum of one element from $F(A)$ and one from the range of f. (This decomposition is unique exactly because the intersection of the two subgroups is trivial.) Thus, A is the internal direct sum of these subgroups.

`Corollary.` *The torsion subgroup $F(A)$ of a finitely generated abelian group is finite.*

`Proof.` It is generated by the images of the finitely many generators of A under the projection $A \cong F(A) \oplus (A/F(A)) \to F(A)$. To prove that a torsion group F with finitely many generators $\mathbf{f}_1, \ldots, \mathbf{f}_n$ is finite, take $N > 0$ such that $N\mathbf{f}_i = \mathbf{0}$ for all i (e.g. take N to be the product of the orders of $\mathbf{f}_i$). The homomorphism $\mathbb{Z}^n \to F$ defined by $(x_1, \ldots, x_n) \mapsto x_1\mathbf{f}_1 + \cdots + x_n\mathbf{f}_n$ is surjective. All vectors with components divisible by N lie in the kernel of it, and therefore the homomorphism factors through the component-wise projection $\mathbb{Z}^n \to (\mathbb{Z}/N\mathbb{Z})^n$, showing that $|F| \leq N^n$.

EXERCISES

74. Compute explicitly the map $\mathbb{Z}_6 \to \mathbb{Z}_3 \oplus \mathbb{Z}_2$ of the further reduction mod 3 and mod 2 of congruence classes mod 6. Repeat this for $\mathbb{Z}_8 \mapsto \mathbb{Z}_4 \oplus \mathbb{Z}_2$.

75. Prove that when $G.C.D.(m, n) > 1$, the cyclic group $\mathbb{Z}_{mn}$ is not isomorphic to $\mathbb{Z}_m \oplus \mathbb{Z}_n$.

76. Let a group $G', G'' \subset \mathcal{G}$ be subgroups with the properties: (i) $G' \cap G'' = \{e\}$, (ii) $g'g'' = g''g'$ for all $g' \in G'$ and $g'' \in G''$, and (iii) every $g \in \mathcal{G}$ can be written as $g'g''$ with $g' \in G$, $g'' \in G''$. Show that G' and G'' are normal subgroups, and $\mathcal{G}$ is their internal direct product.

77. Establish the uniqueness of the direct product (defined abstractly as a "universally repelling object") using only the universality property of it.

78. Prove that every subgroup in $\mathbb{Z}^2$ (the lattice of integer points on the plane) is isomorphic to $\{0\}$, $\mathbb{Z}$, or $\mathbb{Z}^2$. `Warning:` *A priori* it is not obvious at all whether every subgroup in $\mathbb{Z}^2$ is finitely generated.

79. Use induction on n in order to prove that every subgroup in $\mathbb{Z}^n$ is finitely generated.

Lecture 11. Finite abelian groups

A. Decomposition into p-groups. A finite abelian group F of order $p_1^{r_1}\cdots p_k^{r_k}$, where p_i are distinct primes, is canonically decomposed into the direct sum $F_1\oplus\cdots\oplus F_k$ of its (unique due to the commutativity) Sylow subgroups F_i of the orders $p_i^{r_i}$. However, this fact is more elementary than Sylow's theorems, so let us obtain it directly.

One says that an element $\mathbf{x}$ of an abelian group *belongs to exponent* n if $n\mathbf{x}=0$ (the term comes from the multiplicative notation $g^n=e$, and applies to non-abelian groups as well). In a finite group, all elements belong to the exponent $|F|$ by Lagrange's theorem.

Let m and n be coprime integers such that all elements of our finite abelian group F belong to the exponent mn (not-necessarily equal to the order of the group). Define in F two subgroups consisting of all elements belonging to the exponents m and n separately: $F':=\{\mathbf{x}\in F\mid m\mathbf{x}=\mathbf{0}\}$ and $F'':=\{\mathbf{x}\in F\mid n\mathbf{x}=\mathbf{0}\}$. (These are subgroups because F is abelian; in general $(ab)^n\ne a^nb^n$ and has not reason to equal e even if $a^n=e=b^n$.) By the Euclidean algorithm, there exist integers s and t such that $tm+sn=1$. Therefore every $\mathbf{x}\in F$ can be written as $\mathbf{x}=\mathbf{x}'+\mathbf{x}''$ where $\mathbf{x}'=sn\mathbf{x}\in F'$ and $\mathbf{x}''=tm\mathbf{x}\in F''$. On the other hand, the subgroups have trivial intersection: if $\mathbf{x}\in F'\cap F''$ then $\mathbf{x}=tm\mathbf{x}+sn\mathbf{x}=t\mathbf{0}+s\mathbf{0}=\mathbf{0}$. Therefore F decomposes into the internal direct sum $F'\oplus F''$ (where the possibility that, say, $F''=\{0\}$ and $F'=F$ is not excluded).

