This part of the Core A module explores a class of important algebraic objects which are sets that have a particular kind of structure. Such an object will be called a **group**. Familiar examples of groups include the integers (the "structure" referred to above being the addition of integers), the rationals (again with addition), rotations of a circle (where the structure is given by the fact that we can compose such rotations), symmetries of regular polygons (here we have a set consisting of rotations and reflections, which we can compose to give again one of this kind), non-singular $n \times n$ matrices (i.e., with non-zero determinant), the structure in the latter case being matrix multiplication. Groups occur in almost all branches of mathematics.

We are going to look at the rules that govern this structure very soon, and show how the examples named above fit the rules.

After that we shall look at some important new examples, the first resulting from properties relating to the arithmetic of the integers, particularly to division, primes and factorisation.

The second example will be something called permutations, which are ways in which a set of distinct objects can be shuffled. It turns out that any group can be viewed as such a "permutation group", hence it is important to understand the latter.

In each case we shall discover that there are unexpected properties that can be used to do novel things. For instance we shall be able to use the example from the integers to show how one kind of cryptography can be made to work.

After expanding our list of examples of groups, we shall then look at some of the more elementary properties of groups in general, including some methods of counting inside groups.

Finally, we shall study the rotation groups of regular polygons, which have an interesting and easily defined structure.

So what is a group?

Definition 1.1. *A* **group** $(G, \circ)$ *is a non-empty set* G *on which a* **binary operation** $\circ$ *is defined satisfying the following four properties*

(C) **(Closure)** $\forall g_1,\ g_2 \in G$ *we have* $g_1 \circ g_2 \in G$.
(A) **(Associativity)** $\forall g_1,\ g_2,\ g_3 \in G$, *we have* $g_1 \circ (g_2 \circ g_3) = (g_1 \circ g_2) \circ g_3$.
(N) **(Identity)** $\exists e \in G$ *such that* $\forall g \in G$, *we have* $e \circ g = g \circ e = g$. *This* e *is called the* **identity element** *or* **neutral element** *of the group.*
(I) **(Existence of inverses)** $\forall g \in G,\ \exists h \in G$ *such that* $g \circ h = h \circ g = e$. *Such a* h *is called the* **inverse** *of* g *and is often written as* g^{-1}.

Remark 1.2. *(1) $\circ$ is often (but not always) some well-known multiplication or addition. In the latter case g^{-1} is normally $-g$ and e is 0.*

(2) Associativity implies that in any expression $g_1 \circ g_2 \circ \cdots \circ g_n$ with $n \geq 2$ we don't need to write parentheses: any way of doing this will give the same element in G.

(3) If $g_1 \circ g_2 = g_2 \circ g_1$ for all g_1, g_2 in G then G is called **commutative** *(or* **abelian***).*

Example 1.3. *The integers $\mathbb{Z}$. The group operation is addition of integers. Closure and associativity are obvious. 0 is the identity element and the inverse of the integer n is $-n$. We say that the integers is* **a group under addition**.

Example 1.4. *The rationals $\mathbb{Q}$, the reals $\mathbb{R}$, and the complex numbers $\mathbb{C}$ are all groups under addition in the same way.*

Example 1.5. *The non-zero rationals $\mathbb{Q} \setminus \{0\}$, denoted $\mathbb{Q}^*$, form a group when the binary operation is multiplication of numbers. Again closure and associativity are obvious. This time 1 is the identity and the inverse of the rational $q = \frac{m}{n}$ (with $m, n \in \mathbb{Z}$) is $q^{-1} = \frac{1}{q} = \frac{n}{m}$.*

Similarly for $\mathbb{R}^*$ and $\mathbb{C}^*$.

Example 1.6. *But the non-zero* integers $\mathbb{Z} \setminus \{0\}$ *do* **not** *form a group under multiplication because there is not always an inverse. 1 is again would play the role of the identity, but while 2, say, is a non-zero integer, there is no integer n with $2 \cdot n = 1$, so 2^{-1} does not exist.*

Example 1.7. *Let $M(m, n, \mathbb{R})$ (where $m, n \geq 1$) be the set of $m \times n$ matrices with real entries. Then this is a group using matrix* **addition**. *Closure and associativity are obvious, as we only need to check them for each component individually (where it is reduced to the associativity of the addition in $\mathbb{R}$). The identity is the zero matrix and the inverse of the matrix with ij-th entry a_{ij} is the one with ij-th entry $-a_{ij}$.*

Example 1.8. *What if we wanted to use matrix multiplication? First of all we would require $m = n$ or we would not even have multiplication defined. Then closure and associativity would certainly work: for matrices $A = (a_{ij})$, $B = (b_{jk})$ and $C = (c_{k\ell})$ we have the identity*

$$\sum_{k=1}^{n} \left(\sum_{j=1}^{n} a_{ij} b_{jk} \right) c_{k\ell} = \sum_{j=1}^{n} a_{ij} \left(\sum_{j=1}^{n} b_{jk} c_{k\ell} \right).$$

We can easily find an identity matrix I_n with 1s down the diagonal and 0s elsewhere. But inverses would fail. For instance if O is the zero matrix then we can never find an inverse A such that $AO = I_n$, since we always have $AO = O$.

In order to get a group with matrix multiplication, we have to restrict attention to those which have non-zero determinant; these are called the **non-singular matrices**. *Then if A is an*

$n \times n$ matrix with $\det(A) \neq 0$, i.e. non-zero determinant, we can always find an inverse B with $AB = BA = I_n$.

The group $\{A \in M(n, n, \mathbb{R}) \mid \det(A) \neq 0\}$ is called the **general linear group of $n \times n$ matrices over the reals** *and is written $GL(n, \mathbb{R})$. A similar construction works if we use $\mathbb{Q}$ or $\mathbb{C}$ instead of $\mathbb{R}$.*

We can make one further important **observation** *about the general linear group: the order in which we multiply matrices matters. E.g., if we take $n = 2$ and put*

$$A = \begin{pmatrix} 1 & 2 \\ -2 & 3 \end{pmatrix}, \; B = \begin{pmatrix} 2 & -1 \\ 0 & 1 \end{pmatrix},$$

then

$$AB = \begin{pmatrix} 2 & 1 \\ -4 & 5 \end{pmatrix} \neq \begin{pmatrix} 4 & 1 \\ -2 & 3 \end{pmatrix} = BA.$$

This is saying that the general linear group is **not commutative**. *So while almost all of the initial examples we shall see* **are** *commutative, for a general group this need not be the case and you must not assume it is unless you are told otherwise.*

Example 1.9. *Let $n > 1$ be an integer and let $C_n = \{\exp(\frac{2k\pi i}{n}) | \, 0 \leq k \leq n-1\}$. These are complex numbers on the unit circle equally spaced around it. Let the group operation be multiplication (of complex numbers), which here gives*

$$\exp\left(\frac{2k\pi i}{n}\right) \exp\left(\frac{2l\pi i}{n}\right) = \exp\left(\frac{2(k+l)\pi i}{n}\right).$$

We know that we can always ignore excess multiples of $1 = \exp(2\pi i) = \exp(\frac{2n\pi i}{n})$. So if we calculate $k + l$ and get an integer greater than $n - 1$ we can simply subtract n and get a value in the original range.

So we have closure. Associativity is obvious, 1 is the identity, which we obtain for $k = 0$, and the inverse to $\exp(\frac{2k\pi i}{n})$ is $\exp(\frac{2(n-k)\pi i}{n})$.

So we have a group. Note that $\exp(\frac{2k\pi i}{n}) = (\exp(\frac{2\pi i}{n}))^k$, so every element is a power of $\exp(\frac{2\pi i}{n})$. (Later on we shall restate this by saying that $\exp(\frac{2\pi i}{n})$ **generates** *the group C_n and that C_n is a* **cyclic group of order** *n. The* **order of a group** *means the number of elements in it.)*

Example 1.10. *Let $G = \{1, -1\}$ and use normal multiplication of integers. Then this is a group with identity 1 and $(-1)^{-1} = -1$.*

Example 1.11. *Let $H = \{ODD, EVEN\}$ with an addition rule that makes*

$$EVEN + EVEN = EVEN, \quad EVEN + ODD = ODD$$
$$ODD + EVEN = ODD, \quad ODD + ODD = EVEN$$

Then again we have a group where the identity is $EVEN$, and ODD is self-inverse.

Now look at the last two examples and make the associations

$$1 \leftrightarrow EVEN, \qquad -1 \leftrightarrow ODD$$

then after a little thought you should be able to see that these two groups have the same underlying structure.

Example 1.12. *The set of symmetries of a non-square rectangle $ABCD$:*

A little thought shows that there are only 4 possibilities:

I: the identity permutation that does nothing.

H: the reflection about the horizontal axis of symmetry.

V: the reflection about the vertical axis of symmetry.

R: rotation through $180°$.

We can then work out the group table for this which looks like:

$\circ$	I	H	V	R
I	I	H	V	R
H	H	I	R	V
V	V	R	I	H
R	R	V	H	I

This means that if we pick g from the leftmost column, and h from the topmost row, then the slot indicated by those choices contains $g \circ h$. [In this case $g \circ h = h \circ g$ because this group is commutative but in general this will not be the case.]

The point to note is that the square of every element is the identity, so every element is self-inverse. This means that the structure of the group is very different from the cyclic group $C_4 = \{1, -1, i, -i\}$ of Example 1.9 with $n = 4$, since $i^2 = -1 \neq 1$ and you need to raise i to the power 4 to get the identity element.

Notational Conventions

In our original abstract definition of a group, we used the symbol $\circ$ to denote the group operation and so we wrote things like $g_1 \circ g_2$. In practice we do the group operation by *juxtaposition* and so write $g_1 g_2$. We also tend to talk about *group multiplication*.

The important thing to realise is that this is just a convenient way of writing things in an abstract situation. When we get to concrete examples, the group multiplication may in fact be regular addition (e.g., in the integers $\mathbb{Z}$) or regular multiplication (e.g., of matrices in $GL(3,\mathbb{Q})$, or of elements in $\mathbb{R}^*$) or it may be composition of symmetries as in our rectangle example.

So in concrete examples we will use the group operation that is appropriate and say things like "the integers form a group under addition", or "non-singular $n \times n$ matrices form a group under matrix multiplication" or "the symmetries of a rectangle form a group under composition of symmetries".

But when doing things abstractly we shall from now on always use "multiplication" using juxtaposition.

Here are a couple of little examples to show how we can manipulate things abstractly.

Proposition 1.13. (1) *The identity element in a group G is unique.*

(2) *If g is an element of the group G (written $g \in G$), then it has a unique inverse element, written g^{-1}.*

(3) $(g^{-1})^{-1} = g$.

(4) $(gh)^{-1} = h^{-1}g^{-1}$ *(note the order).*

Proof. (1) Suppose there are two elements $e,\ f$ that both act as identity elements in G. Then, since e is an identity, $ef = f$ and since f is an identity $ef = e$. So $e = f$.

(2) Suppose that there are two elements $h,\ k$ which both act as inverses for g, and let e be the identity element in G. Then

$$h \overset{(N)}{=} he \overset{(I)}{=} h(gk) \overset{(A)}{=} (hg)k \overset{(I)}{=} ek \overset{(N)}{=} k$$

So $h = k$.

(3) From the previous part we only have to check that g satisfies the defining properties of $(g^{-1})^{-1}$: $g^{-1}g = e = gg^{-1}$, which hold by definition of g^{-1}.

(4) Similarly we check that $h^{-1}g^{-1}$ satisfies the defining properties of $(gh)^{-1}$, which we leave as an exercise.

□

Note how we have used the various group axioms to make these calculations work.

The first two parts of this result mean that we can talk about *the* identity of a group, and genuinely use the term g^{-1} for the inverse element of g without any ambiguity.

Another notational convention, is that when we are doing abstract work with groups we shall always use the letter e for the identity. If we have more than one group in play, then we would tend to use a subscript to distinguish them. Thus if G, H are two groups, we might use e_G for the identity in G and e_H for the identity in H. But if there is no confusion, we would not bother with a subscript.

So now we have a definition of a group and some examples. We shall now start to develop new examples and see how they help our understanding of other mathematical concepts.

