

# Algebra & Number Theory

[13/05/2003]

A. Baker

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF GLASGOW.

*E-mail address:* `a.baker@maths.gla.ac.uk`

*URL:* `http://www.maths.gla.ac.uk/~ajb`





# Contents

Chapter 1. Basic Number Theory	1
1. The natural numbers	1
2. The integers	3
3. The Euclidean Algorithm and the method of back-substitution	4
4. The tabular method	6
5. Congruences	8
6. Primes and factorization	11
7. Congruences modulo a prime	13
8. Finite continued fractions	16
9. Infinite continued fractions	17
10. Diophantine equations	22
11. Pell's equation	23
Problem Set 1	25
Chapter 2. Groups and group actions	29
1. Groups	29
2. Permutation groups	30
3. The sign of a permutation	31
4. The cycle type of a permutation	32
5. Symmetry groups	33
6. Subgroups and Lagrange's Theorem	35
7. Group actions	38
Problem Set 2	43
Chapter 3. Arithmetic functions	47
1. Definition and examples of arithmetic functions	47
2. Convolution and Möbius Inversion	48
Problem Set 3	52
Chapter 4. Finite and infinite sets, cardinality and countability	53
1. Finite sets and cardinality	53
2. Infinite sets	55
3. Countable sets	55
4. Power sets and their cardinality	57
5. The real numbers are uncountable	59
Problem Set 4	60
Index	61

## CHAPTER 1

# Basic Number Theory

### 1. The natural numbers

The *natural numbers*  $0, 1, 2, \dots$  form the most basic type of number and arise when counting elements of finite sets. We denote the set of all natural numbers by

$$\mathbb{N}_0 = \{0, 1, 2, 3, 4, \dots\}$$

and nowadays this is very standard notation. It is perhaps worth remarking that some people exclude 0 from the natural numbers but we will include it since the empty set  $\emptyset$  has 0 elements! We will use the notation  $\mathbb{Z}^+$  for the set of all positive natural numbers

$$\mathbb{Z}^+ = \{n \in \mathbb{N}_0 : n \neq 0\} = \{1, 2, 3, 4, \dots\},$$

which is also often denoted  $\mathbb{N}$ , although some authors also use this to denote our  $\mathbb{N}_0$ .

We can add and multiply natural numbers to obtain new ones, *i.e.*, if  $a, b \in \mathbb{N}_0$ , then  $a + b \in \mathbb{N}_0$  and  $ab \in \mathbb{N}_0$ . Of course we have the familiar properties of these operations such as

$$a + b = b + a, \quad ab = ba, \quad a + 0 = a = 0 + a, \quad a1 = a = 1a, \quad a0 = 0 = 0a, \quad \text{etc.}$$

We can also compare natural numbers using inequalities. Given  $x, y \in \mathbb{N}_0$  exactly one of the following must be true:

$$x = y, \quad x < y, \quad y < x.$$

As usual, if one of  $x = y$  or  $x < y$  holds then we write  $x \leq y$  or  $y \geq x$ . Inequality is *transitive* in the sense that

$$x < y \quad \text{and} \quad y < z \implies x < z.$$

The most subtle aspect of the natural numbers to deal with is the fact that they form an *infinite set*. We can and usually do list the elements of  $\mathbb{N}_0$  in the sequence

$$0, 1, 2, 3, 4, \dots$$

which never ends. One of the most important properties of  $\mathbb{N}_0$  is

**The Well Ordering Principle (WOP):** Every non-empty subset  $S \subseteq \mathbb{N}_0$  contains a least element.

A *least* or *minimal element* of a subset  $S \subseteq \mathbb{N}_0$  is an element  $s_0 \in S$  for which  $s_0 \leq s$  for all  $s \in S$ . Similarly, a *greatest* or *maximal element* of  $S$  is one for which  $s \leq s_0$  for all  $s \in S$ . Notice that  $\mathbb{N}_0$  has a least element 0, but has no greatest element since for each  $n \in \mathbb{N}_0$ ,  $n + 1 \in \mathbb{N}_0$  and  $n < n + 1$ . It is easy to see that least and greatest elements (if they exist) are always unique. In fact, WOP is logically equivalent to each of the two following statements.

**The Principle of Mathematical Induction (PMI):** Suppose that for each  $n \in \mathbb{N}_0$  the statement  $P(n)$  is defined and also the following conditions hold:

- $P(0)$  is true;
- whenever  $P(k)$  is true then  $P(k + 1)$  is true.

Then  $P(n)$  is true for all  $n \in \mathbb{N}_0$ .

**The Maximal Principle (MP):** Let  $T \subseteq \mathbb{N}_0$  be a non-empty subset which is bounded above, *i.e.*, there exists a  $b \in \mathbb{N}_0$  such that for all  $t \in T$ ,  $t \leq b$ . Then  $T$  contains a greatest element. It is easily seen that two greatest elements must agree and we therefore refer to *the* greatest element.

THEOREM 1.1. *The following chain of implications holds*

$$\text{PMI} \implies \text{WOP} \implies \text{MP} \implies \text{PMI}.$$

*Hence these three statements are logically equivalent.*

PROOF.

PMI  $\implies$  WOP: Let  $S \subseteq \mathbb{N}_0$  and suppose that  $S$  has no least element. We will show that  $S = \emptyset$ . Let  $P(n)$  be the statement

$$P(n): k \notin S \text{ for all natural numbers } k \text{ such that } 0 \leq k \leq n.$$

Notice that  $0 \notin S$  since it would be a least element of  $S$ . Hence  $P(0)$  is true.

Now suppose that  $P(n)$  is true. If  $n + 1 \in S$ , then since  $k \notin S$  for  $0 \leq k \leq n$ ,  $n + 1$  would be the least element of  $S$ , contradicting our assumption. Hence,  $n + 1 \notin S$  and so  $P(n + 1)$  is true.

By the PMI,  $P(n)$  is true for all  $n \in \mathbb{N}_0$ . In particular, this means that  $n \notin S$  for all  $n$  and so  $S = \emptyset$ .

WOP  $\implies$  MP: Let  $T \subseteq \mathbb{N}_0$  have upper bound  $b$  and set

$$S = \{s \in \mathbb{N}_0 : t < s \text{ for all } t \in T\}.$$

Then  $S$  is non-empty since for  $t \in T$ ,

$$t \leq b < b + 1,$$

so  $b + 1 \in S$ . If  $s_0$  is a least element of  $S$ , then there must be an element  $t_0 \in T$  such that  $s_0 - 1 \leq t_0$ ; but we also have  $t_0 < s_0$ . Combining these we see that  $s_0 - 1 = t_0 \in T$ . Notice also that for every  $t \in T$ ,  $t < s_0$ , hence  $t \leq s_0 - 1$ . Thus  $t_0$  is the desired greatest element.

MP  $\implies$  PMI: Let  $P(n)$  be a statement for each  $n \in \mathbb{N}_0$ . Suppose that  $P(0)$  is true and for  $n \in \mathbb{N}_0$ ,  $P(n) \implies P(n + 1)$ .

Suppose that there is an  $m \in \mathbb{N}_0$  for which  $P(m)$  is false. Consider the set

$$T = \{t \in \mathbb{N}_0 : P(n) \text{ is true for all natural numbers } n \text{ satisfying } 0 \leq n \leq t\}.$$

Notice that  $T$  is bounded above by  $m$ , since if  $m \leq k$ ,  $k \notin T$ . Let  $t_0$  be the greatest element of  $T$ , which exists thanks to the MP. Then  $P(t_0)$  is true by definition of  $T$ , hence by assumption  $P(t_0 + 1)$  is also true. But then  $P(n)$  is true whenever  $0 \leq n \leq t_0 + 1$ , hence  $t_0 + 1 \in T$ , contradicting the fact that  $t_0$  was the greatest element of  $T$ .

Hence,  $P(n)$  must be true for all  $n \in \mathbb{N}_0$ . □

An important application of these equivalent results is to proving the following property of the natural numbers.

THEOREM 1.2 (Long Division Property). *Let  $n, d \in \mathbb{N}_0$  with  $0 < d$ . Then there are unique natural numbers  $q, r \in \mathbb{N}_0$  satisfying the two conditions  $n = qd + r$  and  $0 \leq r < d$ .*

PROOF. Consider the set

$$T = \{t \in \mathbb{N}_0 : td \leq n\} \subseteq \mathbb{N}_0.$$

Then  $T$  is non-empty since  $0 \in T$ . Also, for  $t \in T$ ,  $t \leq td$ , hence  $t \leq n$ . So  $T$  is bounded above by  $n$  and hence has a greatest element  $q$ . But then  $qd \leq n < (q+1)d$ . Notice that if  $r = n - qd$ , then

$$0 \leq r = n - qd < (q+1)d - qd = d.$$

To prove uniqueness, suppose that  $q', r'$  is a second such pair. Suppose that  $r \neq r'$ . By interchanging the pairs if necessary, we can assume that  $r < r'$ . Since  $n = qd + r = q'd + r'$ ,

$$0 < r' - r = (q - q')d.$$

Notice that this means  $q' \leq q$  since  $d > 0$ . If  $q > q'$ , this implies  $d \leq (q - q')d$ , hence

$$d \leq r' - r < d - r \leq d,$$

and so  $d < d$  which is impossible. So  $q = q'$  which implies that  $r' - r = 0$ , contradicting the fact that  $0 < r' - r$ . So we must indeed have  $q' = q$  and  $r' = r$ .  $\square$

## 2. The integers

The set of integers is  $\mathbb{Z} = \mathbb{Z}^+ \cup \{0\} \cup \mathbb{Z}^- = \mathbb{N}_0 \cup \mathbb{Z}^-$ , where

$$\mathbb{Z}^+ = \{n \in \mathbb{N}_0 : 0 < n\}, \quad \mathbb{Z}^- = \{n : -n \in \mathbb{Z}^+\}.$$

We can add and multiply integers, indeed, they form a basic example of a *commutative ring*. We can generalize the Long Division Property to the integers.

**THEOREM 1.3.** *Let  $n, d \in \mathbb{Z}$  with  $0 \neq d$ . Then there are unique integers  $q, r \in \mathbb{Z}$  for which  $0 \leq r < |d|$  and  $n = qd + r$ .*

**PROOF.** If  $0 < d$ , then we need to show this for  $n < 0$ . By Theorem 1.2, we have unique natural numbers  $q', r'$  with  $0 \leq r' < d$  and  $-n = q'd + r'$ . If  $r' = 0$  then we take  $q = -q'$  and  $r = 0$ . If  $r' \neq 0$  then take  $q = -1 - q'$  and  $r = d - r'$ .

Finally, if  $d < 0$  we can use the above with  $-d$  in place of  $d$  and get  $n = q'(-d) + r$  and then take  $q = -q'$ .

Once again, it is straightforward to verify uniqueness.  $\square$

Given two integers  $m, n \in \mathbb{Z}$  we say that  $m$  *divides*  $n$  and write  $m \mid n$  if there is an integer  $k \in \mathbb{Z}$  such that  $n = km$ ; we also say that  $m$  *is a divisor of*  $n$ . If  $m$  does *not* divide  $n$ , we write  $m \nmid n$ .

Given two integers  $a, b$  not both 0, an integer  $c$  is a *common divisor* or *common factor* of  $a$  and  $b$  if  $c \mid a$  and  $c \mid b$ . A common divisor  $h$  is a *greatest common divisor* or *highest common factor* if for every common divisor  $c$ ,  $c \mid h$ . If  $h, h'$  are two greatest common divisors of  $a, b$ , then  $h \mid h'$  and  $h' \mid h$ , hence we must have  $h' = \pm h$ . For this reason it is standard to refer to *the* greatest common divisor as the positive one. We can then unambiguously write  $\gcd(a, b)$  for this number. Later we will use Long Division to determine  $\gcd(a, b)$ . Then  $a$  and  $b$  are *coprime* if  $\gcd(a, b) = 1$ , or equivalently that the only common divisors are  $\pm 1$ .

There are many useful algebraic properties of greatest common divisors. Here is one while others can be found in Problem Set 1.