Applying this inductively to a finite abelian group F of order $p_1^{r_1}\cdots p_k^{r_k}$ we obtain the decomposition of F into the direct sum of subgroups $F_i:=\{\mathbf{x}\in F\mid p_i^{r_i}\mathbf{x}=\mathbf{0}\}$. Indeed, taking $m=p_1^{r_1}$, $n=|F|/m$ we find that $F=F'\oplus F''$ where $F'=F_1$ and F'' contains all the other F_i, since $n\mathbf{x}=\mathbf{0}$ for all $\mathbf{x}\in F_i$ with $i>1$. Proceeding this way with F'' and $m=p_2^{r_2}$, we find $F''=F_2\oplus F'''$ where F''' contains all F_i with $i>2$, and so on. At the end we have $F=F_1\oplus\cdots\oplus F_k$. To summarize:

`Proposition.` *A finite abelian group F of order $p_1^{r_1}\cdots p_k^{r_k}$ where p_i are distinct primes, canonically decomposes into the direct sum $F_1\oplus\cdots\oplus F_k$ of its subgroups F_i consisting of all elements belonging to the exponent $p_i^{r_i}$.*

B. Cyclic p-groups and their direct sums. Let F be a non-zero finite abelian group all of whose elements belong to the exponent p^r with some prime p and $r>0$. Eventually we are going to prove (by induction on r) that F can be decomposed into the direct sum

of cyclic subgroups whose orders are powers of p. In particular, this would imply that F is a p-group (which also follows from Sylow's 1st theorem). However, we start with analyzing the structure of such direct sums, and the first step is to look at one cyclic group $C \cong \mathbb{Z}_{p^r}$ of order p^r.

As in any cyclic group of any order n, the only subgroups in $\mathbb{Z}_{p^r}$ are cyclic (because this is true in $\mathbb{Z}$), consisting of all multiples of a divisor of n, which in the case $n = p^r$ are powers of p. Consequently, the subgroups form a tower $\mathbb{Z}_{p^r} \supset \mathbb{Z}_{p^{r-1}} \supset \mathbb{Z}_{p^{r-2}} \cdots \supset \mathbb{Z}_p \supset \{\mathbf{0}\}$ of cyclic groups formed by all multiples of $1, p, p^2, \ldots, p^{r-1}, p^r$ respectively (or, more accurately, of their congruence classes $\mod p^r$). Therefore, if $\mathbf{x}$ is a generator of $C \cong \mathbb{Z}_{p^r}$, then the subgroups of C form such a tower of cyclic subgroups are generated by $\mathbf{x}, p\mathbf{x}, p^2\mathbf{x}, \ldots, p^{r-1}\mathbf{x}, \mathbf{0}$.

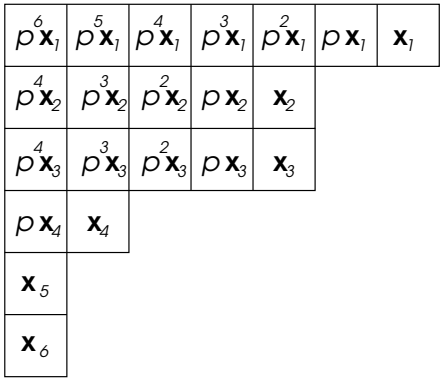


Figure 8: Abelian p-groups.

Consider now an abelian p-group of order p^n which is a direct sum of cyclic p-groups: $F = C_1 \oplus \cdots \oplus C_k$, where $C_i \cong \mathbb{Z}_{p^{n_i}}$. Then $p^n = p^{n_1} \cdots p^{n_k} = p^{n_1+\cdots+n_k}$, i.e. n_i form a partition of n. We will assume that $n_1 \geq n_2 \cdots \geq n_k > 0$, represent the partition by the corresponding Young diagram, and turn it into a Young tableau by filling the rows as shown in Figure 8. Namely, if $\mathbf{x}_i$ is a generator of C_i, then the ith row is filled from right to left by the generators $\mathbf{x}_i, p\mathbf{x}_i, \ldots, p^{n_i-1}\mathbf{x}_i$ of the cyclic subgroups of C_i.

As it is not hard do see, the choice of the generators $\mathbf{x}_i$, and even the mere decomposition of F into the direct sum of its cyclic p-subgroups is not unique (unless F is itself cyclic). Yet we have:

`Proposition.` *The partition $n = n_1 + \cdots + n_k$ (and hence the numbers $l_1, l_2, \ldots$, $n = l_1 + 2l_2 + 3l_3 + \ldots$, of cyclic p-subgroups of orders $p, p^2, \ldots$ in a decomposition of F into the direct sum of cyclic p-groups) does not depend on any choices, but is uniquely determined by the structure of the group,*