2. Numbers

We shall be dealing mainly with the set of integers, $\mathbb{Z}$. A lot of the time we shall be using familiar ideas of addition, subtraction, multiplication and division. Here is a simple start.

Theorem 2.1. *Let n, m be integers with $m > 0$. Then there exist unique integers q (the* **quotient***) and r (the* **remainder***) such that $n = qm + r$ and $0 \leq r < m$.*

I don't intend to prove this, but if you want to see a proof refer to Whitehead [*Guide*^{2} *Abstract Algebra*].

When $r = 0$ we say that m **divides** n and write $m|n$. More generally we say, for m and n in $\mathbb{Z}$, that $m|n$ if there is some q in $\mathbb{Z}$ with $n = qm$.

Lemma 2.2. (1) *for all n in $\mathbb{Z}$, $1|n$;*
(2) *for all n in $\mathbb{Z}$, $n|n$;*
(3) *if $l|m$ and $m|n$ then $l|n$ ("transitivity");*
(4) *for all n in $\mathbb{Z}$, $n|0$;*
(5) *if $n \neq 0$ then 0 does not divide n;*
(6) *if $n|a$ and $n|b$ then $n|(a + b)$;*
(7) *if $n|a$ then $n|ma$ for any m in $\mathbb{Z}$;*
(8) *$n|a$ and $n|b$ implies that $n|(xa + by)$ for all integers x and y.*

Proving those properties from the definition should pose no serious problem and we leave it as an exercise.

To start with we are going to concentrate on **division**, because integers have the particular property that they can often be factorised into smaller ones. For example, $24 = 2{\cdot}2{\cdot}2{\cdot}3$, $399 = 3 \cdot 7 \cdot 19$, $1001 = 7 \cdot 13 \cdot 11$. These factorisations have been chosen to break down the numbers into their basic building blocks; i.e., into their **prime factors**.

Definition 2.3. *A* **prime number** *is a positive integer that is divisible by* **exactly two different positive integers**.

Note the care that we take. It is not enough to say that a prime is only divisible by 1 and itself, because this would include the number 1. **But 1 is not declared to be a prime number** for a reason that will become clear very soon.

It is not hard to start listing the prime numbers in ascending order:

$$2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53, \ldots$$

But there is a problem in general in knowing how to identify which numbers are prime.

Lemma 2.4. *Let $n > 1$. Then n is not prime if and only if there exists a prime $p \leq \sqrt{n}$ such that $p|n$.*

Proof. If n is not prime, it has divisors other than 1 and itself. Every such divisor is less than n. Let p be the least divisor of n that is greater than 1. Then p must be prime, since if it had a divisor other than 1 and p, say k, then we would have $k > 1$, $k|p$ and $p|n$, hence $k|n$ by transitivity of $|$, but this would contradict our choice of p.

Finally as $p|n$ we can find q with $n = pq$. The construction then implies that $p \leq \sqrt{n}$ since otherwise q would provide another smaller non-trivial factor of n.

We leave the proof of the converse as an exercise. □

This gives the following **Test for Primeness: Given a potential prime n, divide n successively by all the integers k, $2 \leq k \leq \sqrt{n}$. If none divide exactly, then n is prime**.

In fact, you do things in ascending order and only do trial division with primes, since if it fails with a given prime, then it certainly fails for any multiple of that prime.

All this is leading us to the following fundamental result.

Theorem 2.5. (Fundamental Theorem of Arithmetic)*(FTA) If n is an integer with $|n| > 1$ then we can write*

$$n = \varepsilon p_1^{r_1} p_2^{r_2} \dots p_k^{r_k}$$

uniquely, where $\varepsilon = \pm 1$ is the sign of n (positive or negative), $k \geq 1$, $p_1 < p_2 < \dots < p_k$ are primes and $\forall i,\ r_i \geq 1$.

Sketch of proof. First note that we have excluded the special cases $n = 0,\ 1,\ -1$. The sign is clearly well-defined and so can be put to one side. So assume $n > 1$.

We have to show both the existence of a prime factorisation and its uniqueness. Existence follows by induction starting with $n = 2$ which we know to be prime. For general n then either n is prime and $k = 1,\ p_1 = n,\ r_1 = 1$, or n is composite with $n = xy$ where both x and y are less than n. We then get prime factorisations for both x and y by induction on n and combine them together to get (after rearranging the order of the factors if necessary) an expression for n of the required type.

Uniqueness is rather more complicated. If you want to see a proof again refer to Whitehead. □

Once we have this result, then another one follows quite quickly.

Theorem 2.6. *There are infinitely many prime numbers.*

Proof. Suppose not. Then there are only finitely many and we can write the total list of them as $0 < p_1 < p_2 < \cdots < p_t$ for some positive integer t.

Form the integer $n = p_1 p_2 \cdots p_t + 1$. Then by the FTA we can factorise n uniquely as a product of primes. But $p_1, \ldots, p_t$ do not appear in this factorisation: when we divide n by p_i we get remainder 1. Thus none of the primes dividing n appear in our supposedly exhaustive list and we have a contradiction. □

Remark 2.7. *We can use Theorem 2.5 to define greatest common divisors (*gcd*'s) and least common multiples (*lcm*'s), for any non-zero numbers a and b. Namely, by allowing $r_i = 0$ in the factorisation in Theorem 2.5 we can write $a = \varepsilon_1 p_1^{r_1} \cdots p_k^{r_k}$ and $b = \varepsilon_2 p_1^{s_1} \cdots p_k^{s_k}$ for the same primes $p_1 < \cdots < p_k$. Then $\gcd(a,b) = p_1^{\min\{r_1,s_1\}} \cdots p_k^{\min\{r_k,s_k\}}$ and $\operatorname{lcm}(a,b) = p_1^{\max\{r_1,s_1\}} \cdots p_k^{\max\{r_k,s_k\}}$. This easily generalizes to $\gcd(a_1, \ldots, a_n)$ and $\operatorname{lcm}(a_1, \ldots, a_n)$ (how?).*

To apply this one would have to write a and b as a product of prime factors, but in general this is quite difficult. The simplest test we have thus far is to divide n by everything up to $\sqrt{n}$. This may sound not too bad, but if we are presented with a number with 200 decimal digits, then the square root has around 100 decimal digits and it takes a long time to count that far! Such a problem is not fanciful as it lies behind some of the algorithms that are used to make secure connections over the internet.

While it may be hard to factor a number (unless all of its factors are small), we can find common factors of two numbers very easily.

Example 2.8. *Consider 336 and 231. We are going to find the greatest integer that divides both 336 and 231. Look at the following sequence of equations:*

$$336 = 1 \cdot 231 + 105$$
$$231 = 2 \cdot 105 + 21$$
$$105 = 5 \cdot 21 \ + 0$$

Let d be a common divisor of both 336 and 231. Then using Lemma 2.2 we see that d also divides 105 and then that it also divides 21. But $21|105$ by the last line and hence, working backwards, also $21|231$ and also $21|336$. So 21 is a common divisor, but any such also divides 21, so 21 is the **greatest common divisor** *of 336 and 231.*

This is our first example of a particular process which is an *algorithm* for finding gcd's.

The Euclidean algorithm

This algorithm starts with two positive integers m, n and results in the greatest common divisor of them both.

(1) If $m > n$ then swap the two numbers around so that $n > m$.
(2) Use Theorem 2.1 to write $n = q \cdot m + r$ where $0 \leq r < m$.
(3) If $r = 0$, then output m as the gcd and *STOP*.
(4) Otherwise, replace n by the current value of m and m by the current value of r, and goto Step (2).

Theorem 2.9. *The Euclidean algorithm does indeed produce the greatest common divisor of two positive integers.*

Proof. The proof is exactly as we described the process with the example. Let d be any common divisor, then $d|m$ and $d|n$ and so $d|(n - qm) = r$. So at each stage the same divisor d divides each of $m,\ n,\ r$ until eventually $r = 0$ and the current value of m is our output number. Conversely, starting with this final output, we can work back through the steps in the same way to ensure that the candidate for the g.c.d. does in fact divide our original two numbers.

One final observation: we can be sure that the process does stop, because we always have $m > r$. This means that the pair of numbers being tested in Step (2) is getting steadily smaller at every cycle of the process. □

We can in fact deduce the following result from the algorithm:

Corollary 2.10. *Given two positive numbers $m,\ n$ with greatest common divisor d, we can find integers $x,\ y$ with $d = xm + yn$.*

Proof. We claim that, at each stage of the algorithm, we can always write each of the current values of m, n and the remainder r as a linear combination of the original pair m, n. The algorithm lets us assume that $n > m$. The first step is to solve $n = qm + r$, so that $r = (-q) \cdot m + n \cdot 1$, $m = m \cdot 1 + n \cdot 0$ and $n = m \cdot 0 + n \cdot 1$ and our claim is true after the first iteration of the algorithm.

Suppose that, putting $n_1 = n$, $m_1 = m$, $q_1 = q$ and $r_1 = r$, we have $n_k = q_k m_k + r_k$ at the k-th stage. Then $n_k = m_{k-1}$ and $m_k = r_{k-1}$, so both these are linear combinations of $m,\ n$. Then $r_k = -q_k m_k + n_k$ is as well.

Thus the property is true at every stage including the final one where the remainder is 0 and d is the final value of m. □

Example 2.11. *Consider the pair 336, 231 again together with our previous calculations. Then*

$$105 = 1 \cdot 336 - 1 \cdot 231$$

$$21 = 1 \cdot 231 - 2 \cdot 105 = 1 \cdot 231 - 2(336 - 231) = 3 \cdot 231 - 2 \cdot 336$$

But at this stage we know that 21 is going to be the g.c.d. and so we have the required formula.

Example 2.12. *Find* $\gcd(98, 77)$ *and express it as a linear combination of 98 and 77.*

$$\begin{aligned} 98 &= 77 + 21 && \Rightarrow 21 = 1 \cdot 98 - 1 \cdot 77 && \\ 77 &= 3 \cdot 21 + 14 && \Rightarrow 14 = 77 - 3(98 - 77) && = 4 \cdot 77 - 3 \cdot 98 \\ 21 &= 14 + 7 && \Rightarrow 7 = (98 - 77) - (4 \cdot 77 - 3 \cdot 98) && = 4 \cdot 98 - 5 \cdot 77 \end{aligned}$$

and 7 is clearly the greatest common divisor here. The key to successful computation here is to ensure that you are tidy and organised with your bookkeeping. Keep things in clear columns so that you can look back through the stages of the computation and copy down things correctly.

Put Theorem 2.9 and Corollary 2.10 together and you get what is called the **Extended Euclidean algorithm**.

An application: finding roots of polynomials

A **polynomial** in the variable x is an object that looks like

$$a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 = \sum_{r=0}^{n} a_r x^r$$

The a_r are the coefficients of the polynomial and are taken to be in some sensible object like the integers, rational, reals or complex numbers. By convention we tend to assume that the highest coefficient (also called the **leading coefficient**) $a_n \neq 0$, although there is the obvious exception where $n = 0$ and $a_0 = 0$ (this is the **zero polynomial**).

Assuming $a_n \neq 0$, then the **degree** of the polynomial is n. A constant $c \neq 0$ is a polynomial of degree 0. The zero polynomial has (by convention) degree $-\infty$. We can do normal arithmetic and algebra with polynomials: addition, subtraction and multiplication.

The set of polynomials over the integers in the variable x is denoted $\mathbb{Z}[x]$, over the rationals $\mathbb{Q}[x]$, over the reals $\mathbb{R}[x]$ etc.

The variable does not have to be called x, it can be y (in which case we would talk about $\mathbb{Z}[y]$) or any other symbol that is convenient. A polynomial in x might be denoted by $f(x)$ or some such (i.e., like a function in x).

We have one result that ought to be obvious:

Lemma 2.13. *If $f(x)$ and $g(x)$ are polynomials of degree m and n respectively, then the product polynomial $f(x)g(x)$ has degree $m+n$. The sum $f(x)+g(x)$ has degree $\leq \max(m,n)$.*

In general we cannot divide two polynomials, but there is one situation where we can always achieve a result of sorts. For this we need the idea of a **monic polynomial** which is one where the leading coefficient is 1.

