2. Groups

2.1. **Groups and monoids.** Let's start out with the basic definitions. We will consider sets with binary operations, which we will usually write multiplicatively, as $a \cdot b$, or, more commonly, just ab.

Before we proceed, let me make a few quick remarks on the terminology: we already encountered *relations,* which accept a fixed number of arguments and then output a truth value, true or false. An *operation* on a set also expects a fixed number of inputs, but it outputs another element of the set. We call relations and operations unary, binary, tertiary, and so on, according to the number of arguments they take.

Definition 2.1. (a) A *semigroup* is a set S with an *associative* binary operation: $(ab)c = a(bc)$ for all $a, b, c \in S$.
(b) A *monoid* is a semigroup M that has an *identity* (or *neutral element*): there exists $e \in M$ such that $ea = ae = a$ for all $a \in M$.
(c) An element a of a monoid is called *invertible* if there exists a $b \in M$ such that $ab = ba = e$. A *group* is a monoid in which every element is invertible.

Here are two quick general observations: The neutral element of a monoid M is unique because if e, e' are both neutral elements, then $e = ee' = e'$. It is common to denote it by 1 (rather than e).

Exercise 2.1. Show that similarly, if a is an invertible element of a monoid, and if $ab = ba = 1$ and also $ab' = b'a = 1$, then $b = b'$. (In particular, this hold for every element of a group.)

We call this unique b the *inverse* of a and denote it by $b = a^{-1}$.

Let's now look at a few examples: $\mathbb{N} = \{1, 2, 3, \ldots\}$ with addition $mn := m + n$ is a semigroup because addition is associative. If we include zero, then we obtain the monoid $(\mathbb{N}_0, +)$ (using self-explanatory notation), with the neutral element $e = 0$. If we also include the negative integers, then we obtain the group $(\mathbb{Z}, +)$; the inverse of $n \in \mathbb{Z}$ is $-n$, since $n + (-n) = (-n) + n = 0$. These examples have the additional property that $ab = ba$ for any two a, b. We say that these semigroups (monoids, groups) are *commutative;* in the case of groups, it is more common to speak of *abelian* groups. In abelian groups G, we often deviate from our notational convention and write the group operation as addition, as in $a + b$, and we denote the neutral element by 0.

The above examples still work if the arithmetic is done modulo k. More precisely, $(\mathbb{Z}_k, +)$ is an abelian group. Indeed, you showed in Exercise 1.11 that addition on $\mathbb{Z}_k$ is (well defined and) associative and commutative, and (0) is a neutral element. Moreover, $(n) + (-n) = (n + (-n)) = (0)$, so every $(n) \in \mathbb{Z}_k$ is invertible, with inverse $(-n)$.

If we instead use multiplication as the operation on $\mathbb{Z}_k$, then we still obtain a (commutative) monoid, by Exercise 1.11 again. The multiplicative identity element is of course given by $e = (1)$. This time $(\mathbb{Z}_k, \cdot)$ is not a group (if $k \geq 2$) because $0a \equiv 0 \mod k$ for all a, so 0 is not invertible. To try to fix this, let's remove zero and consider $\mathbb{Z}_k^\times = \mathbb{Z}_k \setminus \{(0)\}$.

Exercise 2.2. (a) Show that if p is a prime, then $\mathbb{Z}_p^\times$ is still a monoid; show also that if $k \geq 4$ is composite, then there are $a, b \in \mathbb{Z}_k^\times$ with $ab \equiv 0 \mod k$, so multiplication isn't even defined as an operation on $\mathbb{Z}_k^\times$ in this case.

(b) Use Proposition 1.9 to show that if p is prime, then $\mathbb{Z}_p^\times$ is in fact a group.

For an example of a non-commutative monoid or group, we can consider the collection of $n \times n$ matrices with entries in, say, $\mathbb{R}$. We denote this set by $M_n(\mathbb{R})$. Matrix multiplication is associative, so $M_n(\mathbb{R})$ is a semigroup and in fact a monoid with identity element $e = \operatorname{diag}(1, 1, \ldots, 1)$. Since matrix multiplication can depend on the order of the factors, this monoid is not commutative for $n \geq 2$.

Exercise 2.3. Give an explicit example of two matrices $A, B \in M_2(\mathbb{R})$ with $AB \neq BA$.

Since there are non-invertible matrices, $M_n(\mathbb{R})$ is not a group. However, if we only keep the invertible matrices, then this smaller set

$$GL(n, \mathbb{R}) = \{A \in M_n(\mathbb{R}) : A \text{ invertible}\}$$

is a (non-abelian, if $n \geq 2$) group. The notation refers to the usual name *general linear group* for this group. Recall also from linear algebra that an $A \in M_n(\mathbb{R})$ is in $GL(n, \mathbb{R})$ precisely if $\det A \neq 0$.

Another class of examples is obtained by considering maps (or functions) $f : X \to X$ on some set $X \neq \emptyset$. The collection of all such maps becomes a monoid under composition of functions $(f \circ g)(x) = f(g(x))$, and with the identity function $1(x) = x$ as the neutral element. (Since matrices may be identified with linear maps on $\mathbb{R}^n$, the previous examples M_n, $GL(n)$ are actually of this type.)

Exercise 2.4. Show that this is indeed associative, that is, $(f \circ g) \circ h = f \circ (g \circ h)$.

Exercise 2.5. Show that an f is invertible in this monoid precisely if f is bijective. (Note that there is some clash of terminology here:

usually, one calls a function invertible if it has an inverse function on *some* possibly smaller domain, and this condition is equivalent to the function being injective.)

Again, we can obtain a group by only keeping the bijective functions $f : X \to X$. Of particular interest is the case of a finite set X, and then we can just set $X = \{1, 2, \ldots, n\}$, for convenience. The bijective functions $\pi : \{1, \ldots, n\} \to \{1, \ldots, n\}$ are also called *permutations,* and the corresponding group of all permutations on the first n numbers is denoted by S_n and is called the *symmetric group.*

Exercise 2.6. Show that S_n is abelian precisely if $n = 1$ or $n = 2$.

Exercise 2.7. Let G be a group, $a, b \in G$. Show that $(a^{-1})^{-1} = a$ and $(ab)^{-1} = b^{-1}a^{-1}$.

Exercise 2.8. (a) Let G be a group. Suppose that $a, b \in G$, $ab = 1$. Show that then $a = b^{-1}$, $b = a^{-1}$.
(b) Now let M be a monoid. Suppose that $a, b \in M$, $ab = 1$. Show that it does *not* follow that a, b are invertible, by providing suitable counterexamples.
(c) Make sure you understand completely why (a), (b) don't contradict each other.

The procedure we used above to extract a group from a monoid works in general:

Exercise 2.9. Let M be a monoid, and let $U(M)$ be the set of its invertible elements ("units"). Show that $U(M)$ is a group. (What exactly do you need to show here?)

Exercise 2.10. Which of the following sets A are groups (monoids, semigroups) with the specified operation?
(a) $A = \mathbb{Z}$, $ab := a \cdot b$ (multiplication in $\mathbb{Z}$);
(b) fix a set B and let $A = P(B)$, the power set of B (= the collection of all subsets of B), $ab := a \cup b$;
(c) $A = P(B)$, $ab = a \setminus b$;
(d) $A = \mathbb{N}$, $ab = \gcd(a, b)$;
(e) $A = \mathbb{N}$, $ab = \operatorname{lcm}(a, b)$ (the least common multiple of a, b)

Exercise 2.11. In the previous Exercise, in those examples that define monoids, find the group of invertible elements.