`Proof.` To the Young diagram with n cells representing the partition of n into the parts $n_1 \geq n_2 \geq \cdots$ equal to the row lengths of the diagram, associate the *dual partition* of n into the parts $m_1 \geq m_2 \geq \cdots$ equal the column lengths of it. We recover the dual partition from the structure of the group F as follows. Denote by $F_s \subset F$ the subgroup of all elements which belong to the exponent p^s: $F_s = \{\mathbf{a} \in F \mid p^s\mathbf{a} = \mathbf{0}\}$. Then $|F_s| = p^{m_1+\cdots+m_s}$. Indeed, every $\mathbf{a} \in F = \oplus_{i=1}^k C_i$ is written as $\mathbf{a} = k_1\mathbf{x}_1 + k_2\mathbf{x}_2 + \cdots$ uniquely in the sense that $\mathbf{a} = \mathbf{0}$ if and only if $k_i \equiv 0 \mod p^{n_i}$. Therefore $p^s\mathbf{a} = \mathbf{0}$ if and only if $k_i p^s \equiv 0 \mod p^{n_i}$. In other words, $\mathbf{a}$ lies in the direct sum $\bigoplus_{s \leq n_i} p^{n_i - s} C_i \bigoplus_{s > n_i} C_i$. The structure of this direct sum is represented by the left s columns of the Young tableau (Figure 8) associated with $\bigoplus_i C_i$. The order of this direct product equals $p^{m_1+\cdots+m_s}$ (p to the number of cells in the left s columns). Thus, the column lengths $m_s = \log_p |F_s|/|F_{s-1}|$ of the Young diagram are determined by the structure of the group, and hence the row lengths n_i are determined as well.

C. $\mathbb{Z}_p$-vector spaces. Let F be a finite abelian group all of whose elements belong to the exponent p (i.e. $F_1 = F$). One can consider F as a vector space of some finite dimension n over the field of scalars $\mathbb{Z}_p$, and conclude that $F \cong \mathbb{Z}_p^n$, the direct sum of n copies of $\mathbb{Z}_p$. Perhaps this subject is a good illustration to the fact that the basic theory of vector spaces looks the same whether the field of scalars is $\mathbb{R}$ or any other field. Not assuming that the reader is well-familiar with this idea, we provide more explicit details.

Let $\mathbf{f}_1, \ldots, \mathbf{f}_n$ be a set of generators of F with as few elements as possible. Then every $\mathbf{x} \in F$ can be written as $\mathbf{x} = x_1\mathbf{f}_1 + \cdots + x_n\mathbf{f}_n$ uniquely in the sense that if $x_1\mathbf{f}_1 + \cdots + x_n\mathbf{f}_n = \mathbf{0}$ then all $x_i \equiv 0 \mod p$. For if not, i.e. if say $x_1 \not\equiv 0 \mod p$, then for some $s, t \in \mathbb{Z}$, we have $sx_1 + tp = 1$, hence $\mathbf{f}_1 = (sx_1 + tp)\mathbf{f}_1 = s(x_1\mathbf{f}_1) = -s(x_2\mathbf{f}_2 + \cdots + x_n\mathbf{f}_n)$, making $\mathbf{f}_2, \ldots, \mathbf{f}_n$ a smaller set of generators. Therefore, the homomorphism $F \ni \mathbf{x} \mapsto (x_1, \ldots, x_n) \in \mathbb{Z}_p^n$ is bijective. Furthermore, if $F \ni \mathbf{y} = y_1\mathbf{f}_1 + \cdots + y_n\mathbf{f}_n$, then $\mathbf{x} + \mathbf{y} = (x_1 + y_1)\mathbf{f}_1 + \cdots + (x_n + y_n)\mathbf{f}_n$, i.e. the coordinates of $\mathbf{x} + \mathbf{y}$ in $\mathbb{Z}_p^n$ (which are uniquely determined $\mod p$) are $x_i + y_i$. Thus the operation in F translates into component-wise additions of the arrays of coordinates in $\mathbb{Z}_p^n$.

Moreover, as in usual linear algebra, any set $\mathbf{f}_1, \ldots, \mathbf{f}_k \in F$ linearly independent over $\mathbb{Z}_p$ (in the sense that $x_1\mathbf{f}_1 + \cdots + x_k\mathbf{f}_k = \mathbf{0}$ implies $x_i \equiv 0 \mod p$ for each i) can be completed to a basis $\mathbf{f}_1, \ldots, \mathbf{f}_k, \mathbf{f}_{k+1}, \ldots \mathbf{f}_n$. Namely, all linear combinations $x_1\mathbf{f}_1 + \cdots + x_k\mathbf{f}_k$ form in F a subgroup (really a linear subspace isomorphic to $\mathbb{Z}_p^k$). If it is not the whole of F yet (which ultimately means that $k < n$), pick any $\mathbf{f}_{k+1}$ which is not a linear combination of $\mathbf{f}_1, \ldots, \mathbf{f}_k$. If $x_1\mathbf{f}_1 + \cdots x_k\mathbf{f}_k + x_{k+1}\mathbf{f}_{k+1} = \mathbf{0}$, and $x_{k+1} \equiv 0 \mod p$, then all $x_i \equiv 0 \mod p$ by the assumption about linear independence of $\mathbf{f}_1, \ldots, \mathbf{f}_k$. But if $x_{k+1} \not\equiv 0 \mod p$, then $sx_{k+1} + tp = 1$ for some s, t, and we would be able to express $\mathbf{f}_{k+1}$ as a linear combination of $\mathbf{f}_1, \ldots, \mathbf{f}_k$ as before in conflict with our choice of $\mathbf{f}_{k+1}$. (The bottom line here is that in a field, e.g. $\mathbb{Z}_p$, one can divide by any non-zero element.) Therefore $\mathbf{f}_1, \ldots, \mathbf{f}_{k+1}$ are still linearly independent, and one can proceed this way to picking $\mathbf{f}_{k+2}$, etc. until linear combinations of $\mathbf{f}_i$ exhaust the whole of F.