Lemma 2.14. *Let $f(x)$, $g(x)$ be two polynomials where $g(x)$ is monic, then we can find unique polynomials $q(x)$, $r(x)$ such that*

$$f(x) = q(x)g(x) + r(x)$$

where the degree of $r(x)$ is less than the degree of $g(x)$.

I leave this as a fairly straightforward exercise involving nothing more than long division.

Lastly we want to know what a root of a polynomial is. First of all we can take a polynomial $f(x)$ and **evaluate** it for a given constant c, say. We do this simply by substituting c for x and working out the result which we call, naturally, $f(c)$.

A **root of the polynomial** $f(x)$ is a constant c such that $f(c) = 0$. The following should be well known to you.

Theorem 2.15. Remainder Theorem *If c is a root of the polynomial $f(x)$ then we can write $f(x) = (x-c)g(x)$ where g is a polynomial of degree one less than f.*

Proof. Use the previous Lemma. Then we can write $f(x) = (x-c)g(x) + r(x)$ where $r(x)$ has degree less than the degree of $(x-c)$. Thus $r(x)$ has degree less than 1 and so $r(x) = d$ for some constant d.

Now substitute $x = c$ and we get $0 = f(c) = (c-c)g(c) + d$, so $d = 0$. □

We can use certain coefficients of the polynomial to spot rational roots when the coefficients are integers or rationals. Let's start this way:

Lemma 2.16.
(1) *Any non-zero polynomial $f(x) \in \mathbb{Q}[x]$ (or $\mathbb{Z}[x]$) is naturally equivalent to a unique monic polynomial $m(x) \in \mathbb{Q}[x]$ with the same roots.*
(2) *Any non-zero polynomial $f(x) \in \mathbb{Q}[x]$ is naturally equivalent to a polynomial $g(x) \in \mathbb{Z}[x]$ with the same roots where the g.c.d of all the coefficients of g (taken together) is 1.*

Proof. (1) Suppose $f(x) = \sum_{r=0}^{n} a_r x^r$ (with $a_n \neq 0$), then define $m(x) = \frac{1}{a_n} f(x)$.

(2) First of all create the monic polynomial $m(x)$ as in (1) above. Then take all the rational coefficients in $m(x)$ and write them with numerators and denominators as integers in smallest form (so the g.c.d. of any numerator and the corresponding denominator is 1). Let d be the least common multiple of all these denominators and define $g(x) = d.m(x)$.

Then any common divisor of all the coefficients of g will divide the leading coefficient, which is d. The largest prime power in this common divisor will then be present in the denominator of some coefficient of $m(x)$ and so the corresponding numerator will have no such prime in it. But multiplying by d will cancel out all vestige of that prime in the corresponding coefficient of g. So the only common divisor is 1. □

And finally:

Theorem 2.17. *Let $f(x) = \sum_{r=0}^{n} a_r x^r \in \mathbb{Z}[x]$, $n \geq 1$, $a_n \neq 0$, then any rational root θ of f has the form $\theta = \frac{s}{t}$ where $s|a_0$ and $t|a_n$.*

Proof. If θ is a root of $f(x) = \sum_{r=0}^{n} a_r x^r$, then $\theta = s/t$ for integers s and t, with $t \neq 0$ and $\gcd(s,t) = 1$, so $\sum_{r=0}^{n} a_r (\frac{s}{t})^r = 0$. Multiplying through by t^n then gives

$$\begin{aligned} a_n s^n + a_{n-1} s^{n-1} t + \cdots + a_1 s t^{n-1} + a_0 t^n &= 0 \\ \Rightarrow a_0 t^n &= -(a_1 s t^{n-1} + \cdots + a_n s^n) \\ &= (-s)(a_1 t^{n-1} + \cdots + a_n s^{n-1}) \end{aligned}$$

Hence $s|a_0 t^n$, but since $\gcd(s,t) = 1$, we must have $s|a_0$. The proof that $t|a_n$ is very similar and is left as an exercise. □

Example 2.18. *Find the roots of $6x^3 + x^2 - 5x - 2$. The theorem tells us that the list of possible roots is: $\pm 1,\ \pm 2,\ \pm\frac{1}{2},\ \pm\frac{1}{3},\ \pm\frac{2}{3},\ \pm\frac{1}{6}$.*

A quick inspection shows that $x = 1$ is a root and then that

$$6x^3 + x^2 - 5x - 2 = (x-1)(6x^2 + 13x + 6)\,.$$

The quadratic is easy to factor and the final result is

$$(x-1)(2x+3)(3x+2)\,.$$

so the roots are $1,\ -\frac{2}{3},\ -\frac{3}{2}$.

Example 2.19. *The polynomial $f(x) = 2x^3 - 2x + 1$ has no rational roots.*

The theorem tells us that the only possible roots are $\pm 1,\ \pm\frac{1}{2}$. But

$$f(1) = 1,\ f(-1) = 1,\ f(\frac{1}{2}) = \frac{1}{4},\ f(-\frac{1}{2}) = \frac{7}{4}\,.$$

3. Modular Arithmetic

Before we define what modular arithmetic is all about, consider the following two examples:

Example 3.1. *All integers are either even or odd. Even if we do not know which integers we are talking about we know that:*

$$even + even = even,\ even + odd = odd,\ odd + odd = even,$$

$$even \times even = even,\ even \times odd = even,\ odd \times odd = odd.$$

So we can do basic arithmetic using only the concept of odd and even.

Example 3.2. *We know there are infinitely many prime numbers, but we can show more than that:*

Theorem 3.3. *There are infinitely many prime numbers of the form $4k + 3$.*

Proof. Suppose this is false, then there are only finitely many and there is a largest one, n, say (and note that $n \geq 3$).

Think about the number $N = (4n)! - 1 = 4m - 1$ for some integer m. We know that this has a factorisation as a product of prime numbers. It is odd, so all the primes concerned are odd. Now every p_i of our original list of primes of the form $4k + 3$ gives remainder -1 (or rather $p_i - 1$) when we divide into N, so none is a factor.

So all the prime factors of N have the form $4l + 1$ for various numbers l. But it is an easy exercise to show that when you multiply two such numbers together, you get another number of the same form, which implies that this prime product has the form $4m' + 1$ and not $4m - 1$, which is a contradiction. □

Remark 3.4. *It is also true that there are infinitely many primes of the form $4k+1$, but this is very advanced mathematics and well beyond the first or second year of an undergraduate course.*

The first example is all about classifying integers by "what is the remainder when you divide by 2?". The second is about "what is the remainder when you divide by 4?"

Choose a positive integer $n \geq 1$ (the case $n = 1$ is boring but sometimes useful). Then **Arithmetic modulo** n is all about "handling remainders when you divide by n". Effectively, all multiples of n become invisible.

Definition 3.5. *Two integers $a,\ b$ are* **congruent modulo** *n if $n|(a - b)$. We write this as*

$$a \equiv b \pmod{n}.$$

[*For those of you that did Core B2 last term, this is an* **equivalence relation**, *and what we shall be doing is studying the equivalence* **classes** *defined by this relation.*]

The next step is to define sets that represent just the remainders. Remember that n is fixed.

Let a be any integer and define

$$\bar{a} = \{a + kn | k \in \mathbb{Z}\}\,.$$

This is the **congruence class** of a. (Some authors use $[a]$ rather than $\bar{a}$; this is the case in Whitehead's book.) Note that $\bar{a}$ determines a only up to adding a multiple of n.

Let us look at what this means when $n = 5$.

	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	0	1	2	3	4	5	6	7	8	9	10	11
$\bar{0}$	*					*					*					*					*	
$\bar{1}$		*					*					*					*					*
$\bar{2}$			*					*					*					*				
$\bar{3}$				*					*					*					*			
$\bar{4}$					*					*					*					*		
$\bar{5}$	*					*					*					*					*	
$\bar{6}$		*					*					*					*					*
$\bar{7}$			*					*					*					*				
$\bar{8}$				*					*					*					*			
$\bar{9}$					*					*					*					*		

So you can see how the pattern repeats itself, and in fact there are only really five different congruence classes and we can choose them most sensibly by looking at the natural remainders modulo 5; i.e., $\bar{0},\ \bar{1},\ \bar{2},\ \bar{3},\ \bar{4}$.

These are called the **residues** modulo 5.

We can define addition and multiplication on residues by putting $\bar{a} + \bar{b} = \overline{a + b}$ and $\bar{a}\,\bar{b} = \overline{ab}$, but we have to check those operations are *well-defined*: we *choose* a and b in $\bar{a}$ and $\bar{b}$ and we must make sure that the result is independent of those choices. But if we had chosen a' and b' in $\bar{a}$ and $\bar{b}$ instead, then $a' = a + kn$ and $b' = b + ln$ for some integers k and l, and $a' + b' = a + b + (k + l)n$ so that $\overline{a + b} = \overline{a' + b'}$. Similarly $a'b' - ab$ is a multiple of n (check this!) so that $\overline{ab} = \overline{a'b'}$.

For example, when $n = 5$,

$$\bar{1} + \bar{3} = \bar{4}$$
$$\bar{2} + \bar{4} = \bar{6} = \bar{1}$$
$$\bar{2} - \bar{4} = \overline{-2} = \bar{3}$$
$$\bar{2} \times \bar{3} = \bar{6} = \bar{1}$$
$$\bar{3} \times \bar{4} = \overline{12} = \bar{2}$$

Note that in the first line we could also have computed $\bar{1} + \bar{3} = \overline{6+8} = \overline{14} = \bar{4}$, where we picked 6 in $\bar{1}$ and 8 in $\bar{3}$. And in the last we could have used 8 and 9 in $\bar{3}$ and $\bar{4}$ and found $\overline{8 \times 9} = \overline{72} = \bar{2}$.

We can write down addition and multiplication tables:

$+$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$
$\bar{0}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$
$\bar{1}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$	$\bar{0}$
$\bar{2}$	$\bar{2}$	$\bar{3}$	$\bar{4}$	$\bar{0}$	$\bar{1}$
$\bar{3}$	$\bar{3}$	$\bar{4}$	$\bar{0}$	$\bar{1}$	$\bar{2}$
$\bar{4}$	$\bar{4}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$

$\times$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$
$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$
$\bar{1}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$
$\bar{2}$	$\bar{0}$	$\bar{2}$	$\bar{4}$	$\bar{1}$	$\bar{3}$
$\bar{3}$	$\bar{0}$	$\bar{3}$	$\bar{1}$	$\bar{4}$	$\bar{2}$
$\bar{4}$	$\bar{0}$	$\bar{4}$	$\bar{3}$	$\bar{2}$	$\bar{1}$

Now $n = 5$ is just an example, but the general principle works for any value of n. This is summed up in the following

Proposition 3.6. *Let n be a positive integer. Then $\bar{a} = \bar{b}$ if and only if $n|(a-b)$. There are n distinct congruence classes modulo n, namely $\bar{0}, \bar{1}, \ldots, \overline{n-1}$.*

The set of equivalence classes modulo n is called **the integers modulo** n and is denoted $\mathbb{Z}_n$ or $\mathbb{Z}/n\mathbb{Z}$. We shall tend to use the former, although the syllabus on the Maths Department web page uses the latter.

We already defined addition and multiplication. We can similarly define subtraction by $\bar{a} - \bar{b} = \overline{a-b}$, and everything works exactly the way we should think it does. We can even multiply a congruence class by a $k \in \mathbb{Z}$, giving $k\bar{a} = \overline{ka}$, and everything makes sense and works according to all the normal rules of arithmetic (although the results can look strange at first; e.g., $3 + 3 \equiv 0 \pmod 6$).

But we cannot do division – at least, in general we can't.

One of the major rules of normal arithmetic is that the product of two non-zero things is never zero. NOT SO IN MODULAR ARITHMETIC. E.g., $2 \times 4 \equiv 0 \pmod 8$. So if we could divide by $\bar{2}$ then $\bar{2}\,\bar{0} = \bar{0} = \bar{2}\,\bar{4}$ would give $\bar{0} = \bar{4}$, which is not the case in $\mathbb{Z}_8$.

Similarly, we cannot always solve equations. E.g., is there a solution to

$$7x \equiv 17 \pmod{35} \ ?$$

The answer is no, because the only possible values for $\overline{7x}$ are $\bar{0},\ \bar{7},\ \overline{14},\ \overline{21},\ \overline{28}$.