**PROPOSITION 1.4.** *Let  $h$  be a common divisor of the integers  $a, b$ . Then for any integers  $x, y$  we have  $h \mid (xa + yb)$ . In particular this holds for  $h = \gcd(a, b)$ .*

**PROOF.** If we write  $a = uh$  and  $b = vh$  for suitable integers  $u, v$ , then

$$xa + yb = xuh + yvh = (xu + yv)h,$$

and so  $h \mid (xa + yb)$  since  $(xu + yv) \in \mathbb{Z}$ .  $\square$

THEOREM 1.5. *Let  $a, b$  be integers, not both 0. Then there are integers  $u, v$  such that*

$$\gcd(a, b) = ua + vb.$$

PROOF. We might as well assume that  $a \neq 0$  and set  $h = \gcd(a, b)$ . Let

$$S = \{xa + yb : x, y \in \mathbb{Z}, 0 < xa + yb\} \subseteq \mathbb{N}_0.$$

Then  $S$  is non-empty since one of  $(\pm 1)a$  is positive and hence is in  $S$ . By the Well Ordering Principle, there is a least element  $d$  of  $S$ , which can be expressed as  $d = u_0a + v_0b$  for some  $u_0, v_0 \in \mathbb{Z}$ .

By Proposition 1.4, we have  $h \mid d$ ; hence all common divisors of  $a, b$  divide  $d$ . Using Long Division we can find  $q, r \in \mathbb{Z}$  with  $0 \leq r < d$  satisfying  $a = qd + r$ . But then

$$r = a - qd = (1 - qu_0)a + (-qv_0)b,$$

hence  $r \in S$  or  $r = 0$ . Since  $r < d$  with  $d$  minimal, this means that  $r = 0$  and so  $d \mid a$ . A similar argument also gives  $d \mid b$ . So  $d$  is a common divisor of  $a, b$  which is divisible by all other common divisors, so it must be the greatest common divisor of  $a, b$ .  $\square$

This result is theoretically useful but does not provide a practical method to determine  $\gcd(a, b)$ . Long Division can be used to set up the *Euclidean Algorithm* which actually determines the greatest common divisor of two non-zero integers.

### 3. The Euclidean Algorithm and the method of back-substitution

Let  $a, b \in \mathbb{Z}$  be non-zero. Set  $n_0 = a$ ,  $d_0 = b$ . Using Long Division, choose integers  $q_0$  and  $r_0$  such that  $0 \leq r_0 < |d_0|$  and  $n_0 = q_0d_0 + r_0$ .

Now set  $n_1 = d_0$ ,  $d_1 = r_0 \geq 0$  and choose integers  $q_1, r_1$  such that  $0 \leq r_1 < d_1$  and  $n_1 = q_1d_1 + r_1$ .

We can repeat this process, at the  $k$ -th stage setting  $n_k = d_{k-1}$ ,  $d_k = r_{k-1}$  and choosing integers  $q_k, r_k$  for which  $0 \leq r_k < d_k$  and  $n_k = q_kd_k + r_k$ . This is always possible provided  $r_{k-1} = d_k \neq 0$ . Notice that

$$0 \leq r_k < r_{k-1} < \cdots < r_1 < r_0 = b,$$

hence we must eventually reach a value  $k = k_0$  for which  $d_{k_0} \neq 0$  but  $r_{k_0} = 0$ .

The sequence of equations

$$\begin{aligned} n_0 &= q_0d_0 + r_0, \\ n_1 &= q_1d_1 + r_1, \\ &\vdots \\ n_{k_0-2} &= q_{k_0-2}d_{k_0-2} + r_{k_0-2}, \\ n_{k_0-1} &= q_{k_0-1}d_{k_0-1} + r_{k_0-1}, \\ n_{k_0} &= q_{k_0}d_{k_0}, \end{aligned}$$

allows us to express each  $r_k = d_{k+1}$  in terms of  $n_k, r_{k-1}$ . For example, we have

$$r_{k_0-1} = n_{k_0-1} - q_{k_0-1}d_{k_0-1} = n_{k_0-1} - q_{k_0-1}r_{k_0-2}.$$

Using this repeatedly, we can write

$$d_{k_0} = un_0 + vr_0 = ua + vb.$$

Thus we can express  $d_{k_0}$  as an integer linear combination of  $a, b$ . By Proposition 1.4 all common divisors of the pair  $a, b$  divide  $d_{k_0}$ . It is also easy to see that

$$d_{k_0} \mid n_{k_0}, \quad d_{k_0-1} \mid n_{k_0-1}, \dots, r_0 \mid n_0,$$

from which it follows that  $d_{k_0}$  also divides  $a$  and  $b$ . Hence the number  $d_{k_0}$  is the greatest common divisor of  $a$  and  $b$ . So the last non-zero remainder term  $r_{k_0-1} = d_{k_0}$  produced by the Euclidean Algorithm is  $\gcd(a, b)$ .

This allows us to express the greatest common divisor of two integers as a linear combination of them by the *method of back-substitution*.

EXAMPLE 1.6. Find the greatest common divisor of 60 and 84 and express it as an integral linear combination of these numbers.

SOLUTION. Since the greatest common divisor only depends on the numbers involved and not their order, we might as take the larger one first, so set  $a = 84$  and  $b = 60$ . Then

$$\begin{aligned} 84 &= 1 \times 60 + 24, & 24 &= 84 + (-1) \times 60, \\ 60 &= 2 \times 24 + 12, & 12 &= 60 + (-2) \times 24, \\ 24 &= 2 \times 12, & 12 &= \gcd(60, 84). \end{aligned}$$

Working back we find

$$\begin{aligned} 12 &= 60 + (-2) \times 24 \\ &= 60 + (-2) \times (84 + (-1) \times 60) \\ &= (-2) \times 84 + 3 \times 60. \end{aligned}$$

Thus

$$\gcd(60, 84) = 12 = 3 \times 60 + (-2) \times 84. \quad \square$$

EXAMPLE 1.7. Find the greatest common divisor of 190 and  $-72$ , and express it as an integral linear combination of these numbers.

SOLUTION. Taking  $a = 190$ ,  $b = -72$  we have

$$\begin{aligned} 190 &= (-2) \times (-72) + 46, & 46 &= 190 + 2 \times (-72), \\ -72 &= (-2) \times 46 + 20, & 20 &= -72 + 2 \times 46, \\ 46 &= 2 \times 20 + 6, & 6 &= -2 \times 20 + 46, \\ 20 &= 3 \times 6 + 2, & 2 &= 20 + (-3) \times 6, \\ 6 &= 3 \times 2, & 2 &= \gcd(190, -72). \end{aligned}$$

Working back we find

$$\begin{aligned} 2 &= 20 + (-3) \times 6 \\ &= 20 + (-3) \times (-2 \times 20 + 46), \\ &= (-3) \times 46 + 7 \times 20, \\ &= (-3) \times 46 + 7 \times (-72 + 2 \times 46), \\ &= 7 \times (-72) + 11 \times 46, \\ &= 7 \times (-72) + 11 \times (190 + 2 \times (-72)), \\ &= 11 \times 190 + 29 \times (-72). \end{aligned}$$

Thus  $\gcd(190, -72) = 2 = 11 \times 190 + 29 \times (-72)$ . □

This could also be done by using the fact that  $\gcd(190, -72) = \gcd(190, 72)$  and proceeding as follows.

EXAMPLE 1.8. Find the greatest common divisor of 190 and 72 and express it as an integral linear combination of these numbers.

SOLUTION. Taking  $a = 190$ ,  $b = 72$  we have

$$\begin{array}{ll}
 190 = 2 \times 72 + 46, & 46 = 190 + (-2) \times 72, \\
 72 = 1 \times 46 + 26, & 26 = 72 + (-1) \times 46, \\
 46 = 1 \times 26 + 20, & 20 = 46 + (-1) \times 26, \\
 26 = 1 \times 20 + 6, & 6 = 26 + (-1) \times 20, \\
 20 = 3 \times 6 + 2, & 2 = 20 + (-3) \times 6, \\
 6 = 3 \times 2, & 2 = \gcd(190, 72).
 \end{array}$$

Working back we find

$$\begin{aligned}
 2 &= 20 + (-3) \times 6 \\
 &= 20 + (-3) \times (26 + (-1) \times 20), \\
 &= (-3) \times 26 + 4 \times 20, \\
 &= (-3) \times 26 + 4 \times (46 + (-1) \times 26), \\
 &= 4 \times 46 + (-7) \times 26, \\
 &= 4 \times 46 + (-7) \times (72 + (-1) \times 46), \\
 &= (-7) \times 72 + 11 \times 46, \\
 &= (-7) \times 72 + 11 \times (190 + (-2) \times 72), \\
 &= 11 \times 190 + (-29) \times 72.
 \end{aligned}$$

Thus  $\gcd(190, 72) = 2 = 11 \times 190 + (-29) \times 72$ . □

From this we obtain  $\gcd(190, -72) = 2 = 11 \times 190 + 29 \times (-72)$ .

It is usually be more straightforward working with positive  $a, b$  and to adjust signs at the end.

Notice that if  $\gcd(a, b) = ua + vb$ , the values of  $u, v$  are not unique. For example,

$$83 \times 190 + 219 \times (-72) = 2.$$

In general, we can modify the numbers  $u, v$  to  $u + tb, v - ta$  since

$$(u + tb)a + (v - ta)b = (ua + vb) + (tba - tab) = (ua + vb).$$

Thus different approaches to determining the linear combination giving  $\gcd(a, b)$  may well produce different answers.

#### 4. The tabular method

This section describes an alternative approach to the problem of expressing  $\gcd(a, b)$  as a linear combination of  $a, b$ . I learnt this method from Francis Clarke of the University of Wales Swansea. The *tabular method* uses the sequence of quotients appearing in the Euclidean Algorithm and is closely related to the continued fraction method of Theorem 1.42. The tabular method provides an efficient alternative to the method of back-substitution and can also be used check calculations done by that method.

We will illustrate the tabular method with an example. In the case  $a = 267$ ,  $b = 207$ , the Euclidean Algorithm produces the following quotients and remainders.

$$267 = 1 \times 207 + 60,$$

$$207 = 3 \times 60 + 27,$$

$$60 = 2 \times 27 + 6,$$

$$27 = 4 \times 6 + 3,$$

$$6 = 2 \times 3 + 0.$$

The last non-zero remainder is 3, so  $\gcd(267, 207) = 3$ . Back-substitution gives

$$\begin{aligned} 3 &= 27 - 4 \times 6 \\ &= 27 - 4 \times (60 - 2 \times 27) \\ &= -4 \times 60 + 9 \times 27 \\ &= -4 \times 60 + 9 \times (207 - 3 \times 60) \\ &= 9 \times 207 - 31 \times 60 \\ &= 9 \times 207 - 31 \times (267 - 1 \times 207) \\ &= (-31) \times 267 + 40 \times 207. \end{aligned}$$

In the tabular method we form the following table.

		1	3	2	4	2
1	0	1	3	7	31	69
0	1	1	4	9	40	89

Here the first row is the sequence of quotients. The second and third rows are determined as follows. The entry  $t_k$  under the quotient  $q_k$  is calculated from the formula

$$t_k = q_k t_{k-1} + t_{k-2}.$$

So for example, 31 arises as  $4 \times 7 + 3$ . The final entries in the second and third rows always have the form  $b/\gcd(a, b)$  and  $a/\gcd(a, b)$ ; here  $207/3 = 69$  and  $267/3 = 89$ . The previous entries are  $\pm A$  and  $\mp B$ , where the signs are chosen according to whether the number of quotients is even or odd.

Why does this give the same result as back-substitution? The arithmetic involved seems very different. In our example, the value 40 arises as  $31 + 9$  in the back-substitution method and as  $4 \times 9 + 4$  in the tabular method.