D. Abelian p-groups. Let us now assume that F is any finite group whose elements belong to the exponent p^r for some $r > 0$, and prove by induction on r that F is decomposable into a direct sum $\bigoplus_i C_i$ of cyclic p-groups.

When $r = 1$, we have $F \cong \mathbb{Z}_p^n$ for some n, as it was explained in the previous section.

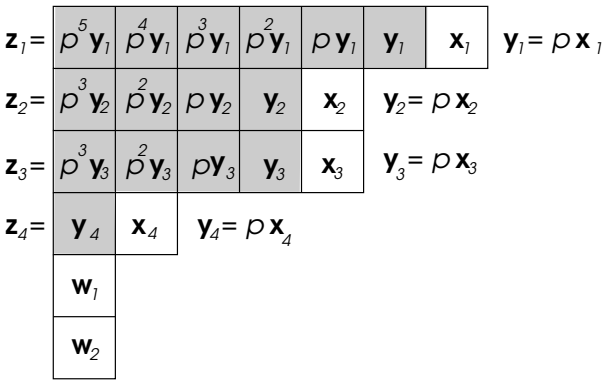


Figure 9: Decomposing into cyclic p-groups.

For $r > 1$, consider in F the subgroup pF consisting of p-multiples $\mathbf{b} = p\mathbf{a}$ of all elements $\mathbf{a} \in F$. Then $p^{r-1}\mathbf{b} = p^r\mathbf{a} = \mathbf{0}$, i.e. all elements of pF belong to the exponent p^{r-1}. By the induction hypoth-

esis, we may assume that pF is a finite direct sum $\bigoplus_\alpha C_\alpha$ of cyclic groups of some orders $p^{r_\alpha-1}$, where of course all $r_\alpha - 1 < r$. The structure of the group pF is illustrated by the shaded region in Figure 9, where each $\mathbf{y}_\alpha$ is some generator of C_α (and $r_\alpha - 1 = 6, 4, 4, 1, 1$ for $\alpha = 1, 2, 3, 4, 5$). Then $\mathbf{y}_\alpha = p\mathbf{x}_\alpha$ for some $\mathbf{x}_\alpha \in F$ which therefore generates in F a cyclic subgroup of order p^{r_α}. Besides, consider in F the subgroup $F_1 := \{\mathbf{a} \in F \mid p\mathbf{a} = \mathbf{0}\}$ of all elements belonging to the exponent p, and treat it as a vector space over $\mathbb{Z}_p$. Note that $\mathbf{z}_\alpha := p^{r_\alpha-2}\mathbf{y}_\alpha$ (the left shaded column) lie in F_1 and are linearly independent in it, because they lie in different direct cyclic summands of the "shaded" group pF. Complete the collection $\mathbf{z}_\alpha$ to a basis of F_1 as explained in the previous section by picking suitable elements $\mathbf{w}_\beta$ (with $\beta = 1, 2$ in the example of Figure 9). Thus, pF becomes the direct sum of cyclic subgroups of order p generated by $\mathbf{z}_\alpha$ (which lie in pF) and $\mathbf{w}_\beta$ (which don't). We claim that F is the direct sum of the cyclic subgroups $\tilde{C}_\alpha$ of orders p^{r_α} generated by $\mathbf{x}_\alpha$ and $\tilde{C}'_\beta$ of order p generated by $\mathbf{w}_\beta$.