So, what sort of structure does $\mathbb{Z}_n$ have? Here are some facts.

$\bar{0}$ always acts like the number 0 acts in the integers. We always have $\bar{0}\bar{a} = \bar{0}$ and $\bar{0} + \bar{a} = \bar{a}$.

In fact we have a group under addition.

Proposition 3.7. *$\mathbb{Z}_n$ is a group under addition.*

Proof. Closure is obvious and we have an identity $\bar{0}$.

We also have inverses since $\overline{-a} + \bar{a} = \bar{0} = \bar{a} + \overline{-a}$. (Note that $\overline{-a} = \overline{n-a}$.)

Associativity follows from associativity of the addition in $\mathbb{Z}$.

□

Note that $\mathbb{Z}_n$ has two operations (addition and multiplication) but only addition is used here. If we try to use multiplication then $\bar{1}\bar{a} = \bar{a} = \bar{a}\bar{1}$ for all $\bar{a}$, but we cannot always solve $\bar{a}\bar{b} = \bar{1}$ so inverses don't always exist. So we do not always get a group under multiplication.

However, we can solve linear equations when we work *modulo a prime number.*

Proposition 3.8. *Let p be prime and let $\bar{a} \neq \bar{0}$ in $\mathbb{Z}_p$, then*

(1) *We can find $\bar{b}$ such that $\bar{a}\bar{b} = \bar{1}$.*
(2) *For any $\bar{c} \in \mathbb{Z}_p$ we can find $\bar{x}$ with $\bar{a}\bar{x} = \bar{c}$*
(3) *If p is prime, then $\mathbb{Z}_p - \{\bar{0}\}$ is a group under multiplication.*

Proof. (1) As p is prime and $\bar{a} \neq \bar{0} \pmod{p}$, then $\gcd(a, p) = 1$ and so we can find $b,\ t$ with $ab + pt = 1$. But

$$1 = ab + pt \equiv ab \pmod{p} \Rightarrow\ \bar{a}\bar{b} = \bar{1} \text{ in } \mathbb{Z}_p\,,$$

and $\bar{a}\bar{b} = \bar{b}\bar{a}$ always.

(2) From (1) we find: $\bar{c} = \bar{c}\bar{1} = \bar{c}\bar{a}\bar{b} = \bar{a}\bar{c}\bar{b} = \bar{a}\overline{cb}$. Hence $\bar{x} = \overline{cb}$ works.

(3) Associativity follows from $(ab)c = a(bc)$ in $\mathbb{Z}$, while $\bar{1}$ is the identity and from part (2) we have inverses. So we just need closure.

For this, let $\bar{a},\ \bar{b}$ be non-zero residues mod p. Neither a nor b is divisible by p and since p is prime it cannot divide ab either. Hence $\overline{ab} \neq \bar{0}$ and we have closure.

□

So greatest common divisors are starting to appear. Indeed, it is worth stating the really key result now.

Lemma 3.9. *Let $n \geq 2$, and a such that $0 < a < n$ and $\gcd(a, n) = 1$. Then there exists a unique integer $0 < b < n$ such that $ab \equiv 1 \pmod{n}$.*

Proof. As $\gcd(a, n) = 1$, by Corollary 2.10 we can find $x,\ y$ such that $ax + ny = 1$. Simply select b in the range that $0, 1, \ldots, n-1$, so that $b \equiv x \pmod{n}$ and then $ab \equiv ax = 1 - ny \equiv 1 \pmod{n}$. Then note that $n \geq 2$ implies that $b \neq 0$.

Now b is unique, since if b' is another, then working in $\mathbb{Z}_n$, $\overline{b'} = \overline{b'} \cdot \bar{1} = \overline{b'}(\bar{a}\bar{b}) = (\overline{b'}\bar{a})\bar{b} = \bar{1} \cdot \bar{b} = \bar{b}$, and with b, b' among $1, \ldots, n-1$, $\bar{b} = \overline{b'}$ in $\mathbb{Z}_n$ implies $b = b'$. □

In general, of course, $\mathbb{Z}_n - \{0\}$ is not a group under multiplication. For example, if $n = 6$, then $\bar{2} \cdot \bar{3} = \bar{6} = \bar{0}$ and the set is not closed under multiplication.

The key component of the proof above was that we could use $\gcd(a, p) = 1$ to construct inverses. This gives us the clue to the next collection of examples.

Definition 3.10. *For $n \geq 2$ a positive integer, let $\mathbb{Z}_n^* = \{\bar{r} \mid 1 \leq r \leq n-1,\ \gcd(r, n) = 1\}$, which is a subset of $\mathbb{Z}_n$.*

Remark 3.11. *We could also have defined $\mathbb{Z}_n^* = \{\bar{r}$ in $\mathbb{Z}_n$ with $\gcd(r, n) = 1\}$. The point is that a residue is not simply represented by one particular number, but by any of the numbers that are congruent to it modulo n. But if k is a number then if $x \equiv k \pmod{n}$, we know that $x = k + tn$ for some $t \in \mathbb{Z}$. It is then clear that $\gcd(k, n) = 1$ if and only if $\gcd(x, n) = 1$.*

Examples

$$\mathbb{Z}_2^* = \{\bar{1}\}$$
$$\mathbb{Z}_3^* = \{\bar{1}, \bar{2}\}$$
$$\mathbb{Z}_4^* = \{\bar{1}, \bar{3}\}$$
$$\mathbb{Z}_5^* = \{\bar{1}, \bar{2}, \bar{3}, \bar{4}\}$$
$$\mathbb{Z}_6^* = \{\bar{1}, \bar{5}\}$$
$$\mathbb{Z}_7^* = \{\bar{1}, \bar{2}, \bar{3}, \bar{4}, \bar{5}, \bar{6}\}$$
$$\mathbb{Z}_8^* = \{\bar{1}, \bar{3}, \bar{5}, \bar{7}\}$$
$$\mathbb{Z}_9^* = \{\bar{1}, \bar{2}, \bar{4}, \bar{5}, \bar{7}, \bar{8}\}$$

and so on.

Here is what $\mathbb{Z}_n^*$ gives us.

Proposition 3.12. *(1) The product of two congruence classes in $\mathbb{Z}_n^*$ is in $\mathbb{Z}_n^*$.*
(2) Any congruence class in $\mathbb{Z}_n^*$ has a multiplicative inverse in $\mathbb{Z}_n^*$.
(3) If $\bar{a}$, $\bar{b}$, $\bar{c}$ are in $\mathbb{Z}_n^*$ such that $\bar{a}\bar{b} = \bar{a}\bar{c}$, then $\bar{b} = \bar{c}$.
(4) $\mathbb{Z}_n^*$ is a group under multiplication.

Proof. (1) $\gcd(ab, n) = 1$: any common prime factor of ab and n must come from either a or b (what are we using here?), and $\gcd(a, n) = \gcd(b, n) = 1$. So we have $\overline{ab}$ in $\mathbb{Z}_n^*$ by our second definition of $\mathbb{Z}_n^*$.

(2) For the inverse, we found our inverse from the existence of x, y with $ax + ny = 1$. This implies that $\gcd(ax, n) = 1$ and hence $\gcd(x, n) = 1$. So the inverse of a residue is also in $\mathbb{Z}_n^*$.

(3) Let x be the inverse residue found above, then $\bar{a}\bar{b} = \bar{a}\bar{c} \Rightarrow \bar{x}\bar{a}\bar{b} = \bar{x}\bar{a}\bar{c} \Rightarrow \bar{1}\bar{b} = \bar{1}\bar{c} \Rightarrow \bar{b} = \bar{c}$

(4) We now have closure; $\bar{1}$ is always present and is the identity; we have inverses; associativity again follows from $a(bc) = (ab)c$ in $\mathbb{Z}$.

□

So how big are the $\mathbb{Z}_n^*$ in general? The above discussion means that we only have to count the number of integers k where $1 \leq k < n$ and $\gcd(k, n) = 1$. This is a special number called the **Euler φ-function** or Euler **totient** function. Here is a table of values:

n	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
$\varphi(n)$	1	2	2	4	2	6	4	6	4	10	4	12	6	8	8	16	6	18	8

Some easy observations are: apart from $n = 2$, it seems that $\varphi(n)$ is always even. Why? Also, if p is prime then $\varphi(p) = p - 1$; this should be obvious since every positive number less than a prime is coprime to it.

In fact, there is a nice formula for $\varphi(n)$ in terms of the prime factorisation of n; see the problem sheet.

Note that, given integers a and c, if x satisfies $ax \equiv c \pmod{n}$ and $x \equiv x' \pmod{n}$ then $ax' \equiv c \pmod{n}$ as well. Similarly, we can replace a with $a' \equiv a \pmod{n}$ and c with $c' \equiv c \pmod{n}$, so the x we find are the elements in $\bar{y} \in \mathbb{Z}_n$ satisfying $\bar{a}\,\bar{y} = \bar{c}$ in $\mathbb{Z}_n$. Every such $\bar{y}$ contains exactly one x with $0 \leq x < n$ so the following result can be translated into the solutions of $\bar{a}\,\bar{y} = \bar{c}$ in $\mathbb{Z}_n$ as well.

Theorem 3.13. *Let $n \geq 1$ be a fixed integer, and let $a > 0$ and c be integers.*

(1) *If* $\gcd(a,n)=1$*, then there is a unique* x*,* $0\le x<n$*, such that* $ax\equiv c \pmod n$*.*
(2) *If* $\gcd(a,n)=d>1$ *and* $d\nmid c$*, then there is no solution for* x *satisfying* $ax\equiv c \pmod n$*.*
(3) *If* $\gcd(a,n)=d>1$ *and* $d\mid c$*, then there are precisely* d *values* x*,* $0\le x<n$ *such that* $ax\equiv c \pmod n$*.*

Proof. (1) As $\gcd(a,n)=1$ we can find s,t with $sa+tn=1$ in $\mathbb{Z}$, so $sa\equiv 1 \pmod n$. Now let x be the remainder of sc when divided by n and we get $ax\equiv asc\equiv 1\cdot c\equiv c \pmod n$. This solution is unique: if x' is another then $ax-ax'\equiv 0 \pmod n$, so

$$x-x'\equiv 1\cdot(x-x')\equiv ya(x-x')\equiv y(ax-ax')\equiv 0 \pmod n$$

so that $x=x'$ since both are among $0,1,\dots,n-1$.

(2) If $ax\equiv c \pmod n$ then $ax=c+kn$ for some integer k, so $c=ax-kn$. Since $d=\gcd(a,n)$ divides a and n, the condition $d\mid c$ is necessary.

(3) Since $\gcd(a,n)=d|c$ we can write $a=a'd$, $c=c'd$ and $n=n'd$ for integers $a',n'>0$ and $c'\ge 0$. Solving $ax\equiv c \pmod n$ is equivalent to finding integers x and k with $ax=c+kn$, which by dividing by d is equivalent to $a'x=c'+kn'$, i.e., $a'x\equiv c' \pmod{n'}$. But $\gcd(a',n')=1$: if $\gcd(a',n')=d'$ then $d'd|a'd=a$ and $d'd|n'd=n$, so $d'd\le\gcd(a,n)=d$. So by (1) there is a unique solution x with $0\le x<n'$. This means the solutions among $0,1,\dots,n-1$ are $x,x+n',x+2n',\dots,x+(d-1)n'$.

□

In fact, we can perform even more intricate calculations like solving simultaneous equations modulo distinct numbers under certain conditions.

Theorem 3.14. (CHINESE REMAINDER THEOREM) *Let* $m,n\ge 1$ *and assume* $\gcd(m,n)=1$*. Then for given* a,b *there is a unique* c *with* $0\le c<mn$*,* $c\equiv a \pmod m$*, and* $c\equiv b \pmod n$*.*

Proof. Since $\gcd(m,n)=1$ we can find x,y with $xm+yn=1$. Then $c=bxm+ayn$ satisfies $c\equiv ayn\equiv a \pmod m$ since $yn\equiv 1 \pmod m$. Similarly $c\equiv bxm\equiv b \pmod n$. So if we replace c with its remainder on division by mn we find a solution.