The key to understanding this is provided by matrix multiplication, in particular the fact that it is associative. Consider the matrix product

$$\begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix}$$

in which the quotients occur as the entries in the bottom right-hand corner. By the associative law, the product can be evaluated either from the right:

$$\begin{aligned} \begin{bmatrix} 0 & 1 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix} &= \begin{bmatrix} 1 & 2 \\ 4 & 9 \end{bmatrix}, \\ \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix} &= \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 4 & 9 \end{bmatrix} = \begin{bmatrix} 4 & 9 \\ 9 & 20 \end{bmatrix}, \\ \begin{bmatrix} 0 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix} &= \begin{bmatrix} 0 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 4 & 9 \\ 9 & 20 \end{bmatrix} = \begin{bmatrix} 9 & 20 \\ 31 & 69 \end{bmatrix}, \\ \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix} &= \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 9 & 20 \\ 31 & 69 \end{bmatrix} = \begin{bmatrix} 31 & 69 \\ 40 & 89 \end{bmatrix}, \end{aligned}$$

or from the left:

$$\begin{aligned} \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 3 \end{bmatrix} &= \begin{bmatrix} 1 & 3 \\ 1 & 4 \end{bmatrix}, \\ \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix} &= \begin{bmatrix} 1 & 3 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 3 & 7 \\ 4 & 9 \end{bmatrix}, \\ \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 4 \end{bmatrix} &= \begin{bmatrix} 3 & 7 \\ 4 & 9 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 4 \end{bmatrix} = \begin{bmatrix} 7 & 31 \\ 9 & 40 \end{bmatrix}, \\ \begin{bmatrix} 0 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 4 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix} &= \begin{bmatrix} 7 & 31 \\ 9 & 40 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 31 & 69 \\ 40 & 89 \end{bmatrix}. \end{aligned}$$

Notice that the numbers occurring as the left-hand columns of the first set of partial products are the same (apart from the signs) as the numbers which arose in the back-substitution method. The numbers in the second set of partial products are those in the tabular method.

Thus back-substitution corresponds to evaluation from the right and the tabular method to evaluation from the left. This shows that they give the same result.

Giving a general proof of this identification of the two methods with matrix multiplication is not too hard. In fact it becomes obvious given the factorization of the matrix  $\begin{bmatrix} 0 & 1 \\ 1 & q \end{bmatrix}$  as the product  $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & q \\ 0 & 1 \end{bmatrix}$  of two elementary matrices. Two elementary row operations are performed when multiplying by  $\begin{bmatrix} 0 & 1 \\ 1 & q \end{bmatrix}$  on the left. Firstly  $q \times (\text{row } 2)$  is added to row 1, then the two rows are swapped. Multiplication by  $\begin{bmatrix} 0 & 1 \\ 1 & q \end{bmatrix}$  on the right performs similar column operations.

The determinant of  $\begin{bmatrix} 0 & 1 \\ 1 & q \end{bmatrix}$  is  $-1$  and so by the multiplicative property of determinants,

$$\det \begin{bmatrix} 0 & 1 \\ 1 & q_1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & q_2 \end{bmatrix} \cdots \begin{bmatrix} 0 & 1 \\ 1 & q_r \end{bmatrix} = (-1)^r.$$

It is this that explains the rule for the choice of signs in the tabular method. The partial products have determinant alternately equal to  $\pm 1$ . This provides a useful check on the calculations.

## 5. Congruences

Let  $n \in \mathbb{N}_0$  be non-zero, so  $n > 0$ . Then for integers  $x, y$ , we say that  $x$  is congruent to  $y$  modulo  $n$  if  $n \mid (x - y)$  and write  $x \equiv y \pmod{n}$  or  $x \equiv_n y$ . Then  $\equiv_n$  is an equivalence relation on

$\mathbb{Z}$  in the sense that the following hold for  $x, y, z \in \mathbb{Z}$ :

$$\begin{aligned} \text{(Reflexivity)} \quad & x \equiv_n x, \\ \text{(Symmetry)} \quad & x \equiv_n y \implies y \equiv_n x, \\ \text{(Transitivity)} \quad & x \equiv_n y \text{ and } y \equiv_n z \implies x \equiv_n z. \end{aligned}$$

The set of equivalence classes is denoted  $\mathbb{Z}/n$ . We will denote the *congruence class* or *residue class* of the integer  $x$  by  $x_n$ ; sometimes notation such as  $\bar{x}$  or  $[x]_n$  is used.

Residue classes can be added and multiplied using the formulæ

$$x_n + y_n = (x + y)_n, \quad x_n y_n = (xy)_n.$$

These make sense because if  $x'_n = x_n$  and  $y'_n = y_n$ , then

$$\begin{aligned} x' + y' &= x + y + (x' - x) + (y' - y) \equiv_n x + y, \\ x'y' &= (x + (x' - x))(y + (y' - y)) = xy + y(x' - x) + x(y' - y) + (x' - x)(y' - y) \equiv_n xy. \end{aligned}$$

We can also define subtraction by  $x_n - y_n = (x - y)_n$ . These operations make  $\mathbb{Z}/n$  into a *commutative ring* with zero  $0_n$  and unity  $1_n$ .

Since for each  $x \in \mathbb{Z}$  we have  $x = qn + r$  with  $q, r \in \mathbb{Z}$  and  $0 \leq r < n$ , we have  $x_n = r_n$ , so we usually list the distinct elements of  $\mathbb{Z}/n$  as

$$0_n, 1_n, 2_n, \dots, (n-1)_n.$$

**THEOREM 1.9.** *Let  $t \in \mathbb{Z}$  have  $\gcd(t, n) = 1$ . Then there is a unique residue class  $u_n \in \mathbb{Z}/n$  for which  $u_n t_n = 1_n$ . In particular, the integer  $u$  satisfies  $ut \equiv_n 1$ .*

**PROOF.** By Theorem 1.5, there are integers  $u, v$  for which  $ut + vn = 1$ . This implies that  $ut \equiv_n 1$ , hence  $u_n t_n = 1_n$ . Notice that if  $w_n$  also has this property then  $w_n t_n = 1_n$  which gives

$$w_n(t_n u_n) = (w_n t_n) u_n = u_n,$$

hence  $w_n = u_n$ . □

We will refer to  $u$  as the *inverse of  $t$  modulo  $n$*  and  $u_n$  as the *inverse of  $t_n$  in  $\mathbb{Z}/n$* . Since  $ut + vn = 1$ , neither  $t$  nor  $u$  can have a common factor with  $n$ .

**EXAMPLE 1.10.** Solve each of the following congruences, in each case giving all (if any) integer solutions:

$$\text{(i) } 5x \equiv_{12} 7; \quad \text{(ii) } 3x \equiv_{101} 6; \quad \text{(iii) } 2x \equiv_{10} 8; \quad \text{(iv) } 2x \equiv_{10} 7.$$

**SOLUTION.**

(i) By use of the Euclidean Algorithm or inspection,  $5^2 = 25 \equiv_{12} 1$ . This gives

$$x \equiv_{12} 5^2 x \equiv_{12} 35 \equiv_{12} 11.$$

(ii) We have  $3 \times 34 = 102 \equiv_{101} 1$ , hence

$$x \equiv_{101} 34 \times 3x \equiv_{101} 34 \times 6 \equiv_{101} 2.$$

(iii) Here  $\gcd(2, 10) = 2$ , so the above method does not immediately apply. We require that  $2(x-4) \equiv_{10} 0$ , giving  $(x-4) \equiv_5 0$  and hence  $x \equiv_5 4$ . So we obtain the solutions  $x \equiv_{10} 4$  and  $x \equiv_{10} 9$ .

(iv) This time we have  $2x \equiv_{10} 7$  so  $2x + 10k = 7$  for some  $k \in \mathbb{Z}$ . This is impossible since  $2 \mid (2x + 10k)$  but  $2 \nmid 7$ , so there are no solutions. □

Another important application is to the simultaneous solution of two or more congruence equations to different moduli. The next Lemma is the key ingredient.

LEMMA 1.11. *Suppose that  $a, b \in \mathbb{N}_0$  are coprime and  $n \in \mathbb{Z}$ . If  $a \mid n$  and  $b \mid n$ , then  $ab \mid n$ .*

PROOF. Let  $a \mid n$  and  $b \mid n$  and choose  $r, s \in \mathbb{Z}$  so that  $n = ra = sb$ . Then if  $ua + vb = 1$ ,

$$n = n(ua + vb) = nua + nvb = su(ab) + rv(ab) = (su + rv)ab.$$

Since  $su + rv \in \mathbb{Z}$ , this implies  $ab \mid n$ . □

THEOREM 1.12 (The Chinese Remainder Theorem). *Suppose  $n_1, n_2 \in \mathbb{Z}^+$  are coprime and  $b_1, b_2 \in \mathbb{Z}$ . Then the pair of simultaneous congruences*

$$x \equiv_{n_1} b_1, \quad x \equiv_{n_2} b_2,$$

*has a unique solution modulo  $n_1n_2$ .*

PROOF. Since  $n_1, n_2$  are coprime, there are integers  $u_1, u_2$  for which  $u_1n_1 + u_2n_2 = 1$ . Consider the integer  $t = u_1n_1b_2 + u_2n_2b_1$ . Then we have the congruences

$$t \equiv_{n_1} u_2n_2b_1 \equiv_{n_1} b_1, \quad t \equiv_{n_2} u_1n_1b_2 \equiv_{n_2} b_2,$$

so  $t$  is a solution for the pair of simultaneous congruences in the Theorem.

To prove uniqueness modulo  $n_1n_2$ , note that if  $t, t'$  are both solutions to the original pair of simultaneous congruences then they satisfy the pair of congruences

$$t' \equiv_{n_1} t, \quad t' \equiv_{n_2} t.$$

By Lemma 1.11,  $n_1n_2 \mid (t' - t)$ , implying that  $t' \equiv_{n_1n_2} t$ , so the solution  $t_{n_1n_2} \in \mathbb{Z}/n_1n_2$  is unique as claimed. □

REMARK 1.13. The general integer solution of the pair of congruences of Theorem 1.12 is

$$x = u_1n_1b_2 + u_2n_2b_1 + kn_1n_2 \quad (k \in \mathbb{Z}).$$

EXAMPLE 1.14. Solve the following pair of simultaneous congruences modulo 28:

$$3x \equiv_4 1, \quad 5x \equiv_7 2.$$

SOLUTION.

Begin by observing that  $3^2 \equiv_4 1$  and  $3 \times 5 = 15 \equiv_7 1$ , hence the original pair of congruences is equivalent to the pair

$$x \equiv_4 3, \quad x \equiv_7 6.$$

Using the Euclidean Algorithm or otherwise we find

$$2 \times 4 + (-1) \times 7 = 1,$$

so the solution modulo 28 is

$$x \equiv_{28} 2 \times 4 \times 6 + (-1) \times 7 \times 3 \equiv_{28} 48 - 21 = 27.$$

Hence the general integer solution is  $27 + 28n$  ( $n \in \mathbb{Z}$ ). □

EXAMPLE 1.15. Find all integer solutions of the three simultaneous congruences

$$7x \equiv_8 1, \quad x \equiv_3 2, \quad x \equiv_5 1.$$

SOLUTION.

We can proceed in two steps.

First solve the pair of simultaneous congruences

$$7x \equiv_8 1, \quad x \equiv_3 2$$

modulo  $8 \times 3 = 24$ . Notice that  $7^2 = 49 \equiv 1 \pmod{8}$ , so the congruences are equivalent to the pair

$$x \equiv 7 \pmod{8}, \quad x \equiv 2 \pmod{3}.$$

Then as  $(-1) \times 8 + 3 \times 3 = 1$ , we have the unique solution

$$(-1) \times 8 \times 2 + 3 \times 3 \times 7 = -16 + 63 = 47 \equiv 23 \equiv -1 \pmod{24}.$$

Now solve the simultaneous congruences

$$x \equiv -1 \pmod{24}, \quad x \equiv 1 \pmod{5}.$$

Notice that  $(-1) \times 24 + 5 \times 5 = 1$ , hence the solution is

$$(-1) \times 24 \times 1 + 5 \times 5 \times (-1) \equiv -24 - 25 \equiv -49 \equiv 71 \pmod{120}.$$

This gives for the general integer solution  $x = 71 + 120n$  ( $n \in \mathbb{Z}$ ). □

## 6. Primes and factorization

**DEFINITION 1.16.** A positive natural number  $p \in \mathbb{N}_0$  for which  $p > 1$  whose only integer factors are  $\pm 1$  and  $\pm p$  is called a *prime*. Otherwise such a natural number is called *composite*.

Some examples of primes are

2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53, 59, 61, 67, 71, 73, 79, 83, 89, 97.

Notice that apart from 2, all primes are odd since every even integer is divisible by 2.

We begin with an important divisibility property of primes.