Exercise 2.12. (a) Show that in a group G, given $a, b \in G$, the equations $ax = b$ and $ya = b$ have (in fact, unique) solutions $x, y \in G$.

(b) Let S be a semigroup. Show that, conversely, S will be a group if for any $a, b \in S$, the equations $ax = b$ and $ya = b$ always have solutions

$x, y \in S$. *Suggestion:* Start out by showing that there are left and right identities, that is, there are $1_L, 1_R \in S$ such that $1_L a = a 1_R = a$ for all $a \in S$.

Associativity allows us to drop parentheses: if a, b, c lie in a semigroup S, then we can unambiguously write abc because the (in principle) two ways of evaluating this product, $(ab)c$ and $a(bc)$, give the same answer. The same property holds for products of arbitrary length: $a_1 a_2 \ldots a_n$ always has the same value, no matter where we put the parentheses.

This is unsurprising and can be checked easily in concrete examples. For example, why is $a((bc)d)$ the same as, say, $((ab)c)d$? Well, we just repeatedly apply associativity in its basic version, for three factors, to obtain that

$$a((bc)d) = (a(bc))d = ((ab)c)d,$$

as desired.

Exercise 2.13. List all possibilities of putting parentheses in a product with four factors (your list should have five entries), and convince yourself that these are all equal to one another.

Now let's try to do the general case of a product $a_1 a_2 \ldots a_n$ with n factors. We will proceed by induction on n. For $n = 3$, this is just the original associative law (= basis of our induction). For the inductive step, assume associativity for products with $\le n - 1$ factors. It is convenient to temporarily agree that a product with no parentheses will be evaluated from left to right, that is, $a_1 \ldots a_n = (\ldots ((a_1 a_2) a_3) \ldots a_n)$. We will now show that any method of evaluating the product gives the same answer $a_1 \ldots a_n$.

Evaluate the individual product that comes first. We might have to make a choice here, in examples such as $(ab)(cd)$; in this case, take the leftmost of these products (ab in the example). Let's say $p = a_k a_{k+1}$ is the product we evaluated. This now leaves us with a product of the form $a_1 \ldots a_{k-1} p a_{k+2} \ldots a_n$, with parentheses (!), with one fewer factor, so the induction hypothesis applies and we may ignore those invisible parentheses and evaluate this from left to right anyway. Let's focus on the first part of this evaluation, until the moment when we reach p. The overall structure of this product is $qp = q(a_k a_{k+1})$, where $q = a_1 \ldots a_{k-1}$. Now by associativity for three factors, we have that $qp = (qa_k) a_{k+1}$, but this is just $a_1 \ldots a_{k+1}$, so our claim follows.

Exercise 2.14. What happens to this argument if $k = 1$? Convince yourself that this case can be handled, too.

This whole treatment provides a rather typical example of a kind of argument that comes up with some regularity. The statement that we established (generalized associativity) looks very obvious, so it ought to have a very quick and easy proof, and indeed no brilliant unexpected ideas are needed, but it actually turns out that organizing a clean formal argument requires some care and mental tidiness.

2.2. **Isomorphisms and Cayley's Theorem.** We first need a few definitions. Let's start with the notion of an isomorphism. We think of isomorphic structures as being the same, except possibly for the names you gave to their elements. An isomorphism is a map that implements this identification. In our setting, it must in particular preserve the algebraic structure, and maps of this type are called *homomorphisms.* In the case of monoids and groups, this takes the following form.

Definition 2.2. Let M, M' be (both) monoids or (both) groups with identity elements 1 and $1'$, respectively. A map $\varphi : M \to M'$ is called a *homomorphism* if $\varphi(1) = 1'$ and $\varphi(ab) = \varphi(a)\varphi(b)$ for all $a, b \in M$.

In the case of groups, we also have inverses as part of the algebraic structure, and we want our maps to preserve this, too, in the sense that $\varphi(a)^{-1} = \varphi(a^{-1})$ for all $a \in G$. This doesn't have to be imposed as an extra condition, though; it follows automatically.

Proposition 2.3. *If $\varphi : M \to M'$ is a homomorphism and $a \in M$ is invertible, then so is $\varphi(a)$ and $\varphi(a)^{-1} = \varphi(a^{-1})$. In particular, this holds for every element of a group.*

Proof. Apply φ to all members of $aa^{-1} = a^{-1}a = 1$ to obtain that

$$\varphi(a)\varphi(a^{-1}) = \varphi(a^{-1})\varphi(a) = 1',$$

and this says that $\varphi(a)$ is invertible in M' with inverse $\varphi(a^{-1})$. □

In fact, in the case of groups, it is also possible to drop the requirement that $\varphi(1) = 1'$ from Definition 2.2:

Exercise 2.15. (a) Let G, G' be groups and let $\varphi : G \to G'$ be a map satisfying $\varphi(ab) = \varphi(a)\varphi(b)$ for all $a, b \in G$. Show that then $\varphi(1) = 1'$.

(b) However, show that if φ is a map as in part (a) between *monoids,* then it can happen that $\varphi(1) \neq 1'$.

Definition 2.4. An *isomorphism* is a bijective homomorphism. Monoids or groups are called *isomorphic* if there is an isomorphism between them.

So in addition to preserving the algebraic structure, an isomorphism also preserves the underlying set; in other words, it preserves the complete structure of a monoid or group. Observe also that if $\varphi : G \to G'$ is an isomorphism between groups or monoids, then so is the inverse map $\varphi^{-1} : G' \to G$, so the condition of being isomorphic is symmetric in G, G', as was already suggested by the way we phrased the definition.

Exercise 2.16. Show that if $\varphi : G \to G'$ is an invertible homomorphism, then $\varphi^{-1} : \varphi(G) \to G$ is a homomorphism, too.

We write $G \cong G'$ to express the fact that G, G' are isomorphic.

Exercise 2.17. Show that $\cong$ is an equivalence relation between groups.

Example 2.1. I claim that the groups $(\mathbb{R}, +)$ and $(\mathbb{R}_+, \cdot)$ are isomorphic. Here, $\mathbb{R}_+ = (0, \infty)$ denotes the positive real numbers. (Convince yourself that these are indeed groups.) An isomorphism is given by the exponential function $\varphi : \mathbb{R} \to \mathbb{R}_+$, $\varphi(x) = e^x$. This map is indeed bijective, by the elementary properties of the exponential function, and it is a homomorphism because $\varphi(xy) = e^{x+y}$, which is equal to $\varphi(x)\varphi(y) = e^x e^y$, as required.

Exercise 2.18. Give a very careful and explicit interpretation of this set of formulae. Note that ab may refer to multiplication in the group, which, to make it more confusing, really is addition in the case of $(\mathbb{R}, +)$, or it may denote multiplication as real numbers.

On the other hand, the groups $(\mathbb{R}, +)$ and $GL(2, \mathbb{R})$ are not isomorphic. This follows because the first group is abelian while the second one isn't (remember: isomorphic groups are really the same abstract structure, except possibly for the names given to the elements). More formally, if $A, B \in GL(2, \mathbb{R})$, then any homomorphism $\varphi : GL(2, \mathbb{R}) \to \mathbb{R}$ has to map AB and BA to the same real number because

$$\varphi(AB) = \varphi(A) + \varphi(B); \qquad \varphi(BA) = \varphi(B) + \varphi(A)$$

and addition of real numbers is commutative, of course. Since there are A, B so that $AB \neq BA$, this means that there aren't any injective homomorphisms.