As for uniqueness, if c,c' are solutions then $m|(c-c')$ and $n|(c-c')$. From Theorem 2.5 we see that then $mn|(c-c')$ because $\gcd(m,n)=1$. So $c=c'$ since both are among $0,1,\dots,mn-1$. □

Example 3.15. *Find* $0\le c<120$ *with* $c\equiv 6 \pmod 8$ *and* $c\equiv 13 \pmod{15}$*. Here we have* $m=8$ *and* $n=15$*, and it is easy to see that* $2\cdot 8+(-1)\cdot 15=1$*. So* $c=13\cdot 2\cdot 8+6(-1)15=118$ *should work – and it does.*

Example 3.16. *Find* c, $0 \leq c < 105$ *such that* $c \equiv 2 \pmod 3$, $c \equiv 3 \pmod 5$, $c \equiv 4 \pmod 7$.

We first get a value c' *mod* $3 \cdot 5 = 15$ *that simultaneously solves the first two. We can use* $3 \cdot 2 + (-1) \cdot 5 = 1$ *so* $c' = 3 \cdot 3 \cdot 2 + 2 \cdot (-1) \cdot 5 = 8$*, which works.*

Next we work to find c *with* $c \equiv 8 \pmod{15}$ *and* $c \equiv 4 \pmod 7$*. Now* $1 \cdot 15 + (-2) \cdot 7 = 1$ *and* $c = 4 \cdot 1 \cdot 15 + 8 \cdot (-2) \cdot 7 = -52 \equiv 53 \pmod{105}$.

So, by now we hope we are starting to get familiar with the use of arithmetic mod n, but our aim is to talk about groups. We have the groups $\mathbb{Z}_n$ which use *addition mod* n, and we have the groups $\mathbb{Z}_n^*$ $(n \geq 2)$ which use *multiplication mod* n. But how many really new *structures* are we creating.

Consider the groups $\mathbb{Z}_4$, $\mathbb{Z}_5^*$ and $\mathbb{Z}_8^*$. These all have 4 elements. Do they have the same underlying structure as groups or are they intrinsically different? Look at the group tables:

$\mathbb{Z}_4$

$+$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$
$\bar{0}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$
$\bar{1}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{0}$
$\bar{2}$	$\bar{2}$	$\bar{3}$	$\bar{0}$	$\bar{1}$
$\bar{3}$	$\bar{3}$	$\bar{0}$	$\bar{1}$	$\bar{2}$

$\mathbb{Z}_5^*$

$\times$	$\bar{1}$	$\bar{2}$	$\bar{4}$	$\bar{3}$
$\bar{1}$	$\bar{1}$	$\bar{2}$	$\bar{4}$	$\bar{3}$
$\bar{2}$	$\bar{2}$	$\bar{4}$	$\bar{3}$	$\bar{1}$
$\bar{4}$	$\bar{4}$	$\bar{3}$	$\bar{1}$	$\bar{2}$
$\bar{3}$	$\bar{3}$	$\bar{1}$	$\bar{2}$	$\bar{4}$

$\mathbb{Z}_8^*$

$\times$	$\bar{1}$	$\bar{3}$	$\bar{5}$	$\bar{7}$
$\bar{1}$	$\bar{1}$	$\bar{3}$	$\bar{5}$	$\bar{7}$
$\bar{3}$	$\bar{3}$	$\bar{1}$	$\bar{7}$	$\bar{5}$
$\bar{5}$	$\bar{5}$	$\bar{7}$	$\bar{1}$	$\bar{3}$
$\bar{7}$	$\bar{7}$	$\bar{5}$	$\bar{3}$	$\bar{1}$

We have deliberately made the ordering of the elements in the middle case unusual to demonstrate that the first two structures are actually the same. Make the elements in the first two table correspond like this:

$$\mathbb{Z}_4 \leftrightarrow \mathbb{Z}_5^*$$
$$\bar{0} \leftrightarrow \bar{1}$$
$$\bar{1} \leftrightarrow \bar{2}$$
$$\bar{2} \leftrightarrow \bar{4}$$
$$\bar{3} \leftrightarrow \bar{3}$$

and everything in the left two tables matches exactly. For example, $\bar{2}+\bar{3} = \bar{1}$ in $\mathbb{Z}_4$ corresponds to $\bar{4} \cdot \bar{3} = \bar{2}$ in $\mathbb{Z}_5^*$.

On the other hand, the table for $\mathbb{Z}_8^*$ is different. All the top left bottom right diagonal elements are the same and this is not the case with the other two, no matter in which order we write the elements in the tables (using the same order in the top row and leftmost column).

A more sophisticated way of saying the same thing is that in $\mathbb{Z}_8^*$ every element satisfies $x \cdot x = \bar{1}$, the identity element, but in $\mathbb{Z}_4$, $\bar{1} + \bar{1} = \bar{2} \neq \bar{0}$ [remember in $\mathbb{Z}_4$ we use addition and the identity element is $\bar{0}$] and in $\mathbb{Z}_5^*$, $\bar{2} \cdot \bar{2} = \bar{4} \neq \bar{1}$.

In fact there are only two distinct group structures for groups with 4 elements and they are represented by $\mathbb{Z}_5^*$ and $\mathbb{Z}_8^*$. For example, $\mathbb{Z}_{10}^*$ also has 4 elements, but $\bar{3}^2 = \bar{9} \neq \bar{1}$, so it cannot have the same structure as $\mathbb{Z}_8^*$ and $\mathbb{Z}_{10}^*$ has the same structure as $\mathbb{Z}_4$ and $\mathbb{Z}_5^*$.

The notion of groups having the same structures is formalised as follows.

Definition 3.17. *Let $(G, \circ)$ and $(H, \star)$ be groups and $\varphi : G \to H$ a map.*

(1) *if $\varphi(g_1 \circ g_2) = \varphi(g_1) \star \varphi(g_2)$ for all g_1, g_2 in G, then φ is called a* **homomorphism***;*
(2) *φ is called* **injective** *if $\varphi(g_1) = \varphi(g_2)$ implies $g_1 = g_2$;*
(3) *φ is called* **surjective** *if for every h in H, $h = \varphi(g)$ for some g in G;*
(4) *φ is called an* **isomorphism** *if it satisfies (1), (2) and (3).*

We write $G \cong H$ when G is isomorphic to H.

Exercise 3.18. If $G \cong H$ then $H \cong G$.

Here are a few properties of homomorphisms:

Lemma 3.19. *If G and H are groups and $\varphi : G \to H$ is a homomorphism then:*

(1) *$\varphi(e_G) = e_H$;*
(2) *$\varphi(g^{-1}) = \varphi(g)^{-1}$ for any g in G;*
(3) *φ is injective if and only if $\varphi(g) = e_H$ implies $g = e_G$.*

Example 3.20. *$\mathbb{R}$ is a group under addition, and $\mathbb{R}_{>0}^* = \{x \text{ in } \mathbb{R} \text{ with } x > 0\}$ is a group under multiplication (check this). Then $\varphi : \mathbb{R} \to \mathbb{R}_{>0}^*$ given by $\varphi(x) = e^x$ is bijective and $\varphi(x + y) = e^{x+y} = e^x e^y = \varphi(x)\varphi(y)$. So $\mathbb{R} \cong \mathbb{R}_{>0}^*$.*

So an important question is how we can recognise when groups are isomorphic or not? Luckily for finite groups associated with things like $\mathbb{Z}_n$ this is not too hard. But first we need to look at a way of constructing new groups from old ones.

Recall that the **Cartesian product** of two sets is the set of **ordered pairs** of elements from the sets. In particular we want to look at the Cartesian product $\mathbb{Z}_m \times \mathbb{Z}_n$ for two integers m, n. So, first of all,

$$\mathbb{Z}_m \times \mathbb{Z}_n = \{(\bar{a}, \bar{b}) \mid \bar{a} \in \mathbb{Z}_m,\ \bar{b} \in \mathbb{Z}_n\}.$$

We define addition **coordinatewise**:

$$(\overline{a_1}, \overline{b_1}) + (\overline{a_2}, \overline{b_2}) = (\overline{a_1} + \overline{a_2}, \overline{b_1} + \overline{b_2}) = (\overline{a_1 + a_2}, \overline{b_1 + b_2}) .$$

It is easy to check that $\mathbb{Z}_m \times \mathbb{Z}_n$ is a group under the addition we've just defined, using that $\mathbb{Z}_m$ and $\mathbb{Z}_n$ are groups.

We also know that:

Lemma 3.21. *$\mathbb{Z}_m \times \mathbb{Z}_n$ has mn elements.*

Example 3.22. *So what does $\mathbb{Z}_2 \times \mathbb{Z}_2$ look like? Here is its addition table:*

$+$	$(\bar{0}, \bar{0})$	$(\bar{1}, \bar{1})$	$(\bar{1}, \bar{0})$	$(\bar{0}, \bar{1})$
$(\bar{0}, \bar{0})$	$(\bar{0}, \bar{0})$	$(\bar{1}, \bar{1})$	$(\bar{1}, \bar{0})$	$(\bar{0}, \bar{1})$
$(\bar{1}, \bar{1})$	$(\bar{1}, \bar{1})$	$(\bar{0}, \bar{0})$	$(\bar{0}, \bar{1})$	$(\bar{1}, \bar{0})$
$(\bar{1}, \bar{0})$	$(\bar{1}, \bar{0})$	$(\bar{0}, \bar{1})$	$(\bar{0}, \bar{0})$	$(\bar{1}, \bar{1})$
$(\bar{0}, \bar{1})$	$(\bar{0}, \bar{1})$	$(\bar{1}, \bar{0})$	$(\bar{1}, \bar{1})$	$(\bar{0}, \bar{0})$

The other example with 4 elements that we have to hand is $\mathbb{Z}_4$. Is this the same? Let's look:

$+$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$
$\bar{0}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$
$\bar{1}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{0}$
$\bar{2}$	$\bar{2}$	$\bar{3}$	$\bar{0}$	$\bar{1}$
$\bar{3}$	$\bar{3}$	$\bar{0}$	$\bar{1}$	$\bar{2}$

Immediately we see a difference: the main diagonal for addition for $\mathbb{Z}_2 \times \mathbb{Z}_2$ always has $(\bar{0}, \bar{0})$, but for $\mathbb{Z}_4$ it is not always $\bar{0}$.

Sometimes there is a relation between $\mathbb{Z}_{mn}$ and $\mathbb{Z}_m \times \mathbb{Z}_n$:

Theorem 3.23. *Let $m,\ n \geq 1$ be integers with $\gcd(m, n) = 1$. Then $\varphi : \mathbb{Z}_{mn} \to \mathbb{Z}_m \times \mathbb{Z}_n$ given by $\varphi(\bar{a}) = (\bar{a}, \bar{a})$ is an isomorphism of groups. It also preserves the multiplication.*

Proof. Remember that in $\varphi(\bar{a}) = (\bar{a}, \bar{a})$ we have that $\bar{a}$ is in $\mathbb{Z}_{mn}$, $\mathbb{Z}_m$ or $\mathbb{Z}_n$, depending on where it occurs.

Before we check that this is a 1-1 correspondence and that all the algebraic structure is preserved by φ we have to check that the definition of φ makes sense: $\bar{a}$ in $\mathbb{Z}_{mn}$ determines a up to a multiple of mn. But if we used $a' = a + kmn$ instead of a in $(\bar{a}, \bar{a})$ then we get $(\overline{a'}, \overline{a'}) = (\overline{a + (kn)m}, \overline{a + (km)n})$, which equals $(\bar{a}, \bar{a})$ in $\mathbb{Z}_m \times \mathbb{Z}_n$.

Now suppose $\varphi(\bar{a}) = \varphi(\bar{b})$, then $a \equiv b \pmod{m}$ and $a \equiv b \pmod{n}$. So $m|(a-b)$ and $n|(a-b)$. But since $\gcd(m,n) = 1$ this means that $mn|(a-b)$ and so $a \equiv b \pmod{mn}$. Thus φ is injective.

The fact that both sets have the same size ($= mn$) implies that it also has to be onto (surjective), so that it is indeed a 1-1 correspondence. (In fact, finding an element $\bar{c}$ of $\mathbb{Z}_{mn}$ that maps to $(\bar{a}, \bar{b})$ in $\mathbb{Z}_m \times \mathbb{Z}_n$ is just an application of the Chinese remainder theorem, Theorem 3.14).