**THEOREM 1.17** (Euclid's Lemma). *Let  $p$  be a prime and  $a, b \in \mathbb{Z}$ . If  $p \mid ab$ , then  $p \mid a$  or  $p \mid b$ .*

**PROOF.** Suppose that  $p \nmid a$ . Since  $\gcd(p, a) \mid p$ , we have  $\gcd(p, a) = 1$  or  $\gcd(p, a) = p$ ; but the latter implies  $p \mid a$ , contradicting our assumption, thus  $\gcd(p, a) = 1$ . Let  $r, s \in \mathbb{Z}$  be such that  $rp + sa = 1$ . Then  $rp + sab = b$  and so  $p \mid b$ . □

More generally, if a prime  $p$  divides a product of integers  $a_1 \cdots a_n$  then  $p \mid a_j$  for some  $j$ . This can be proved by induction on the number  $n$ .

**THEOREM 1.18** (Fundamental Theorem of Arithmetic). *Let  $n \in \mathbb{N}_0$  be a natural number such that  $n \geq 1$ . Then  $n$  has a unique factorization of the form*

$$n = p_1 p_2 \cdots p_t,$$

where for each  $j$ ,  $p_j$  is a prime and  $2 \leq p_1 \leq p_2 \leq \cdots \leq p_t$ .

**PROOF.** We will prove this using the Well Ordering Principle. Consider the set

$$S = \{n \in \mathbb{N}_0 : 1 \leq n \text{ and no such factorization exists for } n\}$$

Now suppose that  $S \neq \emptyset$ . Then by the WOP,  $S$  has a least element  $n_0$  say. Notice that  $n_0$  cannot be prime since then it have such a factorization. So there must be a factorization  $n_0 = uv$  with  $u, v \in \mathbb{N}_0$  and  $u, v \neq 1$ . Then we have  $1 < u < n_0$  and  $1 < v < n_0$ , hence  $u, v \notin S$  and so there are factorizations

$$u = p_1 \cdots p_r, \quad v = q_1 \cdots q_s$$

for suitable primes  $p_j, q_j$ . From this we obtain

$$n_0 = p_1 \cdots p_r q_1 \cdots q_s,$$

and after reordering and renaming we have a factorization of the desired type for  $n_0$ .

To show uniqueness, suppose that

$$p_1 \cdots p_r = q_1 \cdots q_s$$

for primes  $p_i, q_j$  satisfying  $p_1 \leq p_2 \leq \cdots \leq p_r$  and  $q_1 \leq q_2 \leq \cdots \leq q_s$ . Then  $p_r \mid q_1 \cdots q_s$  and hence  $p_r \mid q_t$  for some  $t = 1, \dots, s$ , which implies that  $p_r = q_t$ . Thus we have

$$p_1 \cdots p_{r-1} = q'_1 \cdots q'_{s-1},$$

where we  $q'_1, \dots, q'_{s-1}$  is the list  $q_1, \dots, q_s$  with the first occurrence of  $q_t$  omitted. Continuing this way, we eventually get down to the case where  $1 = q''_1 \cdots q''_{s-r}$  for some primes  $q''_j$ . But this is only possible if  $s = r$ , *i.e.*, there are no such primes. By considering the sizes of the primes we have

$$p_1 = q_1, p_2 = q_2, \dots, p_r = q_s,$$

which shows uniqueness.  $\square$

We refer to this factorization as the *prime factorization* of  $n$ .

**COROLLARY 1.19.** *Every natural number  $n \geq 1$  has a unique factorization*

$$n = p_1^{r_1} p_2^{r_2} \cdots p_t^{r_t},$$

where for each  $j$ ,  $p_j$  is a prime,  $1 \leq r_j$  and  $2 \leq p_1 < p_2 < \cdots < p_t$ .

We call this factorization the *prime power factorization* of  $n$ .

**PROPOSITION 1.20.** *Let  $a, b \in \mathbb{N}_0$  be non-zero with prime power factorizations*

$$a = p_1^{r_1} \cdots p_k^{r_k}, \quad b = p_1^{s_1} \cdots p_k^{s_k},$$

where  $0 \leq r_j$  and  $0 \leq s_j$ . Then

$$\gcd(a, b) = p_1^{t_1} \cdots p_k^{t_k}$$

with  $t_j = \min\{r_j, s_j\}$ .

**PROOF.** For each  $j$ , we have  $p_j^{t_j} \mid a$  and  $p_j^{t_j} \mid b$ , hence  $p_j^{t_j} \mid \gcd(a, b)$ . Then by Lemma 1.11,  $p_1^{t_1} \cdots p_k^{t_k} \mid \gcd(a, b)$ . If

$$1 < m = \frac{\gcd(a, b)}{p_1^{t_1} \cdots p_k^{t_k}},$$

then  $m \mid \gcd(a, b)$  and there is a prime  $q$  dividing  $m$ , hence  $q \mid a$  and  $q \mid b$ . This means that  $q = p_\ell$  for some  $\ell$  and so  $p_\ell^{t_\ell+1} \mid \gcd(a, b)$ . But then  $p_\ell^{r_\ell+1} \mid a$  and  $p_\ell^{s_\ell+1} \mid b$  which is impossible. Hence  $\gcd(a, b) = p_1^{t_1} \cdots p_k^{t_k}$ .  $\square$

We have not yet considered the question of how many primes there are, in particular whether there are finitely many.

**THEOREM 1.21.** *There are infinitely many distinct primes.*

**PROOF.** Suppose not. Let the distinct primes be  $p_0 = 2, p_1, \dots, p_n$  where

$$2 = p_0 < 3 = p_1 < \cdots < p_n.$$

Consider the natural number  $N = (2p_1 \cdots p_n) + 1$ . Notice that for each  $j$ ,  $p_j \nmid N$ . By the Fundamental Theorem of Arithmetic,  $N = q_1 \cdots q_k$  for some primes  $q_j$ . This gives a contradiction since none of the  $q_j$  can occur amongst the  $p_j$ .  $\square$

We can also show that certain real numbers are not rational.

**PROPOSITION 1.22.** *Let  $p$  be a prime. Then  $\sqrt{p}$  is not a rational number.*

PROOF. Suppose that  $\sqrt{p} = \frac{a}{b}$  for integers  $a, b$ . We can assume that  $\gcd(a, b) = 1$  since common factors can be cancelled. Then on squaring we have  $p = \frac{a^2}{b^2}$  and hence  $a^2 = pb^2$ . Thus  $p \mid a^2$ , and so by Euclid's Lemma 1.17,  $p \mid a$ . Writing  $a = a_1p$  for some integer  $a_1$  we have  $a_1^2p^2 = pb^2$ , hence  $a_1^2p = b^2$ . Again using Euclid's Lemma we see that  $p \mid b$ . Thus  $p$  is a common factor of  $a$  and  $b$ , contradicting our assumption. This means that no such  $a, b$  can exist so  $\sqrt{p}$  is not a rational number.  $\square$

Non-rational real numbers are called *irrational*. The set of all irrational real numbers is much 'bigger' than the set of rational numbers  $\mathbb{Q}$ , see Section 5 of Chapter 4 for details. However it is hard to show that particular real numbers such as  $e$  and  $\pi$  are actually irrational.

## 7. Congruences modulo a prime

In this section,  $p$  will denote a prime number. We will study  $\mathbb{Z}/p$ . We begin by noticing that it makes sense to consider a polynomial with integer coefficients

$$f(x) = a_0 + a_1x + \cdots + a_dx^d \in \mathbb{Z}[x],$$

but reduced modulo  $p$ . If for each  $j$ ,  $a_j \equiv b_j \pmod{p}$ , we write

$$a_0 + a_1x + \cdots + a_dx^d \equiv_p b_0 + b_1x + \cdots + b_dx^d$$

and talk about residue class of a polynomial modulo  $p$ . We will denote the residue class of  $f(x)$  by  $f(x)_p$ . We say that  $f(x)$  has *degree  $d$  modulo  $p$*  if  $a_d \not\equiv 0 \pmod{p}$ .

For an integer  $c \in \mathbb{Z}$ , we can evaluate  $f(c)$  and reduce the answer modulo  $p$ , to obtain  $f(c)_p$ . If  $f(c)_p = 0_p$ , then  $c$  is said to be a *root of  $f(x)$  modulo  $p$* . We will also refer to the residue class  $c_p$  as a root of  $f(x)$  modulo  $p$ .

PROPOSITION 1.23. *If  $f(x)$  has degree  $d$  modulo  $p$ , then the number of distinct roots of  $f(x)$  modulo  $p$  is at most  $d$ .*

PROOF. Begin by noticing that if  $c$  is root of  $f(x)$  modulo  $p$ , then

$$f(x) \equiv_p f(x) - f(c) = (a_1 + a_2(x+c) + \cdots + a_d(x^{d-1} + \cdots + c^{d-1}))(x-c).$$

Hence  $f(x) \equiv_p f_1(x)(x-c)$ . If  $c'$  is another root of  $f(x)$  modulo  $p$  for which  $c'_p \neq c_p$ , then since

$$f_1(c')(c' - c) \equiv_p 0$$

we have  $p \mid f_1(c')(c' - c)$  and so by Euclid's Lemma 1.17,  $p \mid f_1(c')$ ; thus  $c'$  is a root of  $f_1(x)$  modulo  $p$ .

If now the integers  $c = c_1, c_2, \dots, c_k$  are roots of  $f(x)$  modulo  $p$  which are all distinct modulo  $p$ , then

$$f(x) \equiv_p (x - c_1)(x - c_2) \cdots (x - c_k)g(x).$$

In fact, the degree of  $g(x)$  is then  $d - k$ . This implies that  $0 \leq k \leq d$ .  $\square$

THEOREM 1.24 (Fermat's Little Theorem). *Let  $t \in \mathbb{Z}$ . Then  $t$  is a root of the polynomial  $\Phi_p(x) = x^p - x$  modulo  $p$ . Moreover, if  $t_p \neq 0_p$ , then  $t$  is a root of the polynomial  $\Phi_p^0(x) = x^{p-1} - 1$  modulo  $p$ .*

PROOF. Consider the function

$$\varphi: \mathbb{Z} \longrightarrow \mathbb{Z}/p; \quad \varphi(t) = (t^p - t)_p.$$

Notice that if  $s \equiv t \pmod p$  then  $\varphi(s) = \varphi(t)$  since  $s^p - s \equiv t^p - t \pmod p$ . Then for  $u, v \in \mathbb{Z}$ ,  $\varphi$  has the following additivity property:

$$\varphi(u + v) = \varphi(u) + \varphi(v).$$

To see this, notice that the Binomial Theorem gives

$$(u + v)^p = u^p + v^p + \sum_{j=1}^{p-1} \binom{p}{j} u^j v^{p-j}.$$

For  $1 \leq j \leq p-1$ ,

$$\binom{p}{j} = \frac{p \cdot (p-1)!}{j!(p-j)!}$$

and as none of  $j!$ ,  $(p-j)!$ ,  $(p-1)!$  is divisible by  $p$ , the integer  $\binom{p}{j}$  is so divisible. This gives the following useful result.

**THEOREM 1.25 (Idiot's Binomial Theorem).** *For a prime  $p$  and  $u, v \in \mathbb{Z}$ ,*

$$(u + v)^p \equiv u^p + v^p \pmod p.$$

From this we deduce

$$\begin{aligned} (u + v)^p - (u + v) &\equiv (u^p + v^p) - (u + v) \pmod p \\ &\equiv (u^p - u) + (v^p - v) \pmod p. \end{aligned}$$

It follows by Induction on  $n$  that for  $n \geq 1$ ,

$$\varphi(u_1 + \cdots + u_n) = \varphi(u_1) + \cdots + \varphi(u_n).$$

To prove Fermat's Little Theorem, notice that  $\varphi(1) = 0_p$  and so for  $t \geq 1$ ,

$$\varphi(t) = \varphi(\underbrace{1 + \cdots + 1}_{t \text{ summands}}) = \underbrace{\varphi(1) + \cdots + \varphi(1)}_{t \text{ summands}} = \underbrace{0_p + \cdots + 0_p}_{t \text{ summands}} = 0_p.$$

For general  $t \in \mathbb{Z}$ , we have  $\varphi(t) = \varphi(t + kp)$  for  $k \in \mathbb{N}_0$ , so we can replace  $t$  by a positive natural number congruent to it and then use the above argument.