Exercise 2.19. Consider the nth roots of unity

$$G = \left\{1, e^{2\pi i/n}, e^{2\pi i\cdot 2/n}, \ldots, e^{2\pi i(n-1)/n}\right\}$$

(in other words, these are the n complex solutions of $z^n = 1$), with multiplication as complex numbers as the operation.
(a) Show that G is a group.
(b) Show that G is isomorphic to $(\mathbb{Z}_n, +)$.

Let G be a group. A *subgroup* of G is, by definition, a subset $H \subseteq G$ that is a group itself, with the same operation as G. Of course, associativity carries over automatically from G, so we need H to contain 1 and whenever $a, b \in H$, we must have that $ab \in H$ and $a^{-1} \in H$. A slightly more elegant formulation is possible:

Proposition 2.5. *Let G be a group. A subset $H \subseteq G$, $H \neq \emptyset$, is a subgroup of G precisely if $a, b \in H$ implies that also $ab^{-1} \in H$.*

Proof. Clearly a subgroup has this property because neither inverses nor products lead us out of H. Conversely, if $ab^{-1} \in H$ whenever $a, b \in H$, then $1 = aa^{-1} \in H$. So, taking $a = 1$, we now see that if $b \in H$, then also $b^{-1} = 1b^{-1} \in H$. Finally, if $a, b \in H$, then $b^{-1} \in H$, as we just saw, so $ab = a(b^{-1})^{-1} \in H$ by assumption. □

Of course, there is an analogous notion of a *submonoid* of a given monoid M. More precisely, we call $N \subseteq M$ a submonoid if $1 \in N$ and $ab \in N$ whenever $a, b \in N$.

Another useful observation is that if $\varphi : G \to G'$ is a homomorphism, then the image $\varphi(G) = \{\varphi(a) : a \in G\}$ is a subgroup of G'.

Exercise 2.20. Prove this.

Recall from the previous section that if X is any (non-empty) set, then the functions $f : X \to X$ form a monoid with composition as the operation, and the bijective functions form a group. Let us denote these by $M(X)$ and $S(X)$, respectively. Cayley's theorem says that any group can be realized as a subgroup of $S(X)$, for a suitable set X (in fact, we are going to take $X = G$ in the proof below). An analogous statement holds for monoids, but we won't make this explicit. We also call such a subgroup of $S(X)$ a *group of transformations.*

Theorem 2.6 (Cayley)**.** *Let G be a group. Then G is isomorphic to a group of transformations.*

Proof. As already announced, we will, more specifically, set up an injective homomorphism $\varphi : G \to S(G)$. As observed above, the image $\varphi(G)$ will then be a subgroup of $S(G)$; it will be the group of transformations we are looking for. Indeed, φ considered as a map from G to $\varphi(G)$ is an isomorphism, so we are done as soon as we have such a φ.

Define $\varphi(a)(x) = ax$. This map $\varphi(a) : G \to G$ is injective because if $ax = ay$, then by multiplying by a^{-1} from the left, we see that $x = y$, and $\varphi(a)$ is also surjective because if $b \in G$ is given, then $\varphi(a)(a^{-1}b) = aa^{-1}b = b$. So $\varphi(a)$ is bijective; in other words, $\varphi(a) \in S(G)$.

I now claim that φ is an injective homomorphism. Certainly φ is injective because if $\varphi(a) = \varphi(b)$, then in particular $\varphi(a)(1) = \varphi(b)(1)$, but these equal a and b, respectively, by the definition of φ, so $a = b$. To verify that φ is a homomorphism, let $a, b \in G$. Then

$$\varphi(ab)(x) = (ab)x = a(bx) = \varphi(a)(\varphi(b)(x)). \tag{2.1}$$

The right-hand side may be interpreted as the composite function $\varphi(a) \circ \varphi(b)$, applied to x. Since (2.1) holds for all $x \in G$, it follows that $\varphi(ab) = \varphi(a) \circ \varphi(b)$, and since composition is the group operation on $S(G)$, this says that φ is a homomorphism. □

Corollary 2.7. *Any finite group is isomorphic to a subgroup of S_n.*

Recall that the symmetric group S_n was defined as the group of permutations on n symbols. Corollary 2.7 really follows from the proof of Cayley's Theorem that was given (not from the statement, at least not immediately); we also obtain that we can take $n = |G|$, the number of elements of G.

Exercise 2.21. Formulate and prove the version of Cayley's Theorem for monoids.

Exercise 2.22. Show that the transition to a subgroup is necessary in general. In other words, find a group G that is not isomorphic to the full group $S(X)$ for any set X.

Exercise 2.23. Let $H_1, H_2 \neq G$ be two subgroups of a group G. Show that $H_1 \cup H_2 \neq G$.

2.3. **Subgroups and cyclic groups.** Given a group G and an arbitrary subset $S \subseteq G$, there is always a smallest subgroup $H = H(S) \subseteq G$ that contains S. More explicitly, the defining properties of H are: (i) $H \supseteq S$; (ii) H is a subgroup of G; (iii) if H' also satisfies (i), (ii), then $H' \supseteq H$.

We first need to make sure that such an object $H = H(S)$ indeed always exists (if you're not convinced that something needs to be shown here, then compare the description of H with the formally analogous but nonsensical "the smallest infinite subset of $\mathbb{N}$ that contains 5353").

Lemma 2.8. *If the $H_\alpha \subseteq G$ are subgroups of G, then so is $\bigcap_\alpha H_\alpha$.*

Proof. This is immediate from the criterion from Proposition 2.5: Write $H = \bigcap H_\alpha$. If $a, b \in H$, then $a, b \in H_\alpha$ for each α, so $ab^{-1} \in H_\alpha$ because H_α is a subgroup, so $ab^{-1} \in H$, as required.

Note also that the intersection is non-empty because $1 \in H_\alpha$ for all α, so 1 is in it. □

Given a subset $S \subseteq G$, we can now let $H(S) = \bigcap H'$, where the intersection is over all subgroups H' of G with $H' \supseteq S$. Observe that one possible choice is $H' = G$, so there are such subgroups H' and the intersection is not over the empty collection. By the Lemma, this set $H(S)$ is a subgroup of G, and it also satisfies properties (i), (iii).

Exercise 2.24. Check this in (slightly) more detail.

We will also use the alternative notation $\langle S\rangle$ instead of $H(S)$; if $S = \{s_1, \ldots, s_n\}$ is finite, we will usually write $\langle s_1, \ldots, s_n\rangle$ instead of the formally correct, but too pedantic $\langle\{s_1, \ldots, s_n\}\rangle$.

The procedure just given builds $\langle S\rangle$ from the top down, so to speak, by starting with $H' = G$ and then cutting it down to size. We can also build $\langle S\rangle$ from the bottom up. We do this by successively putting elements into a set H, but only those that we are sure must definitely be in $H = \langle S\rangle$.