Finally, φ does respect addition as $\varphi(\bar{a} + \bar{b}) = \varphi(\bar{a}) + \varphi(\bar{b})$ (check!). □

NOTE: You need to be clear what Theorem 3.23 is saying. If $\gcd(m,n) = 1$ then we know $\mathbb{Z}_m \times \mathbb{Z}_n \cong \mathbb{Z}_{mn}$. In fact the converse is true, i.e., $\mathbb{Z}_m \times \mathbb{Z}_n \cong \mathbb{Z}_{mn} \Rightarrow \gcd(m,n) = 1$. Hence:

$$\mathbb{Z}_4 \times \mathbb{Z}_8 \ncong \mathbb{Z}_{32},\ \mathbb{Z}_6 \times \mathbb{Z}_{10} \ncong \mathbb{Z}_{60}$$

although, in the latter case we can rearrange things a bit using the rules we have just developed: $\mathbb{Z}_6 \times \mathbb{Z}_{10} \cong \mathbb{Z}_2 \times \mathbb{Z}_3 \times \mathbb{Z}_2 \times \mathbb{Z}_5 \cong \mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_3 \times \mathbb{Z}_5 \cong \mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_{15} \cong \mathbb{Z}_2 \times \mathbb{Z}_{30}$.

FACT: If G is a finite group which is abelian (i.e., such that $\forall g_1,\ g_2 \in G,\ g_1 g_2 = g_2 g_1$) then G is isomorphic to a product of groups of the form $\mathbb{Z}_n$.

FACTS:

(1) There are only two abelian isomorphism types with 4 elements: $\mathbb{Z}_4$ and $\mathbb{Z}_2 \times \mathbb{Z}_2$.
(2) There is only one abelian isomorphism type with 6 elements: $\mathbb{Z}_6$.
(3) There are only three abelian isomorphism types with 4 elements: $\mathbb{Z}_8$, $\mathbb{Z}_2 \times \mathbb{Z}_4$ and $\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2$.
(4) There are only two abelian isomorphism types with 9 elements: $\mathbb{Z}_9$ and $\mathbb{Z}_3 \times \mathbb{Z}_3$.
(5) If p is prime, there is only one abelian isomorphism type with p elements: $\mathbb{Z}_p$.

So, for instance, all our groups $\mathbb{Z}_k^*$ are of this type. How do we recognise which product of these standard mod n groups these are?

Example 3.24. *In an earlier example we saw that $\mathbb{Z}_5^*$ is isomorphic to $\mathbb{Z}_4$. We also saw that $\mathbb{Z}_8^*$, which similarly has 4 elements, was different and a quick comparison back will show that $\mathbb{Z}_8^*$ is isomorphic to $\mathbb{Z}_2 \times \mathbb{Z}_2$.*

The trick to identifying these groups is to use the following.

Definition 3.25. *Let G be a group [written* multiplicatively*] and $g \in G$, then the least positive integer r such that $\bar{x}^r = \bar{1}$ is called the* **order** *of $\bar{x}$.*

FACT: If G has n elements then the order of g will always divide n (see Theorem 5.13).

Of course in $\mathbb{Z}_n$ [where we write the group operation *additively*], the order of an element $\bar{a}$ is the least positive integer r such that $r\bar{a}(=\overline{ra}) = \bar{0}$.

Example 3.26. *So what is the structure of $\mathbb{Z}_9^*$? We quickly check that it has 6 elements, so it is of type $\mathbb{Z}_6$.*

What is the structure of $\mathbb{Z}_{10}^*$? This has 4 elements $\bar{1},\ \bar{3},\ \bar{7},\ \bar{9}$. We spot that $\bar{3}$ does not have order 2, so it must have order 4, which means we must have type $\mathbb{Z}_4$ since $\mathbb{Z}_2 \times \mathbb{Z}_2$ only has elements of order 2.

What is the structure of $\mathbb{Z}_{15}^*$? This has 8 elements. We check that there is no element of order 8 so we do not have type $\mathbb{Z}_8$, but $\bar{2}$ has order 4, so this must be of type $\mathbb{Z}_2 \times \mathbb{Z}_4$, since the only other possibility is $\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2$ where all the elements have order 1 or 2.

So, in practice, by checking the orders of the elements we can match numbers against the orders of elements in our standard types and identify the groups.

These ideas lead on quite neatly to one way of doing *public key cryptography*. But we need one important result to see the connection.

Theorem 3.27. Euler-Fermat *If $n > 1$ and $\gcd(a, n) = 1$, then $a^{\varphi(n)} \equiv 1 \pmod{n}$.*

Proof. Let $[\overline{x_1}, \overline{x_2}, \ldots, \overline{x_{\varphi(n)}}]$ be a list of all the distinct elements of $\mathbb{Z}_n^*$ and let $\overline{z} = \overline{x_1}\,\overline{x_2}\cdots\overline{x_{\varphi(n)}}$ be their product. Now multiply every element of $\mathbb{Z}_n^*$ by $\overline{a}$ and observe that by Lemma 3.6, the resulting elements are all distinct and are all in $\mathbb{Z}_n^*$. So all we have done by multiplying each by $\overline{a}$ is to shuffle the elements of $\mathbb{Z}_n^*$ in some fashion. Because $\mathbb{Z}_n^*$ is abelian this means that the product $\overline{ax}_1\overline{ax}_2\cdots\overline{ax_{\varphi(n)}}$ is also equal to $\overline{z}$. Thus $\overline{z} = \overline{a^{\varphi(n)}z}$ and so $\overline{a^{\varphi(n)}} = \overline{1}$ in $\mathbb{Z}_n^*$, which is what was required. □

Corollary 3.28. Fermat's little theorem *Let p be a prime. Then $\bar{a}^p = \bar{a}$ for any $\bar{a}$ in $\mathbb{Z}_p$.*

Example 3.29. Public Key Cryptography

Alice wants to send a secure message to Bob in code. Bob gives Alice a simple way of encoding her message, but only Bob knows how to decode it again.

In fact Bob can make his encoding procedure public so anyone can send him a message in code and, even if the coded message is intercepted, because only Bob has the decoding key it ought to be secure.

How can this work?

Bob chooses two (very) large primes p, q. He then finds two (large) positive integers d, e such that $de \equiv 1 \pmod{(p-1)(q-1)}$.

Alice converts her message into a sequence of numbers among $0, 1, \ldots, pq-1$. Let T be one of these numbers. Alice computes $M \equiv T^d \pmod{pq}$, with M among $0, 1, \ldots, pq-1$, and sends M to Bob as the coded message.

Bob then works out $U \equiv M^e \pmod{pq}$ with U among $0, 1, \ldots, pq-1$ and, lo and behold, finds that $U = T$ and he has decoded the message.

The mathematics is really quite straightforward now:

$$U \equiv M^e \equiv (T^d)^e \equiv T^{de} \pmod{pq}.$$

Also $de \equiv 1 \pmod{(p-1)(q-1)}$, so $de = 1 + m(p-1)(q-1)$ for some integer m. Then

$$U \equiv T^{de} = T^{1+m(p-1)(q-1)} \equiv T \pmod{p}$$

by Corollary 3.28. Similarly $U \equiv T \pmod{q}$, so $U \equiv T \pmod{pq}$ since p and q are distinct primes. Then $U = T$ since both are among $0, 1, \ldots, pq-1$.

Even for very large p, q, all the processes involved are easy and quick to implement on a computer. The real issues are finding large enough primes (how do you check whether a number with 100 decimal digits is prime?). Once you have selected p, q it is easy to select d, e if you know what p, q are. But without knowing p and q finding e is virtually impossible (even if you know pq and d) since you need to know the value of $(p-1)(q-1)$.

On the other hand you also know the value of the product pq, and so if you can factorise the product (which may have 200 decimal digits), then you can easily crack the code. So this form of public key cryptography is as safe as our inability to factorise very large numbers.

4. Permutations

A permutation is, informally speaking, simply a rearrangement of an ordered set of objects into a (possibly) different order.

We shall always be looking at the set $C_n := \{1, 2, \ldots, n\}$ for some integer n (with the natural order that the integers come in).

Definition 4.1. *A* **permutation** *is then a bijective (i.e., both injective and surjective) map from C_n to itself.*

For example, this is a permutation of C_5:

$$1 \mapsto 5$$
$$2 \mapsto 4$$
$$3 \mapsto 3$$
$$4 \mapsto 2$$
$$5 \mapsto 1$$

This notation would get very tedious if we were dealing with permutations of any size. So, at least to start with we use an alternate notation where we write out the elements of C_n in a line and underneath them write out where they are sent by the permutation. So the above permutation is written $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 4 & 3 & 2 & 1 \end{pmatrix}$.

If $\sigma : C_n \to C_n$ and $\tau : C_n \to C_n$ are permutations then we can compose them to get a third $\tau\sigma : C_n \to C_n$ (first do σ and then τ). How would we work this out with our bracket notation?

Suppose $\tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 4 & 5 & 1 \end{pmatrix}$. Now we can work out the composition $\tau\sigma$ by hand. E.g., $1 \mapsto 5 \mapsto 1$, $2 \mapsto 4 \mapsto 5$, $3 \mapsto 3 \mapsto 4$, $4 \mapsto 2 \mapsto 3$, $5 \mapsto 1 \mapsto 2$ to get $\tau\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 5 & 4 & 3 & 2 \end{pmatrix}$. Perhaps an easier way is to reorganise the top row of τ so that it agrees with the bottom row of σ and then we can "cancel" rows to get the result.

So $\tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 4 & 5 & 1 \end{pmatrix} = \begin{pmatrix} 5 & 4 & 3 & 2 & 1 \\ 1 & 5 & 4 & 3 & 2 \end{pmatrix}$ and then $\tau\sigma = \begin{pmatrix} 5 & 4 & 3 & 2 & 1 \\ 1 & 5 & 4 & 3 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 4 & 3 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 5 & 4 & 3 & 2 \end{pmatrix}$.

We call this process taking the **product of two permutations**.

We shall now show that the set of all permutations of C_n is a group, called the *symmetric* group on n elements and denoted S_n.

Proposition 4.2. (1) *The product of two permutations in S_n is a permutation in S_n.*

(2) *Let e be the* **identity permutation** *where $e(i) = i$ for every i, $1 \leq i \leq n$. Then for every permutation σ in S_n, $\sigma e = e\sigma = \sigma$.*

(3) *For every σ in S_n there exists an inverse permutation σ^{-1} such that $\sigma\sigma^{-1} = \sigma^{-1}\sigma = e$.*

Proof. (1) Obvious

(2) Obvious

(3) A bijective function has an inverse function, which is again bijective. [Note that in the double row notation we obtain σ^{-1} from σ by swapping the two rows.]

□

Example 4.3. *Let* $\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$ *and* $\tau = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}$*, then* $\sigma\tau \neq \tau\sigma$*:*

$$\tau\sigma = \begin{pmatrix} 1 & 3 & 2 \\ 2 & 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \qquad \textit{but} \qquad \sigma\tau = \begin{pmatrix} 2 & 1 & 3 \\ 3 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}.$$

Even the two-row notation is too cumbersome for day-to-day calculations. So we introduce another notation that will simplify calculations greatly as well as allowing us to understand the structures of permutations rather better.