Indeed, given $\mathbf{a} \in F$, we can write $p\mathbf{a}$ as $\sum_\alpha k_\alpha \mathbf{y}_\alpha$ for some $k_\alpha \in \mathbb{Z}$. Then $\mathbf{a} - \sum_\alpha k_\alpha \mathbf{x}_\alpha \in F_1$ and can therefore be written as a linear combination $\sum_\alpha l_\alpha \mathbf{z}_\alpha + l'_\beta \mathbf{w}_\beta$, where $\mathbf{z}_\alpha$ actually lie in $\tilde{C}_\alpha$. We conclude that all $\mathbf{x}_\alpha$ and $\mathbf{w}_\beta$ together generate F. Now suppose $\sum_\alpha k_\alpha \mathbf{x}_\alpha + \sum_\beta l'_\beta \mathbf{w}_\beta = \mathbf{0}$. We want to show that each l'_b is divisible by p and each k_α is divisible by p^{r_α}. This would imply the uniqueness of the decomposition of every element in F into the sum of elements from the cyclic subgroups generated by each $\mathbf{x}_\alpha$ and $\mathbf{w}_\beta$. Multiplying the relation by p we obtain $\sum_\alpha k_\alpha \mathbf{y}_\alpha = \mathbf{0}$ implying, due to our induction hypothesis, that each $k_\alpha = p^{r_\alpha-1} l_\alpha$ for some integer l_α. Since $p^{r_\alpha-1}\mathbf{x}_\alpha = \mathbf{z}_\alpha$, our relation assumes the form $\sum_\alpha l_\alpha \mathbf{z}_\alpha + \sum_\beta l'_\beta \mathbf{w}_\beta = \mathbf{0}$. By construction, $\mathbf{z}_\alpha$ and $\mathbf{w}_\beta$ form a basis in F_1 over $\mathbb{Z}_p$, and hence all l_α and l'_β are divisible by p. Consequently each $k_\alpha = p^{r_\alpha-1} l_\alpha$ is divisible by p^{r_α} as desired.

We arrive at the following main result of this lecture.

`Theorem.` *A finite abelian group canonically decomposes into the direct sum $\bigoplus_i F_i$ of finite abelian p_i-groups over distinct prime factors p_i of the order of the group. Each finite abelian p-group of order p^n is isomorphic to a direct sum $\mathbb{Z}_{p^{n_1}} \oplus \cdots \oplus \mathbb{Z}_{p^{n_k}}$ of cyclic p-groups, and different partitions $n_1 + \cdots + n_k$ of n correspond to pairwise non-isomorphic groups of the same order.*

EXERCISES

80. Prove that $\mathbb{Z}_m \oplus \mathbb{Z}_n \cong \mathbb{Z}_{mn}$ if and only if $G.C.D.(m,n) = 1$.

81. How many cyclic subgroups are there in the group $\mathbb{Z}_5^3$?

82. Show that $\mathbb{Z}_{p^r}$ has $p^r - p^{r-1}$ generators.

83. How many generators are there in a cyclic group of order $p_1^{r_1} \cdots p_k^{r_k}$ where p_i are distinct primes?

84. Classify up to isomorphism all abelian groups of order 32.

85. The same for abelian groups of order 48. Which of them are cyclic and which are not?

86. Let A and A' be abelian p-groups isomorphic to the direct sum of cyclic groups of orders $p^{n_1} \geq p^{n_2} \geq \cdots$ and $p^{n_1'} \geq p^{n_2'} \geq \cdots$ respectively. Prove that if A contains a subgroup isomorphic to A', then $n_i \geq n_i'$ for each i.

87. Find the places of $\mathbb{Z}_{17}^\times$ and $\mathbb{Z}_{32}^\times$ in the classification of abelian 2-groups.

88. Find the order of the multiplicative group $\mathbb{Z}_{144}^\times$, and find the place of this group in the classification of finite abelian group of this order.

89. Is the group $\mathbb{Z}_{1000}^\times$ cyclic?

90. Prove that $\mathbb{Z}_n^\times$ is cyclic if an only if n is one of the following: (a) a power p^r of an odd prime p, (b) $2p^r$, (c) 2 or 4.

91. Classify up to isomorphism all groups of order 8 (possibly non-abelian).

92. The same for orders 15, 45.

93. Classify up to isomorphism all finite groups of orders ≤ 15.

Appendix I: The AC, ZT, and ZL

A. Formulations. A *partial ordering* "$\prec$" on a set P is a binary relation which is transitive, and anti-symmetric (meaning that $x \prec y$ and $y \prec x$ are impossible simultaneously). A *linear* (or *total*) ordering is a partial ordering for which any two distinct elements are comparable, i.e. $x \neq y$ implies either $x \prec y$ or $y \prec x$. A totally ordered set in which every non-empty subset has a least element is called *well-ordered*. A subset $X \subset Y$ in a totally ordered (or well-ordered) set is called an *initial segment* if for every $x \in X$ all elements of Y such that $y \prec x$ are also in X. A totally ordered subset of a partially ordered set is often called a *chain*. An element $x_0 \in P$ is *maximal* if P contains no $x \succ x_0$.

`Axiom of Choice (AC).` *The Cartesian product of non-empty sets is non-empty. Equivalently: for any set X, there is a function φ associating to every non-empty subset of X an element of this subset.*

`Zermelo's Theorem (ZT).` *Every set can be well-ordered.*

`Zorn's lemma (ZL).` *A non-empty partially ordered set, in which every chain has an upper bound, contains at least one maximal element.*

B. ZL implies ZT. Given a set S, consider the set Σ of pairs $(X, \prec)$ where X is a non-empty subset of S, and $\prec$ is a well-ordering on X. Introduce partial ordering on Σ by $(X, \prec) \leq (Y, \prec)$ whenever $(X, \prec)$ is an initial segment of $(Y, \prec)$. Note that here $X \cup Y = Y$ is well-ordered, and the least element in X is the least in Y as well.