Now clearly, we must have $s \in H$ for all $s \in S$ and also $1 \in H$. Next, we must be able to take products and inverses in the subgroup we are looking for, so we must also insist that $s_1 s_2 \ldots s_n \in H$ if $s_j \in H$ or $s_j^{-1} \in H$ for each j (I got these expressions by taking inverses and/or products finitely many times, starting out from the elements of S). Then we must demand that expressions formed with the help of inverses and products from *these* words $s_1 s_2 \ldots s_n$ lie in H, and whatever we get from this can then be made the starting point of a fresh round of operations etc. We now become worried that this process will never stop, but fortunately that is not the case and in fact we have reached the finish line already:

Proposition 2.9. *Let $S \subseteq G$. Then*

$$\langle S\rangle = \{s_1 s_2 \ldots s_n : n \geq 0, s_j \text{ or } s_j^{-1} \in S\} \tag{2.2}$$

For $n = 0$, we interpret this (empty) product as 1.

Proof. Denote the set on the right-hand side of (2.2) by H. We compare H with the description $\langle S\rangle = \bigcap H'$ that was given above. If $H' \supseteq S$ is any subgroup that contributes to the intersection, then clearly H' must contain all the products from (2.2). Thus $\langle S\rangle \supseteq H$.

On the other hand, H is a subgroup because if p, q are products as in (2.2), then pq^{-1} is again of this form. Moreover, $H \supseteq S$, by just taking $n = 1$ in (2.2). So H is a subgroup containing S, and since $\langle S\rangle$ is the smallest such group, this shows that $\langle S\rangle \subseteq H$. □

Depending on what structure exactly you want to build, this kind of procedure (start with the generating set, keep applying the operations that our structure is supposed to be closed under) may or may not stabilize after finitely many steps. An example of the second type that you may be familiar with from Real Analysis is provided by the notion of a σ-algebra generated by a collection of subsets (for example the Borel σ-algebra on $\mathbb{R}$): you start out with the open sets, then you take countable intersections and complements of these sets, then, in the next round, you apply these same operations to the sets you just obtained, etc. etc. It just never stops; the sets keep getting more complicated (you can still build the Borel σ-algebra from the bottom up, but you need *ordinals* for this).

Let's now return to our discussion of $\langle S\rangle$. The simplest example of this should be that of a subgroup $\langle a\rangle$ generated by just one element $a \in G$. We call $\langle a\rangle$ the *cyclic subgroup* generated by a. Similarly, we call a group G a *cyclic group* if $G = \langle a\rangle$ for some $a \in G$.

It will be convenient to use (almost self-explanatory) exponential notation. We denote by $a^n = aa\ldots a$ the n-fold product of a with itself. This is well defined, for $n \ge 1$, by general associativity. In fact, we can naturally define a^n for arbitrary $n \in \mathbb{Z}$, by putting $a^0 := 1$ and, for $n < 0$, $a^n := (a^{-1})^{|n|}$.

Exercise 2.25. Show that we have the power laws $a^{m+n} = a^m a^n$ and $(a^m)^n = a^{mn}$ for arbitrary $m, n \in \mathbb{Z}$.

Exercise 2.26. Is it also always true that $(ab)^n = a^n b^n$ for $n \in \mathbb{Z}$ and $a, b \in G$ for an arbitrary group G?

From Proposition 2.9 and Exercise 2.25, it is then clear that

$$\langle a\rangle = \{a^n : n \in \mathbb{Z}\}.$$

From this and Exercise 2.25 again, we obtain the surjective homomorphism

$$\varphi : (\mathbb{Z}, +) \to \langle a\rangle, \qquad \varphi(n) = a^n.$$

Either φ is also injective and thus an isomorphism, or $a^m = a^n$ for some $m, n \in \mathbb{Z}$, $m \ne n$. In the first case, $\langle a\rangle \cong \mathbb{Z}$.

In the second case, suppose that $n > m$. Then $a^{n-m} = 1$, that is, we also find a positive integer k with $a^k = 1$. Now let $k \ge 1$ be the smallest such integer. Then I claim that

$$\langle a\rangle = \{1, a, a^2, \ldots, a^{k-1}\}, \qquad \langle a\rangle \cong \mathbb{Z}_k;$$

this includes the claim that the k powers of a that are listed are all distinct. This latter claim is indeed clear because if we had $a^m = a^n$, then, as we just saw, also $a^{m-n} = 1$, so if we had $0 \le m, n \le k-1$ here and $m \ne n$, then we would obtain an integer $d = m - n$ or $d = n - m$ with $0 < d < k$ and $a^d = 1$, and this contradicts the definition of k.

It remains to show that for arbitrary $n \in \mathbb{Z}$, we have that $a^n = a^r$ for some $0 \le r < k$. To do this, we divide n by k with remainder: we write $n = qk + r$, and here $q \in \mathbb{Z}$, $0 \le r < k$. Then, by the power laws from Exercise 2.25, $a^n = a^{qk}a^r = (a^k)^q a^r = a^r$, as desired.

Finally, an isomorphism between $\langle a \rangle$ and $\mathbb{Z}_k$ in this situation is set up by mapping $a^n \mapsto n$.

Exercise 2.27. Verify that this indeed defines an isomorphism.

We summarize:

Theorem 2.10. *Let $G = \langle a \rangle$ be a cyclic group. Then $G \cong \mathbb{Z}_k$ or $G \cong \mathbb{Z}$. In the first case, $G = \{1, a, a^2, \ldots, a^{k-1}\}$; in the second case, the powers a^n, $n \in \mathbb{Z}$, are all distinct.*

Cyclic groups are the groups that are easiest to study, and we will establish more results about them in a moment. Let me first introduce some terminology. If G is a finite group, the *order* of G refers to the number of elements of G. We usually denote it by $|G|$. If $a \in G$ is an element of an arbitrary group and $a^n = 1$ for some $n \in \mathbb{Z}$, $n \ne 0$, then we say that a has finite *order,* and we define the order of a as the smallest positive integer n with $a^n = 1$. In this case, we also write $o(a) = n$. Note that if $o(a) = n$, then the cyclic group generated by a has the same order, $|\langle a \rangle| = n$, so these two notions of an order are to some extent compatible.

Exercise 2.28. (a) Show that $a^n = 1$ precisely if $a^{-n} = 1$.
(b) Give an example of a group in which no non-identity element has finite order.
(c) Suppose that $o(a) = n$. Show that $a^k = 1$ precisely if $n|k$.

Theorem 2.11. *(a) The subgroups of $\mathbb{Z}$ are precisely the groups $\langle k \rangle = \{kn : n \in \mathbb{Z}\}$, $k = 0, 1, 2, \ldots$.*

(b) If $|\langle a \rangle| = t$, then for every divisor $n|t$, $1 \le n \le t$, there is exactly one subgroup of $G = \langle a \rangle$ of order n, and there are no other subgroups of G.

Proof. (a) If $H \subseteq \mathbb{Z}$ is a subgroup, then either $H = \{e\} = \{0\}$, corresponding to $k = 0$, or H contains non-identity elements. In this case, let $k \ge 1$ be the smallest positive element of H (why are there positive elements in H?). We can now repeat an argument we already used above, in a slightly different context, to see that $H = \{kn : n \in \mathbb{Z}\}$: Clearly, kn, which is an $|n|$-fold sum of $k \in H$ or (if $n < 0$) $-k \in H$, must be in H. On the other hand, for any $m \in H$, we can write $m = kn + r$ with $0 \le r < k$ (division by k with remainder). Since $m, kn \in H$, it follows that $r \in H$ as well, but this forces $r = 0$ by the definition of k, so $m = kn$, as claimed.

Conversely, it is clear that $\langle k \rangle$ is a subgroup of $\mathbb{Z}$.