Definition 4.4. *A* **cycle** *on a subset of* C_n *is a sequence of* distinct *elements* $a_1, a_2, \ldots, a_k$ *of* C_n *(not necessarily in ascending order). We convert this into a permutation by requiring*

$$a_1 \mapsto a_2 \mapsto a_3 \mapsto \cdots \mapsto a_{k-1} \mapsto a_k \mapsto a_1\,,$$

and $r \mapsto r$ *for every other value of* r*.*

Because the moving elements just cycle around (hence the name), we will use the notation $(a_1\ a_2\ \cdots\ a_k)$ *to represent this cycle. If any number is absent from this list, all it means is that the cycle does not move that number in the rearrangement.*

Example 4.5. *The permutation* $\tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 4 & 5 & 1 \end{pmatrix}$ *above is such a cycle, as* $1 \mapsto 2$*,* $2 \mapsto 3$*,* $3 \mapsto 4$*,* $4 \mapsto 5$ *and* $5 \mapsto 1$*.*

Note that the same cycle can be written in k ways, because we can start the cyclic shift of numbers at any point in the list. So

$$(a_1\ a_2\ \cdots\ a_k) = (a_2\ a_3\ \cdots\ a_k\ a_1) = (a_3\ a_4\ \cdots\ a_k\ a_1\ a_2) = (a_i\ a_{i+1}\ \cdots\ a_k\ a_1\ a_2\ \cdots\ a_{i-1})\,.$$

Definition 4.6. *Two cycles* $(a_1\ a_2\ \cdots\ a_k)$ *and* $(b_1\ \cdots\ b_\ell)$ *are called* **disjoint** *if the two sets of moving elements have nothing in common.*

A cycle $(a_1\ a_2\ \cdots\ a_k)$ *is called a* **cycle of length** k*, or short a* k**-cycle**.

Example 4.7. *The 4-cycle* $(2\ 5\ 1\ 7)$ *is disjoint from the 3-cycle* $(6\ 3\ 4)$*, but not from the 3-cycle* $(6\ 3\ 5)$*.*

Lemma 4.8. *If* $\sigma,\ \tau$ *are two disjoint cycles, then* $\sigma\tau = \tau\sigma$*.*

Proof. The moving elements of the cycle σ are completely independent of the moving elements of τ. □

Theorem 4.9. *Every permutation is a product of disjoint cycles, and this way of writing the permutation is unique up to reordering the product.*

Proof. Remember that a permutation of the set $\{1, \ldots, n\}$ is just a rearrangement of the numbers.

Let σ be our permutation, and consider any number a, say, $1 \leq a \leq n$. Define $\sigma^i(a)$ to be the value $\sigma(\sigma(\cdots\sigma(a)\cdots))$, where the function σ has been applied i times to a. As usual $\sigma^0(a) = a$.

Now look at the sequence $a,\ \sigma(a)\ \sigma^2(a),\ \sigma^3(a), \ldots,\ \sigma^i(a), \ldots$. Since this is taking place in the finite set C_n an element must be repeated in the sequence, say $\sigma^r(a) = \sigma^s(a)$, where $r < s$. Pick the smallest $r \geq 0$ with $\sigma^r(a) = \sigma^s(a)$ for some $s > r$.

If $r > 0$, then $\sigma(\sigma^{r-1}(a)) = \sigma(\sigma^{s-1}(a))$. Since σ is injective we must have $\sigma^{r-1}(a) = \sigma^{s-1}(a)$. But this contradicts the minimality of r since $0 \leq r - 1 < s - 1$ still. Hence $r = 0$.

Now let $s > 0$ be the smallest such that $\sigma^s(a) = a$. Then $\gamma(a) = (a\ \sigma(a)\ \sigma^2(a)\ \cdots\ \sigma^{s-1}(a))$ is a cycle: if $\sigma^i(a) = \sigma^j(a)$ for $0 \leq i < j \leq s - 1$, then $a = \sigma^{j-i}(a)$. But $0 < j - i \leq s - 1$ so this is not possible by our choice of s.

Now we are in a position to decompose σ as a product of disjoint cycles.

We will successively choose $a_i \in \{1, \ldots, n\}$. Choose $a_1 = 1$, and write down $\gamma(a_1)$, the cycle starting with a_1. If this uses all of σ then $\sigma = \gamma(a_1)$ and we are done.

Otherwise, there is a least number a_2 that we have not yet used. Construct the cycle $\gamma(a_2)$. None of the elements of this new cycle can appear in $\gamma(a_1)$, because once you arrive at an element in a cycle you stay in that cycle and indeed you must have come from that cycle.

So $\gamma(a_1)$ and $\gamma(a_2)$ are disjoint cycles making up part of σ. If we have the whole of σ then we can stop, otherwise we choose the lowest number, a_3 say, not yet seen and construct the cycle $\gamma(a_3)$. Again this must comprise of new elements only and be disjoint from $\gamma(a_1)$ and $\gamma(a_2)$.

Proceeding in this way, we continue to pick out new starting points for cycles so long as there are elements available. The process must eventually stop, since we only have a finite supply of numbers to shuffle. At this point we have written σ as a product of disjoint cycles.

The product is essentially unique, since once we look at what $\sigma(a)$ is for any number a, it defines the cycle in which it appears. There is no choice about the cycle even though we can of course write down the numbers in the cycle in a number of equivalent ways. □

Example 4.10. *Consider* $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 5 & 7 & 3 & 8 & 2 & 4 & 1 & 6 \end{pmatrix}$. *Then* $\sigma = (1\ 5\ 2\ 7)(3)(4\ 8\ 6)$.

Once we understand this properly, it becomes much easier to manipulate permutations. Indeed we can do multiplication of permutations very quickly if they are written as product of disjoint cycles.

Convention. A **trivial cycle** is one of length 1, e.g., (a). Conventionally we tend to omit references to cycles of length 1 on the understanding that if a number is absent from a product of disjoint cycles, then the relevant permutation sends that number to itself and so constitutes a trivial cycle. (So in the above example, we would have been very happy to write just $\sigma = (1\ 5\ 2\ 7)(4\ 8\ 6)$.) The identity permutation is denoted e (or sometimes (1)).

Example 4.11. *Let* $\sigma = (1\ 3\ 5)(4\ 8)$ *and* $\tau = (3\ 2\ 1\ 8)(4\ 6)(5\ 7)$ *in* S_8. *How do we compute* $\sigma\tau$*?*

First we write them together in the correct order:

$$(1\ 3\ 5)(4\ 8)(3\ 2\ 1\ 8)(4\ 6)(5\ 7)\,.$$

Then we remind ourselves of two things. First a cycle is just a way of moving numbers around in a circle and secondly that these are functions and when we compose them we do so **from right to left**.

There are 5 separate cycles written down here. We ask what is the effect of each in turn working from the right *on a given number. Consider 1.*

The rightmost cycle does not involve 1 and so $1 \mapsto 1$. *Similarly the second right one does the same. The third from the right sends* $1 \mapsto 8$. *Now we have got to the number 8 and so we must continue with the fourth cycle and ask for its effect on the number 8; here* $8 \mapsto 4$ *and the final cycle then leaves 4 alone. So the total effect is*

$$1 \overset{(5\ 7)}{\mapsto} 1 \overset{(4\ 6)}{\mapsto} 1 \overset{(3\ 2\ 1\ 8)}{\mapsto} 8 \overset{(4\ 8)}{\mapsto} 4 \overset{(1\ 3\ 5)}{\mapsto} 4 \Rightarrow\ 1 \mapsto 4 \text{ overall.}$$

Now the object is to construct the correct cycle structure. So we ask what happens to 4:

$$4 \mapsto 4 \mapsto 6 \mapsto 6 \mapsto 6 \mapsto 6 \ \Rightarrow\ 4 \mapsto 6 \text{ overall.}$$

then

$$6 \mapsto 6 \mapsto 4 \mapsto 4 \mapsto 8 \mapsto 8 \ \Rightarrow\ 6 \mapsto 8 \text{ overall.}$$

$$8 \mapsto 8 \mapsto 8 \mapsto 3 \mapsto 3 \mapsto 5 \ \Rightarrow\ 8 \mapsto 5 \text{ overall.}$$

$$5 \mapsto 7 \mapsto 7 \mapsto 7 \mapsto 7 \mapsto 7 \ \Rightarrow\ 5 \mapsto 7 \text{ overall.}$$

$$7 \mapsto 5 \mapsto 5 \mapsto 5 \mapsto 5 \mapsto 1 \ \Rightarrow\ 5 \mapsto 1 \text{ overall.}$$

and we have the cycle

$$(1\ 4\ 6\ 8\ 5\ 7)$$

We still have to look at any numbers we have not yet worked on. The least of these is 2 so:

$$2 \mapsto 2 \mapsto 2 \mapsto 1 \mapsto 1 \mapsto 3 \;\Rightarrow\; 2 \mapsto 3 \text{ overall.}$$

$$3 \mapsto 3 \mapsto 3 \mapsto 2 \mapsto 2 \mapsto 2 \;\Rightarrow\; 3 \mapsto 2 \text{ overall.}$$

and we have the cycle $(2\ 3)$*. Hence*

$$\sigma\tau = (1\ 4\ 6\ 8\ 5\ 7)(2\ 3)\,.$$

Some easy examples are:
$(1\ 3)(1\ 2) = (1\ 2\ 3)$*;* $(1\ 4)(1\ 3)(1\ 2) = (1\ 2\ 3\ 4)$*;* $(1\ 5)(1\ 4)(1\ 3)(1\ 2) = (1\ 2\ 3\ 4\ 5)$.
In fact we see a pattern here.

Lemma 4.12. *Let* $(a_1 \cdots a_k)$ *be a* k*-cycle, then*

$$(a_1 \cdots a_k) = (a_1\ a_k)(a_1\ a_{k-1}) \cdots (a_1\ a_2)\,.$$

Proof. Just work it out and see (or use formal induction). □

Definition 4.13. *A 2-cycle is usually called a* **transposition**.

So, from the Lemma, the following is very easy.

Theorem 4.14. *Every permutation is a product of transpositions.*

Proof. Write the permutation as a product of disjoint cycles and then every cycle as a product of transpositions. □

Note that the **inverse** σ^{-1} of a k-cycle $\sigma = (a_1\ a_2\ \ldots\ a_k)$ is itself a k-cycle, given simply by mirroring the order in σ, i.e.

$$\sigma^{-1} = (a_k\ a_{k-1}\ \ldots\ a_1)\,.$$

From this it easily follows that the inverse of a product $\sigma_1 \cdot \sigma_2 \cdots \sigma_r$ of cycles is given by the "total mirror" $s_r^{-1} \cdot \sigma_{r-1}^{-1} \cdots \sigma_1$. (Recall that the inverse of a product of grouop elements is the product of its inverses, but **in reverse order**: $(g\,h)^{-1} = h^{-1}\ g^{-1}$.)

Example 4.15. (1) *The inverse of* $(1\ 2\ 3\ 4)$ *is* $(4\ 3\ 2\ 1)$.
(2) *The inverse of* $(1\ 3\ 7\ 4)(1\ 2)(2\ 4\ 3)$ *is given by* $(3\ 4\ 2)(2\ 1)(4\ 7\ 3\ 1)$.

How many transpositions do we need to write a given permutation? Try the following. Let σ be our permutation. and write

$$\sigma = \gamma_1\gamma_2 \cdots \gamma_r$$

as a product of *disjoint* cycles, where γ_i is a cycle of length $k_i > 1$, say. The bracketed sequence of numbers

$$[k_1, k_2, \ldots, k_r],$$

where we demand $1 < k_1 \leq k_2 \leq \cdots \leq k_r$ (this is possible since the g_i are assumed to be disjoint cycles and hence commute with each other), is called the **cycle type** of the permutation. For example, the cycle type of (1 2 3 4)(5 6) is [2,4], and the cycle type of (1 2 3)(4)(6 7 8) is $[3, 3]$ (as we ignore the 1-cycle (4)).

Lemma 4.12 above says that a cycle of length k can be expressed as a product of $k - 1$ transpositions. The theorem then implies

Proposition 4.16. *A permutation of cycle type $[k_1, k_2, \ldots, k_r]$ can be expressed as a product of $k_1 + k_2 + \cdots + k_r - r$ transpositions.*

Definition 4.17. *A permutation is called* **even** *if the number $k_1 + k_2 + \cdots + k_r - r$ is even and* **odd** *if it is odd. This is the* **parity** *of the permutation.*

This parity is important in a number of areas of mathematics. Indeed, we can express the determinant of a matrix using it.

Theorem 4.18. *The determinant of an $(n \times n)-$matrix A is*

$$\det(A) = \sum_{\sigma \in S_n} \varepsilon(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)}$$

where $\varepsilon(\sigma) = +1$ if σ is even and $\varepsilon(\sigma) = -1$ if σ is odd. (Proved in Linear Algebra.) □

We have shown that any permutation can be written as a product of transpositions. There is no unique way of doing it for a given permutation (e.g., the identity permutation in S_n $(n \geq 2)$ satisfies $e = (1\ 2)(1\ 2) = (1\ 2)(1\ 2)(1\ 2)(1\ 2) = (1\ 2)(1\ 2)(1\ 2)(1\ 2)(1\ 2)(1\ 2) = \cdots$). What is true, however, is that the **parity** of the number of permutations is well-defined.