Let $\{(X_\alpha, \prec)\}$ be a chain in Σ. Then $X := \cup_\alpha X_\alpha$ is well-ordered by $\prec$. Indeed, X is totally ordered (since any two $x_1, x_2 \in X$ belong to some X_{α_1} and X_{α_2}, whose union is well-ordered), and all X_α have the same least element. Moreover, $(X_{\alpha_1}, \prec)$ is an initial segment in $(X_{\alpha_1} \cup X_{\alpha_2}, \prec)$. Therefore $x \prec x_1 \in X_{\alpha_1}$ implies that $x \in X_{\alpha_1}$, and therefore $(X_{\alpha_1}, \prec)$ is an initial segment in $(X, \prec)$.

Thus, Zorn's lemma applies to $(\Sigma, \leq)$, meaning that there is a maximal totally ordered subset $(S', \prec)$ in S. But then $S' = S$, for otherwise $x \in S - S'$ could be added to $(S', \prec)$ as an upper bound, contradicting its maximality.

C. ZT implies AC. This is obvious: well-order X, and to each nonempty subset of X, associate its least element.

D. AC implies ZL. Suppose that in a non-empty partially ordered set $(P, \prec)$ which does *not* have any maximal element, every chain C has an upper bound. Then the set of upper bounds not lying in C is non-empty, and using the Axiom of Choice, we can pick one for each C and denote it $\phi(C)$.

Let $\mathcal{C}$ denote the set of *well-ordered* chains C in P such that: $\phi(\emptyset)$ is the least element in C, and for every proper initial segment $D \subset C$, the least element in $C - D$ is $\phi(D)$. We have: $\{\phi(\emptyset)\} \in \mathcal{C}$.

We claim that for $C, C' \in \mathcal{C}$, one of them is an initial segment of the other. Indeed, let D be the maximal common initial segment of C and C' (i.e. the union of all common initial segments). If it is proper in both C and C', then $\phi(D)$, being the least element in both $C - D$ and $C' - D$, can be added to D to form a larger common initial set of C and C' in conflict with the maximality of D. Thus D coincides with one of C or C' making it an initial segment of the other.

Consider now the union U of all chains from $\mathcal{C}$. As in the derivation of Zermelo's theorem from Zorn's lemma, U is well-ordered, contains each $C \in \mathcal{C}$ as an initial segment, and has $\phi(\emptyset)$ as its least element. We claim that $U \in \mathcal{C}$, i.e. for any proper initial segment $D \subset U$, $\phi(D)$ is the least element in $U - D$. Indeed, an element $u \in U - D$ lies in some $C \in \mathcal{C}$, making D a proper initial segment of C, and implying that $\phi(D)$ is the least element of $C - D$, and hence of $U - D$ as well.

By our assumption (on non-existence of maximal elements in P), U must have an upper bound not in U. But U cannot have an upper bound not in U, since in this case $U \cup \{\phi(U)\}$ will be an element of $\mathcal{C}$ not contained in U, contradicting the choice of U. Thus, maximal elements exist.

EXERCISES

237. Prove that every vector space has a basis. (`Hint:` Apply Zorn's lemma to the set of linearly independent subsets of a given vector space, partially ordered by inclusion.)

Appendix II: The FTA

The *Fundamental Theorem of Algebra* says that a one-variable complex coefficient polynomial of positive degree has a complex root. It has many different proofs based on ideas from different subjects, such as, e.g. complex analysis, or algebraic topology. Here we present one of the classical proofs (usually associated with the names of Euler, de Foncenex, Lagrange, Laplace, or Gauss), which seems most suitable for an algebra course.

Any proof of the FTA must use something from analysis, that is — exploit somehow the difference between $\mathbb{R}$ and $\mathbb{Q}$, because the root which it claims to exist in $\mathbb{C} = \mathbb{R}(i)$ cannot, generally speaking, be found in $\mathbb{Q}(i)$ even when the coefficients of the polynomial are rational.

The only fact from analysis that we will use is the property of a polynomial $f = x^n + a_1 x^{n-1} + \cdots + a_n$ with *real* coefficients to have a real root *provided that the degree n is odd*. Indeed, since x^n is an odd function when n is odd, the value $f(x)$ tends to $+\infty$ as x tends to $+\infty$, and to $-\infty$ as x tends to $-\infty$. By the continuity of f, the value $f(x)$ must vanish somewhere in between.

This analytic fact will serve us as the base of induction, while the rest of the argument will be purely algebraic. The first observation is that the general case of polynomials f with complex coefficients can be reduced to the case of real coefficients by replacing f with the product $f\bar{f}$. Here $\bar{f}$ is obtained from f by conjugating its coefficients. Since $f\bar{f}$ is invariant under complex conjugation, the coefficients of it are already real. This special case is sufficient: if it is proved that $f(x_0)\bar{f}(x_0) = 0$ for some $x_0 \in \mathbb{C}$, then either $f(x_0) = 0$ or $\bar{f}(x_0) = 0$ in which case $f(\bar{x}_0) = 0$. In either case, a complex root of f is found.