Part (b) is similar. If $H \subseteq \langle a \rangle$, $H \neq \{1\}$ is a subgroup, we can again define k as the smallest positive integer with $a^k \in H$. In fact, we can also show, in the same way as in part (a), that $H = \langle a^k \rangle$. I now claim that $k|t$. To see this, we again write $t = qk + r$, $0 \le r < k$. Since $a^t = 1$ and $a^k \in H$, we also have that $a^r \in H$, so $r = 0$ from the definition of k, as desired. So we can write $t = qk$ now. This means that we can list the elements of H as

$$H = \langle a^k \rangle = \{1, a^k, a^{2k}, \ldots, a^{(q-1)k}\}.$$

So $|H| = q$, and this divides t, as claimed. We also see from this that the subgroup H is completely determined by k, so two different subgroups must have distinct k's and thus also distinct q's, and thus there is exactly one subgroup of order q for each $q|t$. □

Exercise 2.29. Let $G = \langle a \rangle$ be a cyclic group of finite order $|G| = qk$. Show that the (unique) subgroup of order q consists of exactly those elements $b \in \langle a \rangle$ with $b^q = 1$.

The next result will be important later on, in the theory of fields.

Definition 2.12. Let G be a group. If there exists $n \ge 1$ such that $a^n = 1$ for all $a \in G$, then the smallest such n is called the *exponent* of G and is denoted by $\exp(G)$.

Exercise 2.30. Let G be a finite group. Show that $\exp(G)$ is the least common multiple of $o(a)$, $a \in G$.

Exercise 2.31. Can you give an example of: (a) a (necessarily infinite) group all of whose elements have finite order, but there is no $n \ge 1$ such that $a^n = 1$ for all $a \in G$; (b) an infinite group that has a finite exponent?

Theorem 2.13. *Let G be a finite abelian group. Then G is cyclic if and only if* $\exp(G) = |G|$.

Proof. Clearly, if G is finite and cyclic, say $G = \langle a \rangle$, then $a^{|G|} = 1$, but $a^n \neq 1$ for $1 \le n < |G|$. Since also $a^{n|G|} = 1$, this says that $\exp(G) = |G|$, as claimed.

The proof of the converse will be based on two lemmas that are of some independent interest.

Lemma 2.14. *Let a, b be elements of an abelian group, of finite orders $o(a) = m$, $o(b) = n$, and suppose that $(m, n) = 1$. Then $o(ab) = mn$.*

Proof of Lemma 2.14. First of all, $(ab)^{mn} = a^{mn}b^{mn} = 1$. Here, we use that a, b commute; as you showed in Exercise 2.26, this computation would not be correct in a general group. So ab has finite order, and $o(ab)|mn$, by Exercise 2.28(c).

On the other hand, if $(ab)^k = a^k b^k = 1$, then $a^k = b^{-k} \in \langle a \rangle \cap \langle b \rangle$. I now claim that

$$\langle a \rangle \cap \langle b \rangle = \{1\}. \tag{2.3}$$

This will give the Lemma because it shows that $a^k = b^{-k} = 1$, so $m|o(ab)$, $n|o(ab)$, by Exercise 2.28(c) again. Since m, n are relatively prime, it follows that $mn|o(ab)$.

It remains to establish (2.3). Let c be from the intersection, so $c = a^k = b^j$ for suitable $k, j \geq 0$. Then $c^m = a^{mk} = 1$ and similarly $c^n = 1$, so $o(c)|m$, $o(c)|n$, but since $(m,n) = 1$, this says that $o(c) = 1$ or, equivalently, $c = 1$. □

Exercise 2.32. Show that $\langle a, b \rangle = \langle ab \rangle$ in the situation of Lemma 2.14.

Exercise 2.33. Show that the statement of Lemma 2.14 can fail in non-abelian groups.

Lemma 2.15. *Let G be a finite abelian group. Then there exists an element $a \in G$ such that $o(b)|o(a)$ for all $b \in G$.*

Proof of Lemma 2.15. We will show that given $a, b \in G$, we can find a $c \in G$ so that $o(a), o(b)|o(c)$. This will give the full claim because if $G = \{g_1, \ldots, g_n\}$, then we can apply this step first to g_1, g_2 to produce an element c_1 whose order is a multiple of the orders of g_1, g_2, then we apply it to g_3, c_1 to produce c_2, then to g_4, c_2 and so on.

So write $m = o(a)$, $n = o(b)$, and factor these integers into primes:

$$m = p_1^{e_1} p_2^{e_2} \cdots p_N^{e_N}, \qquad n = p_1^{f_1} p_2^{f_2} \cdots p_N^{f_N},$$

with $e_j, f_j \geq 0$. For convenience, we can assume that these are labeled in such a way that $e_j \leq f_j$ for $j = 1, 2, \ldots, k$ and $e_j > f_j$ for $j = k+1, \ldots, N$, for some $0 \leq k \leq N$. Now let

$$r = p_1^{e_1} \cdots p_k^{e_k}, \qquad s = p_{k+1}^{f_{k+1}} \cdots p_N^{f_N}$$

(where, as usual, we interpret an empty product as 1). Then the integers m/r and n/s are relatively prime. Moreover, $o(a^r) = m/r$, $o(b^s) = n/s$, so Lemma 2.14 may be applied to these two elements, and it follows that

$$o(a^r b^s) = (m/r)(n/s) = p_1^{f_1} \cdots p_k^{f_k} p_{k+1}^{e_{k+1}} \cdots p_N^{e_N}.$$

In other words, each prime now has the larger of the two exponents on offer, and thus this number is a multiple of both m and n (in fact, it is the least common multiple), and we can take $c = a^r b^s$. □

So let's return to the proof of Theorem 2.13 now. Suppose that $\exp(G) = |G|$. Let $a \in G$ be as in Lemma 2.15. Then $b^{o(a)} = 1$ for all $b \in G$, and $o(a)$ is in fact the smallest positive integer with this property for the simple reason that no smaller integer works for $b = a$. In other words, $o(a) = \exp(G)$, so $o(a) = |G|$, and this says that the cyclic subgroup $\langle a \rangle \subseteq G$ has the same order as G, so it must in fact be all of G. □

2.4. **Cosets.** Let $H \subseteq G$ be a subgroup of a group G. A set of the form $aH = \{ab : b \in H\}$, with $a \in G$ fixed, is called a (left) *coset;* similary, Ha is called a right coset. Let's focus on left cosets for now; of course, analogous observations will apply to right cosets. I claim that the collection of all cosets aH, $a \in G$, forms a partition of G.

Obviously, every $a \in G$ is in some coset (namely, aH), so what we must show is that for any two cosets, we have either $aH = bH$ or $aH \cap bH = \emptyset$. To confirm this, suppose that $aH \cap bH \neq \emptyset$, let's say $c \in aH \cap bH$. Then $c = ah_1 = bh_2$ for suitable $h_1, h_2 \in H$. Thus $b = ah_1h_2^{-1} = ak$ with $k = h_1h_2^{-1} \in H$, since H is a subgroup. So if bh is an arbitrary element of bH, then $bh = akh \in aH$ as well because $kh \in H$ also. This says that $bH \subseteq aH$, and $aH \subseteq bH$ is shown in the same way.