If  $t_p \neq 0_p$ , then we have  $p \mid t(t^{p-1} - 1)$  and so by Euclid's Lemma 1.17,  $p \mid (t^{p-1} - 1)$ .  $\square$

The second part of Fermat's Little Theorem can be used to elucidate the multiplicative structure of  $\mathbb{Z}/p$ .

Let  $t$  be an integer not divisible by  $p$ . By Theorem 1.9, since  $\gcd(t, p) = 1$ , there is an inverse  $u$  of  $t$  modulo  $p$ . The set

$$P_t = \{t_p^k : k \geq 1\} \subseteq \mathbb{Z}/p$$

is finite with at most  $p-1$  elements. Notice that in particular we must have  $t_p^r = t_p^s$  for some  $r < s$  and so  $t_p^{s-r} = 1_p$ . The *order of  $t$  modulo  $p$*  is the smallest  $d > 0$  such that  $t_p^d \equiv 1$ . We denote the order of  $t$  by  $\text{ord}_p t$ . Notice that the order is always in the range  $1 \leq \text{ord}_p t \leq p-1$ .

**LEMMA 1.26.** *For  $t \in \mathbb{Z}$  with  $p \nmid t$ , the order of  $t$  modulo  $p$  divides  $p-1$ . Moreover, for  $k \in \mathbb{N}_0$ ,  $t_p^k \equiv 1$  if and only if  $\text{ord}_p t \mid k$ .*

**PROOF.** Let  $d = \text{ord}_p t$  be the order of  $t$  modulo  $p$ . Writing  $p-1 = qd + r$  with  $0 \leq r < d$ , we have

$$1 \equiv t_p^{p-1} \equiv t_p^{qd+r} = t_p^{qd} t_p^r \equiv t_p^r,$$

which means that  $r = 0$  since  $d$  is the least positive integer with this property.

If  $t^k \equiv 1 \pmod{p}$ , then writing  $k = q'd + r'$  with  $0 \leq r' < d$ , we have

$$1 \equiv t^{q'd+r'} \equiv t^{r'} \pmod{p},$$

hence  $r' = 0$  by the minimality of  $d$ . So  $d \mid k$ .  $\square$

**THEOREM 1.27.** *For a prime  $p$ , there is an integer  $g$  such that  $\text{ord}_p g = p - 1$ .*

**PROOF.** Proofs of this result can be found in many books on elementary Number Theory. It is also a consequence of our Theorem 2.28.  $\square$

Such an integer  $g$  is called a *primitive root modulo  $p$* . The distinct powers of  $g$  modulo  $p$  are then the  $(p - 1)$  residue classes

$$1_p = g_p^0, g_p, g_p^2, \dots, g_p^{p-2}.$$

This implies the following result.

**PROPOSITION 1.28.** *Let  $g$  be a primitive root modulo the prime  $p$ . Then for any integer  $t$  with  $p \nmid t$ , there is a unique integer  $r$  such that  $0 \leq r < p - 1$  and  $t \equiv g^r \pmod{p}$ .*

Notice that the power  $g^{(p-1)/2}$  satisfies  $(g^{(p-1)/2})^2 \equiv 1 \pmod{p}$ . Since this number is not congruent to 1 modulo  $p$ , Proposition 1.23 implies that  $g^{(p-1)/2} \equiv -1 \pmod{p}$ .

**PROPOSITION 1.29.** *If  $p$  is an odd prime then the polynomial  $x^2 + 1$  has*

- *no roots modulo  $p$  if  $p \equiv 3 \pmod{4}$ ,*
- *two roots modulo  $p$  if  $p \equiv 1 \pmod{4}$ .*

**PROOF.** Let  $g$  be a primitive root modulo  $p$ .

If  $p \equiv 3 \pmod{4}$ , suppose that  $u^2 + 1 \equiv 0 \pmod{p}$ . Then if  $u \equiv g^r \pmod{p}$ , we have  $g^{2r} \equiv -1 \pmod{p}$ , hence  $g^{2r} \equiv g^{(p-1)/2} \pmod{p}$ . But then  $(p - 1) \mid (2r - (p - 1)/2)$  which is impossible since  $(p - 1)/2$  is odd.

If  $p \equiv 1 \pmod{4}$ ,  $(g^{(p-1)/4})^4 - 1 \equiv 0 \pmod{p}$ , so the polynomial  $x^4 - 1$  has four distinct roots modulo  $p$ , namely

$$1_p, -1_p, g_p^{(p-1)/4}, g_p^{3(p-1)/4}.$$

By Proposition 1.23, this means that  $g^{(p-1)/4}, g^{3(p-1)/4}$  are roots of  $x^2 + 1$  modulo  $p$ .  $\square$

**THEOREM 1.30 (Wilson's Theorem).** *For a prime  $p$ ,*

$$(p - 1)! \equiv -1 \pmod{p}.$$

**PROOF.** This is trivially true when  $p = 2$ , so assume that  $p$  is odd. By Fermat's Little Theorem 1.24, the polynomial  $x^{p-1} - 1$  has for its  $p - 1$  distinct roots modulo  $p$  the numbers  $1, 2, \dots, p - 1$ . Thus

$$(x - 1)(x - 2) \cdots (x - p + 1) \equiv x^{p-1} - 1 \pmod{p}.$$

By setting  $x = 0$  we obtain

$$(-1)^{p-1}(p - 1)! \equiv -1 \pmod{p}.$$

As  $(p - 1)$  is even, the result follows.  $\square$

### 8. Finite continued fractions

Let  $a, b \in \mathbb{Z}$  with  $b > 0$ . If the Euclidean Algorithm for these integers produces the sequence

$$\begin{aligned} a &= q_0 b + r_0, \\ b &= q_1 r_0 + r_1, \\ r_0 &= q_1 r_1 + r_2, \\ &\vdots \\ r_{k_0-2} &= q_{k_0-1} r_{k_0-1} + r_{k_0}, \\ r_{k_0-1} &= q_{k_0} r_{k_0}. \end{aligned}$$

Then

$$\frac{a}{b} = q_0 + \frac{r_0}{b} = q_0 + \frac{1}{b/r_0} = q_0 + \frac{1}{q_1 + \frac{1}{q_2 + \cdots + \frac{1}{q_{k_0-1} + \frac{1}{q_{k_0}}}}}$$

and this expression is called the *continued fraction expansion* of  $a/b$ , written  $[q_0; q_1, \dots, q_{k_0}]$ ; we also say that  $[q_0; q_1, \dots, q_{k_0}]$  *represents*  $a/b$ .

In general,  $[a_0; a_1, a_2, a_3, \dots, a_n]$  gives a *finite continued fraction* if each  $a_k$  is an integer with all except possibly  $a_0$  being positive. Then

$$[a_0; a_1, a_2, a_3, \dots, a_n] = a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \cdots}}}$$

Notice that this expansion for  $a/b$  is not necessarily unique since if  $q_{k_0} > 1$ , then  $q_{k_0} = (q_{k_0} - 1) + 1$  and we obtain the different expansion

$$\frac{a}{b} = q_0 + \frac{r_0}{b} = q_0 + \frac{1}{b/r_0} = q_0 + \frac{1}{q_1 + \frac{1}{q_2 + \cdots + \frac{1}{q_{k_0-1} + \frac{1}{(q_{k_0} - 1) + \frac{1}{1}}}}}$$

which shows that  $[q_0; q_1, \dots, q_{k_0}] = [q_0; q_1, \dots, q_{k_0} - 1, 1]$ . For example,

$$\begin{aligned} \frac{21}{13} &= 1 + \frac{8}{13} = 1 + \frac{1}{\frac{13}{8}} = 1 + \frac{1}{1 + \frac{5}{8}} = 1 + \frac{1}{1 + \frac{1}{1 + \frac{3}{5}}} \\ &= 1 + \frac{1}{1 + \frac{1}{1 + \frac{1}{1 + \frac{2}{3}}}} = 1 + \frac{1}{1 + \frac{1}{1 + \frac{1}{1 + \frac{1}{1 + \frac{1}{2}}}}} = 1 + \frac{1}{1 + \frac{1}{1 + \frac{1}{1 + \frac{1}{1 + \frac{1}{1}}}}} \end{aligned}$$

so  $21/13 = [1; 1, 1, 1, 1, 2] = [1; 1, 1, 1, 1, 1, 1]$ . Analogous considerations show that every rational number has exactly two such continued fraction expansions related in a similar fashion.

The *convergents* of the above continued fraction expansion are the numbers

$$\begin{aligned} A_0 &= 1, & A_1 &= 1 + \frac{1}{1} = 2, & A_2 &= 1 + \frac{1}{1 + \frac{1}{1}} = \frac{3}{2}, & A_3 &= 1 + \frac{1}{1 + \frac{1}{1 + \frac{1}{1}}} = \frac{5}{3}, \\ A_4 &= 1 + \frac{1}{1 + \frac{1}{1 + \frac{1}{1 + \frac{1}{1}}}} = \frac{8}{5}, & A_5 &= 1 + \frac{1}{1 + \frac{1}{1 + \frac{1}{1 + \frac{1}{1 + \frac{1}{2}}}}} = \frac{21}{13}, \end{aligned}$$

which form a sequence tending to  $21/13$ . They also satisfy the inequalities

$$A_0 < A_2 < A_4 < A_5 < A_3 < A_1.$$

In general, the even convergents of a finite continued fraction expansion always form a strictly increasing sequence, while the odd ones form a strictly decreasing sequence.

### 9. Infinite continued fractions

The continued fraction expansions considered so far are all *finite*, however *infinite continued fraction* (icf) expansions turn out to be interesting too. Such an infinite continued fraction expansion has the form

$$[a_0; a_1, a_2, a_3, \dots] = a_0 + \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{a_3 + \dots}}}$$

where  $a_0, a_1, a_2, a_3, \dots$  are integers with all except possibly  $a_0$  being positive. Of course, we might expect to have to consider questions of convergence for such an infinite expansion and we will discuss this point later.

EXAMPLE 1.31. Assuming it makes sense, what real number  $\alpha$  must the following infinite continued fraction  $[1; 1, 1, 1, \dots]$  represent?

SOLUTION. If

$$\alpha = [1; 1, 1, 1, \dots] = 1 + \frac{1}{1 + \frac{1}{1 + \frac{1}{1 + \dots}}}$$

then

$$\alpha = 1 + \frac{1}{\alpha}, \quad \text{i.e., } \alpha^2 - \alpha - 1 = 0,$$

which has solutions  $\alpha = \frac{1 \pm \sqrt{5}}{2}$ . It is 'obvious' that  $\alpha > 0$ , hence  $\alpha = (1 + \sqrt{5})/2$ .  $\square$

Let  $A = [a_0; a_1, a_2, a_3, \dots]$  be an infinite continued fraction expansion. Then for each  $k \geq 0$ , the finite continued fraction  $A_k = [a_0; a_1, a_2, a_3, \dots, a_k]$  is called the  $k$ -th convergent of  $A$ . In Example 1.31, the first few convergents are

$$\begin{aligned} A_0 &= \frac{1}{1}, & A_1 &= 1 + \frac{1}{1} = \frac{2}{1}, & A_2 &= 1 + \frac{1}{1 + \frac{1}{1}} = \frac{3}{2}, \\ A_3 &= 1 + \frac{1}{1 + \frac{1}{1 + \frac{1}{1}}} = \frac{5}{3}, & A_4 &= 1 + \frac{1}{1 + \frac{1}{1 + \frac{1}{1 + \frac{1}{1}}}} = \frac{8}{5}. \end{aligned}$$

Here the numerators and denominators form the famous *Fibonacci sequence*  $\{u_n\}$ ,

$$1, 1, 2, 3, 5, 8, \dots$$

which is given by the *recurrence relation*

$$u_1 = u_2 = 1, \quad u_n = u_{n-1} + u_{n-2} \quad (n \geq 3).$$

Using the convergents of a continued fraction, we might define  $A = [a_0; a_1, a_2, a_3, \dots]$  to be  $\lim_{n \rightarrow \infty} A_n$ , provided this limit exists. We will show that such limits do always exist and we will then say that  $A = [a_0; a_1, a_2, a_3, \dots]$  represents the value of this limit.