The following is proved in Exercise 35. As above, $\varepsilon(\sigma) = +1$ if σ is even and -1 if σ is odd.

Theorem 4.19. *If a permutation σ can be written as a product of t transpositions then $\varepsilon(\sigma) = (-1)^t$. In particular, t is always odd for odd permutations and always even for even permutations.*

This theorem implies that the product of two odd or two even permutations is even; an even permutation multiplied by an odd one is odd. So the parity property on multiplication of permutations acts like the addition property in $\mathbb{Z}_2$.

To conclude this section we work out the order of a permutation from its cycle type (remember that the order of a group element is the smallest positive power that gets you back to the identity element).

Theorem 4.20. *Let $\sigma = \gamma_1\gamma_2\cdots\gamma_r$ be the permutation σ expressed as a product of disjoint cycles, where, for each j, γ_j is a cycle of length k_j. Then the order of σ is* $\operatorname{lcm}\{k_1, k_2, \ldots, k_r\}$.

5. Groups revisited

Now we have a really good supply of examples of groups. Apart from groups like $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$, we have $\mathbb{Q}^* = \mathbb{Q}\backslash\{0\}$, $\mathbb{R}^* = \mathbb{R}\backslash\{0\}$ and $\mathbb{C}^* = \mathbb{C}\backslash\{0\}$ as well as $\mathbb{Z}_n$, $\mathbb{Z}_n^*$ and S_n. Furthermore, we have encountered the alternating groups $A_n = \big\{\pi \in S_n \,|\, \pi \text{ is an } \textit{even} \text{ permutation}\big\}$ and the cyclic groups $\mathcal{C}_n = \big\{\exp(2\pi i\, k/n) \;|\; 0 \le k \le n-1\big\}$. (Here we use the standard notation for $i \in \mathbb{C}$ such that $i^2 = -1$.)

As we can see, the actual "group multiplication" can take a number of forms. It may be normal addition or normal multiplication of numbers, it can be composition of functions. There are lots of possibilities depending on the context.

For example, we can take the above groups as building blocks, and form new groups out of them by *taking products*, whose definition is recalled here:

Definition 5.1. *Let G and H be abstract groups, then we can form a new group $G \times H$, the* **Cartesian product** *of G and H, where $G \times H = \{(g, h) \,|\, g$ in G, h in $H\}$ and the multiplication is given by*

$$(g_1, h_1)(g_2, h_2) = (g_1g_2, h_1h_2)\,.$$

Checking that the group axioms are satisfied is very easy and is left as an exercise.

Example 5.2. *The following groups are examples for Cartesian products of smaller groups:*

$$\mathbb{Z}_2 \times \mathbb{Z}_4\,, \qquad \mathbb{Z}\times\mathbb{Z}\times\mathbb{Z}\,, \qquad \mathbb{R}\times\mathbb{R} \quad (= \mathbb{R}^2\,,\ \textit{the points in the } xy-\textit{plane})\,.$$

We can *relate groups* by homomorphisms.

Example 5.3. *The following maps are homomorphisms:*

$$\begin{aligned} \psi_1 : \mathbb{Z} &\to \mathbb{Z}\,, \\ a &\mapsto 2a\,, \qquad \textit{(not surjective, so no isomorphism)} \end{aligned}$$

$$\begin{aligned} \psi_2 : \mathbb{Z}_2 \times \mathbb{Z}_4 &\to \mathbb{Z}_{15}^*\,, \\ (\overline{a}, \overline{b}) &\mapsto \overline{(-1)^a \cdot 2^b} \qquad \textit{(this gives an isomorphism),} \end{aligned}$$

$$\begin{aligned} \psi_3 : \mathbb{Z}_n &\to \mathcal{C}_n\,, \\ k &\mapsto \exp(2\pi i k/n) \qquad \textit{(this also gives an isomorphism).} \end{aligned}$$

$$\begin{aligned} \varepsilon : S_n &\to \{\pm 1\}, \\ \sigma &\mapsto \varepsilon(\sigma), \qquad \textit{(cf. Thm 4.18).} \end{aligned}$$

Furthermore, we can define *subgroups* of a given group:

Definition 5.4. *Let G be a group. A* **subgroup** *of G is a non-empty subset H of G that is a group with the* same *multiplication as G. We denote this by $H \leq G$.*

Example 5.5.
- *$A_n \leq S_n$ ($n \geq 2$);*
- *$2\mathbb{Z} = \{2k \mid k \in \mathbb{Z}\} \leq \mathbb{Z}$;*
- *$n\mathbb{Z} \leq \mathbb{Z}$;*
- *$\mathbb{Z} \leq \mathbb{Q} \leq \mathbb{R} \leq \mathbb{C}$;*
- *$\mathbb{Q}^* \leq \mathbb{R}^* \leq \mathbb{C}^*$;*
- *$S_m \leq S_n$ if $m \leq n$;*
- *There are "stupid" subgroups for any group G, given by $\{e\} \leq G$ and $G \leq G$.*

A subgroup inherits the identity and inverses:

Lemma 5.6. *If $H \leq G$ then the identity of H is the identity of G and for any element h in H, the inverse in H is the inverse in G.*

Proof. The identity e_H for H satisfies $e_H \cdot h = h$ for any $h \in H$. In particular, for $h = e_H$ we get: $e_H \cdot e_H = e_H$. For the latter, there is an inverse e_H^{-1} in G, hence

$$e_H = e_H \cdot \underbrace{e_H \cdot e_H^{-1}}_{=e_G} = \underbrace{e_H \cdot e_H^{-1}}_{=e_G} = e_G .$$

For the inverse, suppose h has an inverse g in G and an inverse h' in H. Then $h' = e_G \cdot h' = (gh)h' = g(hh') = ge_G = g$. □

We can give the following criterion for a subset of a group to be a subgroup:

Lemma 5.7. *If H is a non-empty subset of G, then $H \leq G$ if and only if, for all h_1 and h_2 in H, the element $h_1 h_2^{-1}$ is in H.*

Proof. We need to show that a subset satisfying the stated property is in fact a group.

Associativity comes from the fact that $H \subseteq G$ and G is associative.

H is non-empty, so we have some element h in H. Then $hh^{-1} = e$ in H and we have the identity in H.

Let h in H, then, putting $h_1 = e$ and $h_2 = h$, we find $eh^{-1} = h^{-1}$ in H and we have inverses.

Let $h,\ h'$ in H, then h'^{-1} in H and so $h(h'^{-1})^{-1} = hh'$ in H and we have closure. □

For any group G the subset $\{e\}$ is the **trivial subgroup**.

A **proper subgroup** is a subgroup other than the trivial subgroup and the whole group.

Definition 5.8. *Let G be a group.*

(1) *The order of G is the number of elements of G, denoted by $|G|$.*
(2) *Let g be in G. If, for some $n > 0$, $g^n = e$, then g has* **finite order***, and the smallest such n is called the order of g. (In particular, the identity element e has order 1.) Elements not of finite order have* **infinite order***.*

Proposition 5.9. *Let g be in G. Then the set $\langle g\rangle = \{g^k$ with k in $\mathbb{Z}\}$ is a subgroup of G.*

Definition 5.10. *For any $g \in G$, the subgroup $\langle g\rangle \leq G$ is called the* **cyclic subgroup generated by** *g.*

Proof. Certainly $g^1 = g$ is in $\langle g\rangle$ so it is non-empty. If h_1, h_2 are in $\langle g\rangle$ then $h_1 = g^{m_1}$ and $h_2 = g^{m_2}$ for some integers m_1, m_2. Then $h_2^{-1} = g^{-m_2}$, so $h_1 h_2^{-1} = g^{m_1} g^{-m_2} = g^{m_1 - m_2}$, which lies in $\langle g\rangle$. □

Remark 5.11. *If g has finite order r, then $\langle g\rangle = \{e, g, \ldots, g^{r-1}\}$ and those r elements are distinct. In particular, the order of the subgroup generated by g equals the order of g.*

GROUPS OF SMALL ORDER

We shall "classify" all groups of order up to 7. This means that we shall display some standard examples so that any group of order up to 7 is isomorphic to exactly one of these examples. First of all we can deal simply with groups of prime order.

Theorem 5.12. *Any group of prime order is cyclic and generated by any non-identity element.*

We shall leave the proof of this to later, but it does mean that we only have one choice for the standard example when the order is 2, 3, 5 or 7. Examples of cyclic groups are the $\mathcal{C}_n$ that we constructed as rotations in $\mathbb{C}$, or the $\mathbb{Z}_n$ under addition.

For groups of order 4, it turns out that there are just two distinct group structures: the cyclic group of order 4, and the product of two cyclic groups of order 2. For groups of order 6 there are again 2 types. One is the cyclic group of order 6 and the second is S_3. We know these are distinct because the first has the property that $gh = hg$ is always true, whereas this is false in S_3 where if you multiply two transpositions one way you get a different answer from reversing the process: (1 2)(1 3) = (1 3 2), but (1 3)(1 2) = (1 2 3).

Theorem 5.13. (Lagrange) *Let H be a subgroup of a finite group G. Then $|H|$ divides $|G|$.*

Once we have this then another important result becomes clear:

Proof of Theorem 5.12 As p is prime, $p \geq 2$, so $\exists g \in G,\ g \neq e$. We claim that $G =< g >$.

By Lagrange's Theorem $| < g > |$ divides $|G| = p$. But p is prime and so the only divisors are 1 and p. Since $g \neq e$, we do not have $| < g > | = 1$, so $| < g > | = p$ and $< g >$ must be the whole group (it is a subset og G with the same number of elements as G). □

Proof of Theorem 5.13. This is quite straightforward once we have set up the right machinery.

Definition 5.14. *We start with $H \leq G$. Let $g \in G$ and define $gH = \{gh | h \in H\}$ and $Hg = \{hg | h \in H\}$ to be the* **left coset** *and* **right coset** *respectively of H by g.*

We prove a series of small results.

Lemma 5.15. *Let X be a finite subset of a group G, $g \in G$ and define $gX = \{gx | x \in X\}$ and $Xg = \{xg | x \in X\}$. Then both gX and Xg have the same number of elements as X.*

Proof. Consider gX. The only thing that could go wrong is if $x \neq x'$ but $gx = gx'$. But $gx = gx' \Rightarrow g^{-1}gx = g^{-1}gx' \Rightarrow ex = ex' \Rightarrow x = x'$, so this never happens and gX has the same size as X. Similarly for Xg. □

So all the cosets (left and right) of H in G have the same size as H.

Lemma 5.16. *If $gH \cap g'H \neq \emptyset$ then $gH = g'H$. Similarly for right cosets.*

Proof. We shall show that, if the intersection is non-empty, $gH \subseteq g'H$ and $g'H \subseteq gH$.

Suppose that $x \in gH \cap g'H$, so we can find $h,\ h' \in H$ with $x = gh = g'h'$. Then $g = g'h'h^{-1}$ (multiply on the right by h^{-1}).

Let $y \in gH$. Then we can find $h'' \in H$ with $y = gh'' = (g'h'h^{-1})h'' = g'(h'h^{-1}h'')$. But $h,\ h',\ h''$ are all in H which is a group and so the product $h'h^{-1}h'' \in H$ and $y \in g'H$ also. The converse inclusion is exactly similar. □

Lemma 5.17. *G is the union of the set of left (respectively right) cosets of H in G.*

Proof. Let $g \in G$, then $g = ge = eg,\ e \in H$ and so $\forall g \in G,\ g \in gH$ and $g \in Hg$. □

Putting all this together, the left cosets make up all of G, distinct cosets are disjoint and they all have the same size $|H|$. So $|G|$ must be equal to the number of distinct cosets multiplied by $|H|$, hence $|H|$ divides $|G|$. $\square$

Note: The number of left (respectively right) cosets of a subgroup is called the **index** of the subgroup in the group and is denoted $|G : H|$. Thus Lagrange's Theorem can be restated as

$$|G| = |G : H| \, |H| \, .$$