This has a very important consequence:

Theorem 2.16 (Lagrange). *Let G be a finite group. Then the order of any subgroup H divides $|G|$, and we have that $|G| = |H|[G : H]$.*

Here we denote by $[G : H]$ the number of (distinct) cosets of H in G. (This number is the same for left and right cosets, so we don't need to specify which type we are considering.) This is also called the *index* of the subgroup H in G.

Proof. This is immediate from the preceding discussion: partition G into (let's say: left) cosets $G = a_1H \cup \ldots \cup a_nH$, and count elements on both sides. Notice that $n = [G : H]$, by the definition of the index, and that $|a_jH| = |H|$ (why is that true?). □

Corollary 2.17. *Let G be a finite group. Then $a^{|G|} = 1$ for all $a \in G$, and thus $\exp(G)$ is a divisor of $|G|$.*

Proof. For any $a \in G$, we can consider the cyclic subgroup $\langle a \rangle \subseteq G$. By Lagrange's Theorem, its order $o(a)$ divides $|G|$, so $a^{|G|} = a^{o(a)n} = 1$. □

Exercise 2.34. Give an explicit argument for the final claim, on $\exp(G)$.

Exercise 2.35. Use Corollary 2.17 to give a new (and extremely short) proof of Theorem 1.10. *Hint:* Recall that $(\mathbb{Z}_p^\times, \cdot)$ is a group.

Exercise 2.36. (a) Let G be a finite abelian group, and form the product $x = a_1 \dots a_n$ of all elements of G. Show that $x^2 = 1$.
(b) Prove *Wilson's Theorem:* $(p-1)! \equiv -1 \mod p$ for any prime p.
Suggestion: Use part (a) for inspiration. You will probably have to show that ± 1 are the only solutions of $x^2 \equiv 1 \mod p$, and for this, the identity $x^2 - 1 = (x+1)(x-1)$ should be useful.

Exercise 2.37. Let G be a finite group with subgroups $K \subseteq H \subseteq G$. Show that then $[G : K] = [G : H][H : K]$.

Exercise 2.38. Let H be a subgroup of G, with both G and H possibly infinite now.
(a) Show that $aH \mapsto Ha^{-1}$ sets up a bijection between left and right cosets.
(b) Conclude that if the number of left (say) cosets is finite, then so is the number of right cosets, and these two numbers agree (so we can still unambiguously define $[G : H]$ in this situation).

Exercise 2.39. Let H_1, H_2 be subgroups of G. Show that $H_1 \cap H_2$ is a subgroup, too, and the cosets satisfy $a(H_1 \cap H_2) = aH_1 \cap aH_2$.

Exercise 2.40. Suppose that $|G| = p$ is a prime. Show that G is a cyclic group.

We know from our general discussion from Section 1.2 that partitions are essentially the same thing as equivalence relations. Now we just saw that a subgroup $H \subseteq G$ produces a partition of G into cosets; in fact, it produces two partitions because here we can work with left or right cosets. What is the corresponding equivalence relation?

Proposition 2.18. *Let H be a subgroup of G and partition G into left cosets cH. Define an equivalence relation $\sim$ on G by declaring $a, b \in G$ equivalent if they lie in the same coset: $a, b \in cH$. Then $a \sim b$ if and only if $b^{-1}a \in H$.*

Exercise 2.41. Prove Proposition 2.18. Also, formulate and prove a version for right cosets.

2.5. **Congruences.** A *congruence* on a monoid or group (or any algebraic structure, for that matter) is an equivalence relation that is compatible with the algebraic structure. More specifically:

Definition 2.19. Let M be a monoid or a group. We call an equivalence relation $\equiv$ on M a *congruence* if $a \equiv a'$, $b \equiv b'$ implies that $ab \equiv a'b'$.

Recall that for any equivalence relation, we can form the new set $\overline{M} = (M/\equiv) = \{\overline{a} : a \in M\}$ of equivalence classes; these were defined as $\overline{a} = \{b : b \equiv a\}$. If our equivalence relation is a congruence, then $\overline{M}$ inherits a monoid structure from M in a natural way: we define a binary operation on $\overline{M}$ by $\overline{a}\,\overline{b} := \overline{ab}$. This is indeed well defined because no matter which representatives $a' \in \overline{a}$, $b' \in \overline{b}$ we choose, $a'b'$ will always be in the same equivalence class, so the right-hand side does not depend on an arbitrary choice of $a \in \overline{a}$, $b \in \overline{b}$.

Next, observe that this operation is associative because

$$(\overline{a}\,\overline{b})\overline{c} = \overline{ab}\,\overline{c} = \overline{(ab)c} = \overline{a(bc)} = \overline{a}\,\overline{bc} = \overline{a}(\overline{b}\,\overline{c}).$$

Similarly, $\overline{1}\,\overline{a} = \overline{1a} = \overline{a}$ and, in the same way, $\overline{a}\,\overline{1} = \overline{a}$, so $\overline{1} \in \overline{M}$ is a neutral element. We call $\overline{M}$ the *quotient monoid.*

If $M = G$ is a group, then $\overline{G} = G/\equiv$ will be a group also, which we (unsurprisingly) call the *quotient group.* We already know that $\overline{G}$ is a monoid, so we only need to show that every $\overline{a} \in \overline{G}$ is invertible, but this is obvious because $\overline{a}\,\overline{a^{-1}} = \overline{aa^{-1}} = \overline{1}$ and similarly $\overline{a^{-1}}\,\overline{a} = \overline{1}$, so $\overline{a}$ is indeed invertible, with inverse $\overline{a}^{-1} = \overline{a^{-1}}$.

Recall that equivalence relations are the same thing as partitions, so congruences can also be described in terms of the partitions they induce. In the case of groups (and all other algebraic structures we are going to discuss later on, but not for monoids), it actually suffices to know the equivalence class of 1.

Basically, we are now reversing the steps that led us to Proposition 2.18. As we will see in a moment, congruences will lead us back to cosets when we focus on the partition that is induced by the congruence, but the subgroups that are involved will now have an additional property. We need a definition:

Definition 2.20. A subgroup K of a group G is called *normal* if $aka^{-1} \in K$ for all $a \in G$, $k \in K$. We write $K \trianglelefteq G$ to indicate that K is a normal subgroup of G.

Proposition 2.21. *Let K be a subgroup of G. Then the following statements are equivalent:*
(a) $K \trianglelefteq G$;
(b) $aKa^{-1} \subseteq K$ for all $a \in G$;
(c) $aKa^{-1} = K$ for all $a \in G$;
(d) $aK = Ka$ for all $a \in G$

Here, we write $aKa^{-1} = \{aka^{-1} : k \in K\}$, and the cosets aK, Ka are defined similarly and were used earlier.

Exercise 2.42. Prove Proposition 2.21.

If G is abelian, then every subgroup is normal, but things are not so clear in general groups.

Exercise 2.43. Let H be subgroup of G of index 2. Show that H is normal.

Exercise 2.44. Consider the symmetric group S_3, and let $H \subseteq S_3$ be the (cyclic) subgroup generated by the permutation $\pi(1) = 2$, $\pi(2) = 1$, $\pi(3) = 3$.
(a) What is the order of π? List all elements of $H = \langle \pi \rangle$.
(b) Show that H is not normal in S_3.

Exercise 2.45. Let $\equiv$ be a congruence on a group G. Show that $a \equiv b$ precisely if $ab^{-1} \equiv 1$.