The first few convergents of  $A = [a_0; a_1, a_2, a_3, \dots]$  are

$$\begin{aligned} A_0 &= \frac{a_0}{1}, \\ A_1 &= \frac{a_1 a_0 + 1}{a_1}, \\ A_2 &= \frac{a_2(a_1 a_0 + 1) + a_0}{a_2 a_1 + 1}, \\ A_3 &= \frac{a_3(a_2 a_1 a_0 + a_2 + a_0) + a_1 a_0 + 1}{a_3(a_2 a_1 + 1) + a_1}. \end{aligned}$$

The general pattern is given in the next result.

**THEOREM 1.32.** *Given the infinite continued fraction  $A = [a_0; a_1, a_2, a_3, \dots]$ , set  $p_0 = a_0$ ,  $q_0 = 1$ ,  $p_1 = a_1 a_0 + 1$ ,  $q_1 = a_1$ , while for  $n \geq 2$ ,*

$$p_n = a_n p_{n-1} + p_{n-2}, \quad q_n = a_n q_{n-1} + q_{n-2}.$$

*Then for each  $n \geq 0$  the  $n$ -th convergent of  $[a_0; a_1, a_2, a_3, \dots]$  is  $A_n = \frac{p_n}{q_n}$ .*

In the proof and later in this section we will make use of *generalized finite continued fractions*  $[a_0; a_1, a_2, a_3, \dots, a_{n-1}, a_n]$  for which  $a_0 \in \mathbb{Z}$ ,  $0 < a_k \in \mathbb{N}_0$ ,  $1 \leq k \leq n-1$ , and  $0 < a_n \in \mathbb{R}$ .

PROOF. The cases  $n = 0, 1, 2$  clearly hold. We will prove the result by Induction on  $n$ . Suppose that for some  $k \geq 2$ ,  $A_k = \frac{p_k}{q_k}$ . Then

$$A_{k+1} = [a_0; a_1, a_2, a_3, \dots, a_k, a_{k+1}] = [a_0; a_1, a_2, a_3, \dots, a_k + 1/a_{k+1}],$$

which gives us the inductive step

$$\begin{aligned} A_{k+1} &= \frac{(a_k + 1/a_{k+1})p_k + p_{k-1}}{(a_k + 1/a_{k+1})q_k + q_{k-1}} \\ &= \frac{a_{k+1}(a_k p_{k-1} + p_{k-2}) + p_{k-1}}{a_{k+1}(a_k q_{k-1} + q_{k-2}) + q_{k-1}} \\ &= \frac{a_{k+1}p_k + p_{k-1}}{a_{k+1}q_k + q_{k-1}} \\ &= \frac{p_{k+1}}{q_{k+1}}. \end{aligned} \quad \square$$

COROLLARY 1.33. *The convergents of  $A = [a_0; a_1, a_2, a_3, \dots]$  satisfy*

i) *for  $n \geq 1$ ,*

$$p_n q_{n-1} - p_{n-1} q_n = (-1)^{n-1}, \quad A_n - A_{n-1} = \frac{(-1)^{n-1}}{q_{n-1} q_n};$$

ii) *for  $n \geq 2$ ,*

$$p_n q_{n-2} - p_{n-2} q_n = (-1)^n a_n, \quad A_n - A_{n-2} = \frac{(-1)^n a_n}{q_{n-2} q_n}.$$

PROOF. We will use Induction on  $n$ . We can easily verify the cases  $n = 1, 2$ . Assume that the equations hold when  $n = k$  for some  $k \geq 2$ . Then

$$\begin{aligned} p_{k+1} q_k - p_k q_{k+1} &= (a_{k+1} p_k + p_{k-1}) q_k - p_k (a_{k+1} q_k + q_{k-1}) \\ &= p_{k-1} q_k - p_k q_{k-1} \\ &= (-1)(-1)^{k-1} = (-1)^k, \end{aligned}$$

giving the inductive step required to prove (i). Similarly, for (ii) we have

$$\begin{aligned} p_k q_{k-2} - p_{k-2} q_k &= (a_k p_{k-1} + p_{k-2}) q_{k-2} - p_{k-2} (a_k q_{k-1} + q_{k-2}) \\ &= a_k (p_{k-1} q_{k-2} - p_{k-2} q_{k-1}) \\ &= a_k (-1)^{k-2} = (-1)^k a_k. \end{aligned} \quad \square$$

COROLLARY 1.34. *The convergents of  $A = [a_0; a_1, a_2, a_3, \dots]$  satisfy the inequalities*

$$A_{2r} < A_{2r+2s} < A_{2r+2s-1} < A_{2s-1}$$

*for all integers  $r, s$  with  $s > 0$ . Hence each  $A_{2m}$  is less than each  $A_{2n-1}$  and the sequence  $\{A_{2n}\}$  is strictly increasing while the sequence  $\{A_{2n-1}\}$  is strictly decreasing, i.e.,*

$$A_0 < A_2 < \dots < A_{2m} < \dots < A_{2n-1} < \dots < A_3 < A_1.$$

THEOREM 1.35. *The convergents of the infinite continued fraction  $[a_0; a_1, a_2, a_3, \dots]$  form a sequence  $\{A_n\}$  which has a limit  $A = \lim_{n \rightarrow \infty} A_n$ .*

PROOF. Notice that the increasing sequence  $\{A_{2n}\}$  is bounded above by  $A_1$ , hence it has a limit  $\ell$  say. Similarly, the decreasing sequence  $\{A_{2n-1}\}$  is bounded below by  $A_0$ , hence it has a limit  $u$  say. Notice that

$$\ell = \lim_{n \rightarrow \infty} A_{2n} \leq \lim_{n \rightarrow \infty} A_{2n-1} = u.$$

In fact,

$$u - \ell = \lim_{n \rightarrow \infty} A_{2n-1} - \lim_{n \rightarrow \infty} A_{2n} = \lim_{n \rightarrow \infty} (A_{2n-1} - A_{2n}) = \lim_{n \rightarrow \infty} \frac{1}{q_{2n-1} q_{2n}}.$$

Notice that for  $n \geq 1$  we have  $a_k > 0$  and hence  $q_k < q_{k+1}$ . Since  $q_k \in \mathbb{Z}$ ,  $\lim_{n \rightarrow \infty} \frac{1}{q_n} = 0$ , hence  $u - \ell = 0$ . Thus  $\lim_{n \rightarrow \infty} A_n$  exists and is equal to  $\lim_{n \rightarrow \infty} A_{2n} = \lim_{n \rightarrow \infty} A_{2n-1}$ .  $\square$

EXAMPLE 1.36. Determine the real number which is represented by the infinite continued fraction  $[1; 2, 2, 2, \dots]$  and calculate its first few convergents.

SOLUTION. Let  $\gamma$  be this number. Then

$$\gamma - 1 = \frac{1}{1 + \gamma},$$

giving the equation  $\gamma^2 - 1 = 1$ . Thus  $\gamma = \pm\sqrt{2}$  and since  $\gamma$  is clearly positive, we get  $\gamma = \sqrt{2}$ .

We have  $a_0 = 1$ ,  $2 = a_1 = a_2 = a_3 = \dots$ , giving  $p_0 = 1$ ,  $q_0 = 1$ ,  $p_1 = 3$ ,  $q_1 = 2$  and for  $n \geq 2$ ,

$$p_n = 2p_{n-1} + p_{n-2}, \quad q_n = 2q_{n-1} + q_{n-2}.$$

The first few convergents are

$$A_0 = 1, \quad A_1 = \frac{3}{2}, \quad A_2 = \frac{7}{5}, \quad A_3 = \frac{17}{12}, \quad A_4 = \frac{41}{29}, \quad A_5 = \frac{99}{70}, \quad A_6 = \frac{239}{169}, \quad A_7 = \frac{577}{408}. \quad \square$$

THEOREM 1.37. *Each irrational number  $\gamma$  has a unique representation as an infinite continued fraction expansion  $[c_0; c_1, c_2, \dots]$  for which  $c_j \in \mathbb{Z}$  with  $c_j > 0$  if  $j > 0$ .*

PROOF. We begin by setting  $\gamma_0 = \gamma$  and  $c_0 = [\gamma_0]$ . Then if

$$\gamma_1 = \frac{1}{\gamma_0 - c_0},$$

we can define  $c_1 = [\gamma_1]$ . Continuing in this way, we can inductively define sequences of real numbers  $\gamma_n$  and integers  $c_n$  satisfying

$$\gamma_n = \frac{1}{\gamma_{n-1} - c_{n-1}}, \quad c_n = [\gamma_n].$$

Notice that for  $n > 0$ ,  $c_n > 0$ . Also, if  $\gamma_n$  is rational then so is  $\gamma_{n-1}$  since

$$\gamma_{n-1} = c_{n-1} + \frac{1}{\gamma_n},$$

and this would imply that  $\gamma_0$  was rational which is false. In particular this shows that  $\gamma_n \neq 0$  at each stage and  $\gamma_n > c_n$ .

Using the generalized continued fraction notation we have  $\gamma = [c_0; c_1, \dots, c_n, \gamma_{n+1}]$  with convergents satisfying the conditions

$$C_n = \frac{p_n}{q_n}, \quad \gamma = \frac{\gamma_{n+1}p_n + p_{n-1}}{\gamma_{n+1}q_n + q_{n-1}}.$$

Then

$$\begin{aligned} |\gamma - C_n| &= \left| \frac{\gamma_{n+1}p_n + p_{n-1}}{\gamma_{n+1}q_n + q_{n-1}} - \frac{p_n}{q_n} \right| \\ &= \left| \frac{p_{n-1}q_n - p_nq_{n-1}}{(\gamma_{n+1}q_n + q_{n-1})q_n} \right| \\ &= \left| \frac{(-1)^n}{(\gamma_{n+1}q_n + q_{n-1})q_n} \right| \\ &= \frac{1}{q_{n+1}q_n} < \frac{1}{q_n^2}. \end{aligned}$$

Since the  $q_n$  form a strictly increasing sequence of integers,  $1/q_n^2 \rightarrow 0$  as  $n \rightarrow \infty$ , hence  $C_n \rightarrow \gamma$ . Thus the infinite continued fraction  $[c_0; c_1, c_2, \dots]$  represents  $\gamma$ .

It is easy to see that if  $\gamma$  is represented by the infinite continued fraction  $[a_0; a_1, a_2, \dots]$  then  $a_0 = [\gamma]$ , and in general  $c_n = a_n$  for all  $n$ , hence this representation is unique.  $\square$

EXAMPLE 1.38. Find the continued fraction expansion of  $\sqrt{2}$ .

SOLUTION. Let  $\gamma_0 = \sqrt{2}$  and so  $c_0 = [\sqrt{2}] = 1$ . Then

$$\gamma_1 = \frac{1}{\sqrt{2} - 1} = \frac{\sqrt{2} + 1}{2 - 1} = \sqrt{2} + 1$$

and so  $c_1 = [\sqrt{2} + 1] = 2$ . Repeating this gives

$$\gamma_2 = \frac{1}{\sqrt{2} + 1 - 2} = \frac{1}{\sqrt{2} - 1} = \sqrt{2} + 1$$

and  $c_2 = [\sqrt{2} + 1] = 2$ . Clearly we get for each  $n > 0$ ,

$$\gamma_n = \frac{1}{\sqrt{2} - 1} = \sqrt{2} + 1, \quad c_n = 2.$$

So the infinite continued fraction representing  $\sqrt{2}$  is  $[1; 2, 2, \dots] = [1; \overline{2}]$ , where  $\overline{2}$  means 2 repeated infinitely often.  $\square$

We will write  $\overline{a_1, a_2, \dots, a_p}$  to denote the sequence  $a_1, a_2, \dots, a_p$  repeated infinitely often as in the last example.

EXAMPLE 1.39. Find the continued fraction expansion of  $\sqrt{3}$ .