Thus, without any loss of generality, it suffices to prove that a real coefficient polynomial of degree $n > 0$ has a complex root. The 18th century argument we are about to explain was criticized at the beginning of the 19th century for being based on the following bold yet poorly grounded assumption: that the polynomial f with real coefficients $a_1, \ldots, a_n$ is already factored as $f = (x - x_1) \cdots (x - x_n)$, and the FTA merely claims that at least one of the roots x_i lies in $\mathbb{C}$. The theory of algebraic field extensions sets this assumption on solid grounds: the abstract construction of the splitting field of a given polynomial guarantees that over some finite extension of $\mathbb{R}$ the polynomial can indeed be factored into linear factors.

Let $n = 2^k m$ where m is odd. The induction will be on $k = 0, 1, 2, \ldots$, with $k = 0$ as the base. So, we want to prove for $k > 0$ that at least one of the $2^k m$ roots x_i of any real coefficient polynomial f is complex, assuming that the same is true for any real coefficient polynomial of degree $2^{k-1}m'$ where m' is odd.

Consider the polynomial $g = \prod_{i<j}(x - y_{ij})$ whose roots are $y_{ij} = x_i x_j + c(x_i + x_j)$, and c is a real number yet to be selected. The degree of g equals $\binom{n}{2} = n(n-1)/2 = 2^{k-1}m(2^k m - 1)$ (where the third factor is odd since $k > 0$) and satisfies the requirement of the induction hypothesis.

The coefficients of g are symmetric functions of y_{ij}, but each y_{ij} is a symmetric expression of x_i, x_j. Permutations of $x_1, \ldots, x_n$ act on the set of pairs $\{x_i, x_j\}$ and thus merely permute y_{ij} leaving the coefficients of g invariant. By the theorem about symmetric functions (Lecture 17), the coefficients of g are expressible therefore as polynomials with real (since $c \in \mathbb{R}$) coefficients in elementary symmetric functions $\sigma_l(x_1, \ldots, x_n) = (-1)^l a_l \in \mathbb{R}$. Thus, the coefficients of g are real.

By the induction hypothesis, for any c at least one of the expressions $x_i x_j + c(x_i + x_j)$ is complex. Now we can apply the *Pigeonhole Principle*. There are infinitely many possible values of c, and finitely many pairs $i \neq j$. Therefore, for at least one pair $i \neq j$ we can find at least two values of c such that $x_i x_j + c(x_i + x_j)$ is complex. This means that for this i, j both $\alpha := x_i + x_j$ and $\beta := x_i x_j$ are complex. Therefore x_i, x_j are the solutions of the quadratic equation $x^2 - \alpha x + \beta = 0$ with complex coefficients. They can be expressed in terms of α and β by the quadratic formula. Since complex numbers have well-defined complex square roots, we conclude that x_i, x_j are complex.

Index

R-module, 73
p-group, 57
p-subgroup, 57

Abel's theorem, 145
abelian group, 21
action of group on set, 51
affine transformations, 18
algebra, 81
algebra homomorphism, 81
algebra of quaternions, 77
algebra of regular functions, 85
algebraic closure, 120
algebraic element, 111
algebraic extension, 112
algebraic number field, 113
algebraic set, 84
algebraically closed field, 119
alternating group, 29
anti-homomorphism, 51
anti-symmetric polynomial, 108
anti-symmetry, 14
ascending chain, 90
associativity, 20
automorphism group, 35
axiom of choice, 5

Bézout's theorem, 99
base of induction, 14
binary operation, 20
binary relation, 5

cancellation property, 84
cancellation rule, 20
Cardano's formula, 148
Cartesian product, 2
casus irreducibilis, 148
category, 25
category of rings, 81
category of unital rings, 81
center, 56
centralizer, 56
chain, 151
character, 139
characteristic, 111
Chinese Remainder Theorem, 13
class formula, 56
codomain, 2
commutative group, 21
commutative ring, 76
commutator, 36, 47
commutator subgroup, 36
complement, 2, 4
complex number, 1
composite element, 86
congruence, 7
conjugacy class, 40
conjugate elements, 127
conjugate subfields, 128
conjugated subgroups, 38
conjugation, 33
content, 96
content-free polynomial, 96
coprime integers, 10
countable set, 5, 15
cycle, 39
cycle decomposition, 39
cyclic extension, 138
cyclic subgroup, 26
cyclotomic field, 132
cyclotomic polynomial, 131

degree of extension, 111
derivative, 123
dihedral group, 29, 55
direct product, 61, 62
direct sum, 61, 62
discriminant, 107, 109, 147
distributive law, 73
divisibility, 14
domain, 2
dual partition, 69