This already confirms what I announced above: congruences on groups are completely determined as soon as we know what is (and isn't) equivalent to 1. It pays to elaborate some more on this:

Theorem 2.22. *(a) Let $\equiv$ be a congruence on a group G. Then $K = \overline{1}$ is a normal subgroup of G, and the equivalence class of $a \in G$ is given by the coset $\overline{a} = aK = Ka$.*

(b) Conversely, if $K \trianglelefteq G$, then the relation $\equiv$ defined as $a \equiv b$ precisely if $ab^{-1} \in K$ is a congruence, and $K = \overline{1}$.

A very quick proof can be given if we introduce some additional material that is of fundamental importance anyway.

Definition 2.23. Let $\varphi : G \to G'$ be a homomorphism. The *kernel* of φ is defined as

$$\ker(\varphi) = \{a \in G : \varphi(a) = 1'\}.$$

Proposition 2.24. *Let $\varphi : G \to G'$ be a homomorphism. Then $\ker(\varphi) \trianglelefteq G$. Moreover, φ is injective precisely if $\ker(\varphi) = \{1\}$.*

Proof. Let's first check that $\ker(\varphi)$ is a subgroup: if $a, b \in \ker(\varphi)$, then $\varphi(ab^{-1}) = \varphi(a)\varphi(b)^{-1} = 1'1'^{-1} = 1'$, so $ab^{-1} \in \ker(\varphi)$ as well, as required.

To see that $\ker(\varphi)$ is normal, let $k \in \ker(\varphi)$, $a \in G$, and consider $\varphi(aka^{-1}) = \varphi(a)1'\varphi(a)^{-1} = 1'$, so $aka^{-1} \in \ker(\varphi)$, again as required.

Everything in the kernel gets mapped to the same image $1'$, so clearly $\ker(\varphi)$ cannot contain other elements (than 1) if φ is injective. Conversely, if $\ker(\varphi) = \{1\}$ and $\varphi(a) = \varphi(b)$, then $\varphi(ab^{-1}) = 1'$, so $ab^{-1} \in \ker(\varphi)$ and hence $a = b$, that is, φ is injective. □

Exercise 2.46. Recall from Exercise 2.20 (or better yet, show it again) that $\varphi(G)$ is a subgroup of G'. Is this always a normal subgroup, too?

Proposition 2.25. *Let $\equiv$ be a congruence on a group G. Then the natural map ("quotient map") $G \to \overline{G}$, $a \mapsto \overline{a}$ is a surjective homomorphism.*

Exercise 2.47. Prove this. (When you write it out, you should really find that Proposition 2.25 just restates the definition of the group operation on the quotient group.)

We are now ready for the

Proof of Theorem 2.22. (a) Let $\varphi : G \to \overline{G}$, $\varphi(a) = \overline{a}$ be the homomorphism from Proposition 2.25, and define $K = \overline{1}$, as in the statement of the Theorem. Observe that $K = \ker(\varphi)$. Indeed, $a \in \ker(\varphi)$ precisely if $\overline{a} = \overline{1}$, and this is equivalent to $a \equiv 1$ or $a \in \overline{1}$.

So K is indeed a normal subgroup, by Proposition 2.24. We already saw in Exercise 2.45 that $b \in \overline{a}$ precisely if $ba^{-1} \in \overline{1} = K$ or $b \in Ka$. Also, $Ka = aK$ since K is normal.

(b) We first check that $\equiv$, defined by $a \equiv b$ precisely if $ab^{-1} \in K$ or, equivalently, $a \in Kb$ is an equivalence relation. In fact, we already know this from the previous section because the cosets Kg, $g \in G$ form a partition and since, trivially, $b \in Kb$, we can now say that $a \equiv b$ happens precisely if a, b lie in a common set of this partition (this part does not use that K is normal, it works for arbitrary subgroups).

Now let's check that $\equiv$ is a congruence. So let $a, a', b, b' \in G$, with $a \equiv a'$, $b \equiv b'$. In other words, $aa'^{-1}, bb'^{-1} \in K$. We must check that then $ab(a'b')^{-1} \in K$. We can write

$$ab(a'b')^{-1} = abb'^{-1}a'^{-1} = a[bb'^{-1}]a^{-1}[aa'^{-1}],$$

and the products in square brackets are in K, by assumption, so $a[\ldots]a^{-1} \in K$ as well, by the normality of K, and thus we are looking at a product of two elements of K, which is in K because K is a subgroup.

The final claim is clear from observing that $a \equiv 1$ precisely if $a1^{-1} = a \in K$. □

Exercise 2.48. Give an alternative direct proof of part (a) of the Theorem that does not make use of homomorphisms.

Exercise 2.49. Show that congruences on monoids can *not* necessarily be reconstructed from $\overline{1}$, the equivalence class of the neutral element. *Suggestion:* Find two distinct congruences on $(\mathbb{N}_0, +)$ with $\overline{0} = \{0\}$.

If $K \trianglelefteq G$, we also write G/K for the quotient group. As we just saw, its elements are the cosets aK, $a \in G$; it doesn't matter which type of coset is used here because the fact that K is normal implies that $aK = Ka$. Two cosets are multiplied as follows $aK \cdot bK = (ab)K$, the neutral element of G/K is $1K = K$, and $(aK)^{-1} = a^{-1}K$.

There is an alternative interpretation of the multiplication in the quotient group G/K. We can, more generally, introduce a product of arbitrary subsets $A, B \subseteq G$ of a group G in a natural way, by setting

$$AB = \{ab : a \in A, b \in B\}. \tag{2.4}$$

It is easy to see that this product is associative: this property is just inherited from G. Or, rephrasing this slightly (and adding the observation that there is a neutral element):

Exercise 2.50. Let M be a monoid (in particular, $M = G$ could be a group). Show that the power set $P(M)$ with the product (2.4) is a monoid, too.

Observe also that if the first set has one element and the second is a subgroup, then the set product $\{a\}H = \{ah : h \in H\}$ recovers the coset; we will continue to denote this by aH.

Exercise 2.51. Show that a subset $H \subseteq G$, $H \neq \emptyset$, is a subgroup precisely if: (1) $HH \subseteq H$; (2) $H^{-1} \subseteq H$, where $H^{-1} = \{h^{-1} : h \in H\}$. Show also that in this case, $HH = H$.

Now let's return to the quotient group G/K, with $K \trianglelefteq G$. What happens if we multiply two cosets aK, bK as *sets?* By associativity of the set product, Exercise 2.51, and Proposition 2.21(d), we have that

$$(aK)(bK) = a(Kb)K = a(bK)K = (ab)(KK) = abK,$$

so we obtain the satisfying conclusion that the group operation in G/K can be viewed as set multiplication of cosets.

Exercise 2.52. Let $G = \{(a, b) : a, b \in \mathbb{R}, a \neq 0\}$, with the operation $(a, b)(c, d) = (ac, ad + b)$.
(a) Show that G is a group.
(b) Let $K = \{(1, b) : b \in \mathbb{R}\}$. Show that $K \trianglelefteq G$.
(c) What is G/K? (Find a familiar group that is isomorphic to G/K.)

Exercise 2.53. Let $H, K \trianglelefteq G$. Show that then $H \cap K$ and HK are also normal subgroups of G.

Exercise 2.54. Suppose that $K \trianglelefteq G$, $[G : K] = n$. Show that then $a^n \in K$ for all $a \in G$.