SOLUTION. Let  $\gamma_0 = \sqrt{3}$  and so  $c_0 = [\sqrt{3}] = 1$ . Then

$$\gamma_1 = \frac{1}{\sqrt{3} - 1} = \frac{\sqrt{3} + 1}{3 - 1} = \frac{\sqrt{3} + 1}{2}$$

and so  $c_1 = 1$ . Repeating gives

$$\gamma_2 = \frac{1}{\gamma_1 - c_1} = \frac{2}{\sqrt{3} - 1} = \frac{2(\sqrt{3} + 1)}{3 - 1} = \sqrt{3} + 1$$

and  $c_2 = 2$ . Repeating again gives

$$\gamma_3 = \frac{1}{\gamma_2 - c_2} = \frac{1}{\sqrt{3} - 1} = \frac{\sqrt{3} + 1}{3 - 1} = \frac{\sqrt{3} + 1}{2} = \gamma_1$$

and  $c_3 = 1 = c_1$ . From now on this pattern repeats giving

$$\gamma_n = \begin{cases} \frac{\sqrt{3} + 1}{2} & \text{if } n \text{ is odd,} \\ \sqrt{3} + 1 & \text{if } n \text{ is even,} \end{cases} \quad c_n = \begin{cases} 1 & \text{if } n \text{ is odd,} \\ 2 & \text{if } n \text{ is even.} \end{cases}$$

So the infinite continued fraction representing  $\sqrt{3}$  is  $[1; 1, 2, 1, 2, \dots] = [1, \overline{1, 2}]$ . The first few convergents are

$$1, 2, \frac{5}{3}, \frac{7}{4}, \frac{19}{11}, \frac{26}{15}, \frac{71}{41}, \frac{97}{56}. \quad \square$$

This example illustrates a general phenomenon.

THEOREM 1.40. *For a natural number  $n$  which is not a square, the irrational number  $\sqrt{n}$  has an infinite continued fraction expansion of the form  $[a_0; \overline{a_1, a_2, \dots, a_p}]$ .*

*Furthermore, if  $p$  is the smallest such number, then the continued fraction expansion of  $\sqrt{n}$  also has the symmetry*

$$\sqrt{n} = [a_0; \overline{a_1, a_2, \dots, a_p}] = [a_0; \overline{a_1, a_2, \dots, a_2, a_1, 2a_0}].$$

The smallest  $p$  for which the expansion has periodic part of length  $p$  is called the *period* of the continued fraction expansion of  $\sqrt{n}$ . Here are some more examples whose periods are indicated.

$$\begin{aligned}\sqrt{5} &= [2; \overline{4}] && \text{(period 1)} \\ \sqrt{6} &= [2; \overline{2, 4}] && \text{(period 2)} \\ \sqrt{7} &= [2; \overline{1, 1, 1, 4}] && \text{(period 4)} \\ \sqrt{8} &= [2; \overline{1, 4}] && \text{(period 2)} \\ \sqrt{10} &= [3; \overline{6}] && \text{(period 1)} \\ \sqrt{11} &= [3; \overline{3, 6}] && \text{(period 2)} \\ \sqrt{12} &= [3; \overline{2, 6}] && \text{(period 2)} \\ \sqrt{13} &= [3; \overline{1, 1, 1, 1, 6}] && \text{(period 5)} \\ \sqrt{97} &= [9; \overline{1, 5, 1, 1, 1, 1, 1, 1, 5, 1, 18}] && \text{(period 11)}\end{aligned}$$

## 10. Diophantine equations

Consider the following problem:

Find all integer solutions  $x, y$  of the equation  $35x + 61y = 1$ .

Such problems in which we are only interested in *integer* solutions are called *Diophantine problems* and are named after the Greek *Diophantus* in whose book many examples appeared. Diophantine problems were also studied in several ancient civilizations including those of China, India and the Middle East.

Since  $\gcd(35, 61) = 1$ , we can use the Euclidean Algorithm to find a specific solution of this problem.

$$\begin{aligned}61 &= 1 \times 35 + 26, & 26 &= 61 + (-1) \times 35, \\ 35 &= 1 \times 26 + 9, & 9 &= 35 + (-1) \times 26, \\ 26 &= 2 \times 9 + 8, & 8 &= 26 + (-2) \times 9, \\ 9 &= 1 \times 8 + 1, & 1 &= 9 + (-1) \times 8, \\ 8 &= 1 \times 8.\end{aligned}$$

Hence a solution is obtained from

$$\begin{aligned}1 &= 9 + (-1) \times 8 \\ &= 9 + (-1) \times (26 + (-2) \times 9) = (-1) \times 26 + 3 \times 9 \\ &= (-1) \times 26 + 3 \times (35 + (-1) \times 26) = 3 \times 35 + (-4) \times 26 \\ &= 3 \times 35 + (-4) \times (61 + (-1) \times 35) = 7 \times 35 + (-4) \times 61.\end{aligned}$$

So  $x = 7$ ,  $y = -4$  is an integer solution. To find *all* integer solutions, notice that another solution  $x, y$  must satisfy  $35(x - 7) + 61(y + 4) = 0$ , hence we have  $35 \mid 61(y + 4)$  so by Euclid's Lemma 1.17,  $35 \mid (y + 4)$ . Thus  $y = -4 + 35k$  for some  $k \in \mathbb{Z}$  and then  $x = 7 - 61k$ . The general integer solution is

$$x = 7 - 61k, \quad y = -4 + 35k \quad (k \in \mathbb{Z}).$$

Here is the general result about this kind of problem.

**THEOREM 1.41.** *If  $a, b, c \in \mathbb{Z}$  and  $h = \gcd(a, b)$  set  $a = uh$  and  $b = vh$ .*

*a) If  $h \nmid c$ , then the equation  $ax + by = c$  has no integer solutions.*

b) If  $h \mid c$ , then the equation  $ax + by = c$  has integer solutions. If  $x_0, y_0$  is a particular integer solution, then the general integer solution is  $x = x_0 - vk$ ,  $y = y_0 + uk$  ( $k \in \mathbb{Z}$ ).

PROOF. a) This is obvious.

b) Dividing through by  $h$  gives the equivalent equation  $ux + vy = w$ , where  $c = wh$ . This can be solved as in the preceding discussion to obtain the stated general solution.  $\square$

We end this section by showing how continued fractions can be used to find one solution of the above Diophantine problem,  $35x + 61y = 1$ . Consider the continued fraction expansion

$$\frac{61}{35} = 1 + \frac{1}{1 + \frac{1}{2 + \frac{1}{1 + \frac{1}{8}}}} = [1; 1, 2, 1, 8].$$

The penultimate convergent is  $[1; 1, 2, 1] = 7/4$ . Apart from the signs involved, the numbers 7, 4 are those appearing in the above solution. This illustrates a general result which is closely related to the tabular method of §4.

**THEOREM 1.42.** *If  $a, b$  are coprime positive integers, then a solution of  $ax + by = 1$  is obtained from the continued fraction expansion*

$$\frac{b}{a} = [c_0; c_1, \dots, c_m]$$

by taking the penultimate convergent  $\frac{p_{m-1}}{q_{m-1}}$  and setting

$$x = (-1)^m p_{m-1}, \quad y = (-1)^{m-1} q_{m-1}.$$

PROOF. This is a consequence of Corollary 1.33(ii) where we take  $n = m$ . This gives

$$bq_{m-1} - ap_{m-1} = (-1)^{m-1}, \quad \text{i.e., } (-1)^m p_{m-1}a + (-1)^{m-1} q_{m-1}b = 1. \quad \square$$

## 11. Pell's equation

Another important Diophantine problem is the solution of *Pell's Equation*  $x^2 - dy^2 = 1$ , where  $d$  is an integer which is not a square. It turns out that the integer solutions  $x, y$  of this equation can be found using continued fractions. We will describe the method without detailed proofs.

From now on, let  $d$  be a non-square natural number. Let  $[a_0; \overline{a_1, \dots, a_p}]$  be the infinite continued fraction expansion of period  $p$  for  $\sqrt{d}$ , with  $n$ -th convergent  $A_n = p_n/q_n$  using the notation of Theorem 1.32.

**THEOREM 1.43.** *If  $x = u, y = v$  is a positive integer solution of the equation  $x^2 - dy^2 = 1$ , then  $u/v$  is a convergent of the continued fraction expansion of  $\sqrt{d}$ .*

**THEOREM 1.44.**

a) *If the period  $p$  is even then all positive integer solutions of the equation  $x^2 - dy^2 = 1$  are given by*

$$x = p_{kp-1}, \quad y = q_{kp-1} \quad (k = 1, 2, 3, \dots).$$

b) *If the period  $p$  is odd then all positive integer solutions of the equation  $x^2 - dy^2 = 1$  are given by*

$$x = p_{2kp-1}, \quad y = q_{2kp-1} \quad (k = 1, 2, 3, \dots).$$

**EXAMPLE 1.45.** Find all positive integer solutions of  $x^2 - 2y^2 = 1$ .

SOLUTION. From Example 1.38 we have  $\sqrt{2} = [1; \overline{2}]$  with  $p = 1$ . So the positive integer solutions are

$$x = p_{2k-1}, y = q_{2k-1} \quad (k = 1, 2, 3, \dots).$$

The first few are  $(x, y) = (3, 2), (17, 12), (99, 70)$ .  $\square$

EXAMPLE 1.46. Find all positive integer solutions of  $x^2 - 3y^2 = 1$ .

SOLUTION. From Example 1.39 we have  $\sqrt{3} = [1, \overline{1, 2}]$  with  $p = 2$ . So the positive integer solutions are

$$x = p_{2k-1}, y = q_{2k-1} \quad (k = 1, 2, 3, \dots).$$

The first few are  $(x, y) = (2, 1), (7, 4), (26, 15)$ .  $\square$

The *fundamental solution* of  $x^2 - dy^2 = 1$  is the positive integral solution  $x_1, y_1$  with  $x_1, y_1$  minimal. Thus since the convergents of  $\sqrt{d}$  have  $p_n, q_n$  strictly increasing, we have

$$(x_1, y_1) = \begin{cases} (p_{p-1}, q_{p-1}) & \text{if the period } p \text{ is even,} \\ (p_{2p-1}, q_{2p-1}) & \text{if the period } p \text{ is odd.} \end{cases}$$

THEOREM 1.47. *The positive integral solutions of  $x^2 - dy^2 = 1$  are precisely the pairs of integers  $(x_n, y_n)$  for which*

$$x_n + y_n\sqrt{d} = (x_1 + y_1\sqrt{d})^n.$$

For  $d = 2$ ,

$$\begin{aligned} x_1 + y_1\sqrt{2} &= 3 + 2\sqrt{2}, \\ x_2 + y_2\sqrt{2} &= 17 + 12\sqrt{2}, \\ x_3 + y_3\sqrt{2} &= 99 + 70\sqrt{2}, \\ x_4 + y_4\sqrt{2} &= 577 + 408\sqrt{2}. \end{aligned}$$

For  $d = 3$ ,

$$\begin{aligned} x_1 + y_1\sqrt{3} &= 2 + 1\sqrt{3}, \\ x_2 + y_2\sqrt{3} &= 7 + 4\sqrt{3}, \\ x_3 + y_3\sqrt{3} &= 26 + 15\sqrt{3}, \\ x_4 + y_4\sqrt{3} &= 97 + 56\sqrt{3}. \end{aligned}$$

EXAMPLE 1.48. Find the fundamental solution of  $x^2 - 97y^2 = 1$  and hence find 3 other positive integral solutions.

SOLUTION. We have  $\sqrt{97} = [9; \overline{1, 5, 1, 1, 1, 1, 1, 1, 5, 1, 18}]$  which has period  $p = 11$ . The fundamental solution is

$$(x_1, y_1) = (p_{21}, q_{21}) = (62809633, 6377352),$$

while the first few solutions are given by

$$\begin{aligned} x_1 + y_1\sqrt{97} &= 62809633 + 6377352\sqrt{97}, \\ x_2 + y_2\sqrt{97} &= 7890099995189377 + 801118277263632\sqrt{97}, \\ x_3 + y_3\sqrt{97} &= 991148570062293006927649 + 100635889969041933956760\sqrt{97}, \\ x_4 + y_4\sqrt{97} &= 124507355868174813917084187296257 \\ &\quad + 12641806631167809665750389674528\sqrt{97}. \end{aligned} \quad \square$$

### Problem Set 1

1-1. If  $a, b, c$  are non-zero integers, show that each of the following statements is true.

- (a) If  $a \mid b$  and  $b \mid c$ , then  $a \mid c$ .
- (b) If  $a \mid b$  and  $b \mid a$ , then  $b = \pm a$ .
- (c) If  $k \in \mathbb{Z}$  is non-zero, then  $\gcd(ka, kb) = |k| \gcd(a, b)$ .

1-2. Use the Euclidean Algorithm and the method of back-substitution to find the following greatest common divisors and in each case express  $\gcd(a, b)$  as an integer linear combination of  $a, b$ :

$$\gcd(76, 98), \gcd(108, 120), \gcd(1008, -520), \gcd(936, -876), \gcd(-591, 691).$$

Use the tabular method of §4 to check your results.