Eisenstein's criterion, 101
element, 1
elementary symmetric function, 144
elementary symmetric functions, 105
embedding, 111
empty set, 6
epimorphism, 25

equivalence class, 6
equivalence relation, 6
equivalent divisors, 87
equivariant map, 53
Euclidean algorithm, 9
Euclidean domain, 87
Euler formula, 27
Euler's φ-function, 12
Euler's formula, 13
Euler's function, 33
Euler's theorem, 33
even permutation, 29
expressible in radicals, 137

factor group, 34
Fermat number, 135
Fermat's Little Theorem, 16, 32
field extension, 111
field of fractions, 93
finite cyclic group, 26
finite extension, 112
finitely generated group, 63
floor function, 59
free abelian group, 63
Frobenius automorphism, 121
function, 2
Fundamental Theorem of Algebra, 153

Galois group, 128
Gauss sum, 134
Gaussian integers, 77, 89
general linear group, 18, 27
graph, 2
greatest common divisor, 9
group, 15, 20
group action, 51
group algebra, 75
group of diffeomorphisms, 18
group of permutations, 17
group of units, 12, 87

Hamilton's relations, 77
homeomorphism, 17
homomorphism theorem, 34

ideal, 81
ideal generated by S, 86
identity, 17, 20
identity map, 3
index, 35
induction hypothesis, 14
infinite cyclic group, 26
initial segment, 151
integer, 1
integer part, 59
integral domain, 76
interior automorphism, 35, 38
intersection, 2
inverse, 20
inverse function, 3
inverse image, 4
irreducible element, 83, 87
irreducible number, 11
isomorphism, 22, 25
isomorphism classes, 22

kernel, 26
Klein subgroup, 23

Lagrange resolvent, 138
least common multiple, 39
left action, 51
left congruence, 31
left coset, 31
left ideal, 82
left inverse, 4
left translation, 31
Leibniz' rule, 123
length of permutation, 28
lexico-graphical ordering, 106
linear map, 27
linear ordering, 14, 151
localization, 93

map, 2
mapping, 2
mathematical induction, 13
maximal ideal, 83
minimal polynomial, 111
module of a ring, 73
monic polynomial, 99
monomorphism, 25
morphisms, 25
multiplication by scalars, 73
multiplicative system, 93

natural number, 1

Newton's identity, 110
Newton's polynomials, 110
non-separable elements, 124
norm, 11, 101
norm of quaternion, 78
normal extension, 127
normal subgroup, 33
normalizer, 60

objects, 25
odd permutation, 29
onto function, 3
opposite group, 52
order, 22
order of a group element, 32
orientation, 18
orthogonal group, 19
orthogonal transformations, 19

partial ordering, 14, 91, 151
partition, 6, 39
perfect field, 124
permutation, 17
PID, 87
Pigeonhole Principle, 154
prime element, 83
prime number, 11
primitive element, 124
primitive polynomial, 96
principal, 83
principal ideal domain, 87
principle of mathematical induction, 13
product, 86
projection map, 6

quadratic extension, 116
quadratic formula, 137
quaternionic conjugate, 78
quotient group, 34
quotient ring, 82
quotient set, 6

range, 3
rank, 64
rational functions, 95
rational number, 1
real number, 1
reducible element, 86
reflexivity, 5
regular maps, 85
remainder, 9
representation, 75
right action, 51
right congruence, 31
right coset, 31
right ideal, 82
right inverse, 4
right translation, 31
rigid motions, 19
ring, 73
ring homomorphism, 81
rotation group of the cube, 37

Schur's polynomial, 108
second isomorphism theorem, 141
separable element, 124
separable extension, 124
separable polynomial, 124
set, 1
sign of permutation, 28
simple group, 47
skew-field, 77
smooth structure, 18
solvable group, 141
special linear group, 18
special orthogonal group, 19
special unitary group, 78
splitting field, 127
stabilizer, 53
step of induction, 14
subgroup, 17, 25
subgroup generated by S, 26
subring, 78
subset, 2
sum of ideals, 86
surjective function, 3
Sylow p-subgroup, 57
Sylow's theorems, 57
symmetric group, 17
symmetric polynomial, 105
symmetricity, 5

third isomorphism theorem, 142
topological structure, 17
torsion subgroup, 63

total ordering, 14, 151
transcendental element, 111
transfinite induction, 14
transitivity, 5
transposition, 29
trivial ideals, 82
trivial representation, 75
trivial subgroups, 35, 46

UFD, 87
unary operation, 20
uncountable set, 5
undefinable notions, 1
union, 2
unique factorization domain, 87
unital R-module, 79
unital ring, 76
units, 87
upper bound, 91

value, 2
Vandermonde determinant, 108
Vandermonde's identity, 108
vector space, 73
Vieta's formulas, 105

Waring's algorithm, 105
well-ordering, 14, 151
Wilson's theorem, 49, 102

Young diagram, 39
Young tableau, 39

Zermelo's theorem, 15
zero divisor, 49, 76
Zorn's Lemma, 15
Zorn's lemma, 91