2.6. **Permutations.** In this section, we study the elements of S_n in more detail. A very useful tool is the *cycle decomposition* of a permutation. Given $\pi \in S_n$, pick an integer k_1 and keep track of how this gets moved around by π: So apply π to k_1, and let's call this $\pi(k_1) =: k_2$. Then look at $\pi(k_2) =: k_3$, then at $\pi(k_3) =: k_4$ and so on, until we return to one of these integers. In fact, this cycle can only close at k_1; otherwise, we would obtain a contradiction to the injectivity of π. Let's say $\pi(k_r) = k_1$. This permutation that maps k_j to k_{j+1} for $j = 1, \ldots, r-1$ and k_r back to k_1 and fixes the other integers from $\{1, 2, \ldots, n\}$ (if any) is called a *cycle* and is denoted by $(k_1 k_2 \ldots k_r)$. (Note that this cycle may or may not agree with the original permutation π.) For example, if $n = 4$, then (241) is the permutation π that sends $\pi(2) = 4$, $\pi(4) = 1$, $\pi(1) = 2$, $\pi(3) = 3$. Notice also that $(241) = (412) = (124)$, and of course this observation works for arbitrary cycles.

Let's return to the original permutation $\pi \in S_n$. We have extracted a cycle from this, but there may be others integers left that are not part of this cycle. If so, then pick one of these and form another cycle. Note that the integers in this new cycle will be distinct from the ones from the first cycle because everything that gets mapped to an integer from the first cycle automatically becomes part of that cycle. If $\{1, 2, \ldots, n\}$ is still not exhausted by these first two cycles, pick one the integers that is left and form a third cycle. Continue in this way until every integer belongs to a unique cycle. We have proved most of:

Proposition 2.26. *Every permutation $\pi \in S_n$ can be decomposed into disjoint cycles:*

$$\pi = (k_1 \ldots k_r)(m_1 \ldots m_s) \ldots (p_1 \ldots p_t) \tag{2.5}$$

This representation is unique except for the order in which the cycles appear (and we know that we may cyclically permute within each cycle).

The product of cycles on the right-hand side is taken in S_n; in other words, the individual cycles are composed as maps. The cycles are called disjoint here because every integers belongs to precisely one cycle. If there are k with $\pi(k) = k$, then any such integer will contribute a cycle (k) of length one. These could of course be dropped from (2.5).

Exercise 2.55. Complete the proof of the Proposition. In particular, show that two disjoint cycles commute in S_n.

Exercise 2.56. Let $\pi \in S_4$ be the permutation $\pi = (24)(413)(123)$. Find $\pi(j)$ for $j = 1, 2, 3, 4$, and give the cycle decomposition of π.

Exercise 2.57. Let $r_1, \ldots, r_m$ be the lengths of the cycles in the cycle decomposition of $\pi \in S_n$. Show that $o(\pi) = \text{lcm}(r_1, \ldots, r_m)$. Then find $\exp(S_6)$.

A *transposition* is a cycle (jk) of length 2. I claim that S_n as a group is generated by the transpositions. To show this, it is enough to verify that any cycle is in the subgroup generated by the (jk), $1 \leq j, k \leq n$, and this is immediate from the formula

$$(k_1 k_2 \dots k_r) = (k_1 k_r)(k_1 k_{r-1}) \dots (k_1 k_3)(k_1 k_2). \tag{2.6}$$

Exercise 2.58. Prove (2.6). (Don't forget that when composing maps, the ones on the right act first.)

By applying (2.6) to all cycles from the cycle decomposition, we obtain a representation of an arbitrary $\pi \in S_n$ as a product of transpositions. Unlike the cycle decomposition, this representation will not be unique. For example, $(123) = (13)(12) = (12)(23)$. Moreover, the transpositions are not disjoint, in general, and thus may not commute (what is $(12)(13)$?). However, what is determined by the permutation is the parity of the number of transpositions.

Theorem 2.27. *Let $\pi \in S_n$ be a permutation. Then either all representations of π as a product of transpositions have an odd number of factors, or they all have an even number of factors.*

We call π an *odd* or *even* permutation, according to which alternative holds. We also define the *sign* of π as $\sigma(\pi) = 1$ if π is even and $\sigma(\pi) = -1$ if π is odd.

Exercise 2.59. Show (with the help of Theorem 2.27, to be proved in a moment) that $\sigma : S_n \to \{-1, 1\}$ is a homomorphism; here, the group operation on ± 1 is multiplication.

Definition 2.28. The set of even permutations in S_n is called the *alternating group* and is denoted by A_n.

Exercise 2.60. Show (directly) that $A_n \trianglelefteq S_n$. Then give a different argument with the help of Exercise 2.59.

Proof of Theorem 2.27. Our key tool will be the number $N(\pi)$ of transpositions in the specific representation that we constructed above. More precisely, if the cycle decomposition of π has cycles of lengths $r_1, \dots, r_m$, then we set $N(\pi) = \sum (r_j - 1)$; this is motivated by (2.6), which writes a cycle of length r as a product of $r - 1$ transpositions.

I now claim that for any $\pi \in S_n$ and any transposition (jk), we have that

$$N((jk)\pi) = N(\pi) + 1 \text{ or } N((jk)\pi) = N(\pi) - 1 \tag{2.7}$$

(which case we are in will depend on the details of the situation). This will give the Theorem because if π is written as a product of T transpositions in an arbitrary way, then I can successively multiply from the left by these transpositions (which are their own inverses), and

I will eventually reach the identity element $1 \in S_n$. Clearly $N(1) = 0$, and since I got there in T steps, starting from π, and each individual step changes the parity of N, it follows that $N(\pi)$ has the same parity as T. Since I considered an *arbitrary* representation of π as a product of transpositions, it follows that these all have the same parity as $N(\pi)$, as claimed.

It remains to establish (2.7). This argument will depend on the identity

$$(jk)(jx_1 \ldots x_s k y_1 \ldots y_t) = (jx_1 \ldots x_s)(ky_1 \ldots y_t). \tag{2.8}$$

Exercise 2.61. Verify (2.8).

I will now show that, more specifically, we are in the first case in (2.7) if j, k belong to disjoint cycles of π, and we are in the second case otherwise. Let's do the second case first: suppose that the cycle decomposition of π reads

$$\pi = (jx_1 \ldots x_s k y_1 \ldots y_t) \ldots ,$$

where $\ldots$ indicates the other cycles (if any), which will act as spectators here. (Why can I put j into the first slot of this cycle?) Now (2.8) gives that $(jk)\pi = (jx_1 \ldots x_s)(ky_1 \ldots y_t) \ldots$, so the net effect of multiplying by (jk) was to split this large cycle with j, k in it into two smaller cycles. So if we now work out $N((jk)\pi)$ and compare it with $N(\pi)$, we see that the contribution $s + t + 2 - 1 = s + t + 1$ to $N(\pi)$ changes to $(s+1-1)+(t+1-1) = s+t$, and everything else in the sum defining N stays the same. Thus $N((jk)\pi) = N(\pi) - 1$, as claimed. I leave the other case to the reader. □

Exercise 2.62. Finish the proof by discussing the other case (j, k in distinct cycles) in the same style.

Exercise 2.63. (a) Show that S_n is generated by $(12), (13), \ldots, (1n)$.
(b) Show that S_n is generated by $(123 \ldots n), (12)$.
(c) Show that S_n is not generated by a single element if $n \geq 3$.

Exercise 2.64. (a) Show that every 3 cycle (jkm) is even.
(b) Show that A_n is generated by the 3 cycles.