1-3. For non-zero integers  $a, b$ , show that the set

$$S = \{xa + yb : x, y \in \mathbb{Z}, 0 < xa + yb\}$$

defined in the proof of Theorem 1.5 agrees with the set

$$T = \{t \gcd(a, b); t \in \mathbb{N}_0, 0 < t\}.$$

1-4. Use the method in the proof of Theorem 1.5, show that if  $n \in \mathbb{Z}$ ,  $\gcd(12n + 5, 5n + 2) = 1$ .

1-5. Find all integer solutions  $x$  (if there are any) of each of the following congruences:

$$(a) 9x \equiv_{245} 23; (b) 21x \equiv_{77} 7; (c) 21x \equiv_{77} 8; (d) 210x \equiv_{1007} 97; (e) 13x \equiv_{37} 36.$$

1-6. Find all integer solutions  $x$  (if there are any) of each of the following pairs of simultaneous congruences:

$$(a) 9x \equiv_{245} 23, 210x \equiv_{1007} 97; (b) 21x \equiv_{77} 7, 13x \equiv_{37} 36; (c) 9x \equiv_{245} 23, 21x \equiv_{77} 7.$$

1-7. [Challenge question] Using Maple, for a collection of values  $n = 2, 3, 4, 5, 6, 7, 8, \dots$  determine all the solutions modulo  $n$  of the congruence  $x^3 - x \equiv_n 0$ . Can you spot anything systematic about the *number* of solutions modulo  $n$ ?

1-8. Show that if a prime  $p$  divides a product of integers  $a_1 \cdots a_n$ , then  $p \mid a_j$  for some  $j$ .

1-9. Let  $p, q$  be a pair of distinct prime numbers. Show that each of the following is irrational:

$$(a) \sqrt[n]{p} \text{ for } n > 1; (b) \sqrt{\frac{p}{q}}; (c) \frac{\sqrt[n]{p}}{\sqrt[s]{q}} \text{ for any coprime pair of natural numbers } r, s.$$

1-10. Let  $p_1, p_2, \dots, p_r$  and  $q_1, q_2, \dots, q_s$  be primes which satisfy the congruences

$$p_i \equiv_4 1 \quad (1 \leq i \leq r), \quad q_j \equiv_4 3 \quad (1 \leq j \leq s).$$

Show that  $p_1 p_2 \cdots p_r q_1 q_2 \cdots q_s \equiv_4 (-1)^s$ .

Use this result to show that for any natural number  $n$ ,  $4n + 3$  is divisible by at least one prime  $p$  with  $p \equiv_4 3$ .

1-11. (a) Find two roots of the polynomial  $f(x) = x^4 + 22$  modulo 23. Hence find three factors of  $f(x)$  modulo 23 and explain why you would not expect there to be any other monic linear factors.

(b) Find two roots of the polynomial  $g(x) = x^4 + 4x^2 + 43x^3 + 43x + 3$  modulo 47. Hence find three factors of  $g(x)$  modulo 47 and explain why you would not expect there to be any other monic linear factors.

1-12. For each of the primes  $p = 5, 7, 11, 13, 17, 19, 23, 37$  find a primitive root modulo  $p$ .

1-13. Let  $p$  be an *odd* prime. For  $t \in \mathbb{Z}$  with  $p \nmid t$ , define

$$\left(\frac{t}{p}\right) = \begin{cases} 1 & \text{if there is a } u \in \mathbb{Z} \text{ such that } t \equiv u^2, \\ -1 & \text{otherwise.} \end{cases}$$

Use the proof of Proposition 1.29 to show that

$$\left(\frac{-1}{p}\right) = (-1)^{(p-1)/2}.$$

1-14. Let  $p$  be an odd prime. Show that  $(1+p)^p \equiv 1 \pmod{p^2}$ .

More generally, show by Induction that for  $n \geq 2$ , the following congruences are true:

$$(1+p)^{p^{n-2}} \equiv 1 + p^{n-1} \pmod{p^n}, \quad (1+p)^{p^{n-1}} \equiv 1 \pmod{p^n}.$$

What can you say about the case  $p = 2$ ?

1-15. Determine the two continued fraction expansions of each of the numbers

$$1/3, 2/3, 3.14159, 3.14160, 51/11, 1725/1193, 1193/1725, -1193/1725, 30031/16579, 1103/87.$$

In each case determine all the convergents.

1-16. If  $n$  is a positive integer, what are the continued fraction expansions of  $-n$  and  $1/n$ ? What about when  $n$  is negative? [Hint: *Try a few examples first then attempt to formulate and prove general results.*]

Try to find a relationship between the continued fraction expansions of  $a/b$  and  $-a/b$ ,  $b/a$  when  $a, b$  are non-zero natural numbers.

1-17. If  $A = [a_0; a_1, \dots, a_n]$  with  $A > 1$ , show that  $1/A = [0; a_0, a_1, \dots, a_n]$ .

Let  $x > 1$  be a real number. Show that the  $n$ -th convergent of the continued fraction representation of  $x$  agrees with the  $(n-1)$ -th convergent of the continued fraction representation of  $1/x$ .

1-18. Find the continued fraction expansions of  $\frac{1}{\sqrt{5}}$  and  $\frac{1}{\sqrt{5}-1}$ . Determine as many convergents as you can.

1-19. Investigate the continued fraction expansions of  $\sqrt{6}$  and  $\frac{1}{\sqrt{6}}$ . Determine as many convergents as you can.

1-20. [*Challenge question*] Try to determine the first 10 terms in the continued fraction expansion of  $e$  using the series expansion

$$e = 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \dots$$

1-21. Find all the solutions of each of the following Diophantine equations:

$$(a) 64x + 108y = 4, \quad (b) 64x + 108y = 2, \quad (c) 64x + 108y = 12.$$

1-22. Let  $n$  be a positive integer.

a) Prove the identities

$$n + \sqrt{n^2 + 1} = 2n + (\sqrt{n^2 + 1} - n) = 2n + \frac{1}{n + \sqrt{n^2 + 1}}.$$

b) Show that  $[\sqrt{n^2 + 1}] = n$  and that the infinite continued fraction expansion of  $\sqrt{n^2 + 1}$  is  $[n; \overline{2n}]$ .

c) Show that  $[\sqrt{n^2 + 2}] = n$  and that the infinite continued fraction expansion of  $\sqrt{n^2 + 2}$  is  $[n; \overline{n, 2n}]$ .

d) Show that  $[\sqrt{n^2 + 2n}] = n$  and that the infinite continued fraction expansion of  $\sqrt{n^2 + 2n}$  is  $[n; \overline{1, 2n}]$ .

1-23. Find the fundamental solutions of Pell's equation  $x^2 - dy^2 = 1$  for each of the values  $d = 5, 6, 8, 11, 12, 13, 31, 83$ . In each case find as many other solutions as you can.



## CHAPTER 2

# Groups and group actions

### 1. Groups

Let  $G$  be set and  $*$  a *binary operation* which combines each pair of elements  $x, y \in G$  to give another element  $x * y \in G$ . Then  $(G, *)$  is a *group* if the following conditions are satisfied.

Gp1: for all elements  $x, y, z \in G$ ,  $(x * y) * z = x * (y * z)$ ;

Gp2: there is an element  $\iota \in G$  such that for every  $x \in G$ ,  $\iota * x = x = x * \iota$ ;

Gp3: for every  $x \in G$ , there is a unique element  $y \in G$  such that  $x * y = \iota = y * x$ .

Gp1 is usually called the associativity law.  $\iota$  is usually called the identity element of  $(G, *)$ . In Gp3, the unique element  $y$  associated to  $x$  is called the inverse of  $x$  and is denoted  $x^{-1}$ .

EXAMPLE 2.1. The following are examples of groups.

(1)  $G = \mathbb{Z}$ ,  $* = +$ ,  $\iota = 0$  and  $x^{-1} = -x$ .

(2)  $G = \mathbb{Q}$ ,  $* = +$ ,  $\iota = 0$  and  $x^{-1} = -x$ .

(3)  $G = \mathbb{R}$ ,  $* = +$ ,  $\iota = 0$  and  $x^{-1} = -x$ .

EXAMPLE 2.2. Let  $n > 0$  be a natural number. Then  $(\mathbb{Z}/n, +)$  is a group with

$$\begin{aligned}\iota &= 0_n, \\ x^{-1} &= -x_n = (-x)_n.\end{aligned}$$

EXAMPLE 2.3. Let  $R = \mathbb{Q}, \mathbb{R}, \mathbb{C}$ . Then each of these choices gives a group  $(\text{GL}_2(R), *)$  with

$$\begin{aligned}\text{GL}_2(R) &= \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} : a, b, c, d \in R, ad - bc \neq 0 \right\}, \\ * &= \text{multiplication of matrices,} \\ \iota &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2, \\ \begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} &= \begin{bmatrix} \frac{d}{ad - bc} & -\frac{b}{ad - bc} \\ -\frac{c}{ad - bc} & \frac{a}{ad - bc} \end{bmatrix}.\end{aligned}$$

EXAMPLE 2.4. Let  $X$  be a finite set and let  $\text{Perm}(X)$  be the set of all bijections  $f: X \rightarrow X$ . Then  $(\text{Perm}(X), \circ)$  is a group where

$$\begin{aligned}\circ &= \text{composition of functions,} \\ \iota &= \text{Id}_X = \text{the identity function on } X, \\ f^{-1} &= \text{the inverse function of } f.\end{aligned}$$

$(\text{Perm}(X), \circ)$  is called the *permutation group* of  $X$ . We will study these and other examples in more detail.

If a group  $(G, *)$  has a finite underlying set  $G$ , then the number of elements in the  $G$  is called the *order* of  $G$ , written  $|G|$ .

## 2. Permutation groups

We will follow the ideas of Example 2.4 and consider the *standard set with  $n$  elements*

$$\mathbf{n} = \{1, 2, \dots, n\}.$$

The  $S_n = \text{Perm}(\mathbf{n})$  is called the *symmetric group on  $n$  objects* or the *symmetric group of degree  $n$*  or the *permutation group on  $n$  objects*.

THEOREM 2.5.  $S_n$  has order  $|S_n| = n!$ .

PROOF. Defining an element  $\sigma \in S_n$  is equivalent to specifying the list

$$\sigma(1), \sigma(2), \dots, \sigma(n)$$

consisting of the  $n$  numbers  $1, 2, \dots, n$  taken in some order with no repetitions. To do this we have

- $n$  choices for  $\sigma(1)$ ,
- $n - 1$  choices for  $\sigma(2)$  (taken from the remaining  $n - 1$  elements),
- and so on.

In all, this gives  $n \times (n - 1) \times \dots \times 2 \times 1 = n!$  choices for  $\sigma$ , so  $|S_n| = n!$  as claimed. We will often describe  $\sigma$  using the notation

$$\sigma = \begin{pmatrix} 1 & 2 & \dots & n \\ \sigma(1) & \sigma(2) & \dots & \sigma(n) \end{pmatrix}. \quad \square$$

EXAMPLE 2.6. The elements of  $S_3$  are the following,

$$\iota = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}.$$

We can calculate the composition  $\tau \circ \sigma$  of two permutations  $\tau, \sigma \in S_n$ , where  $\tau \sigma(k) = \tau(\sigma(k))$ . Notice that we apply  $\sigma$  to  $k$  first then apply  $\tau$  to the result  $\sigma(k)$ . For example,

$$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} = \iota.$$

In particular,

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}^{-1}.$$

Let  $X$  be a set with exactly  $n$  elements which we list in some order,  $x_1, x_2, \dots, x_n$ . Then there is an *action* of  $S_n$  on  $X$  given by

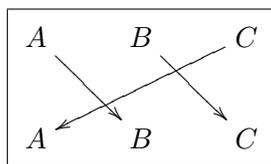
$$\sigma \cdot x_k = x_{\sigma(k)} \quad (\sigma \in S_n, k = 1, 2, \dots, n).$$

For example, if  $X = \{A, B, C\}$  we can take  $x_1 = A, x_2 = B, x_3 = C$  and so

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \cdot A = B, \quad \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \cdot B = C, \quad \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \cdot C = A.$$

Often it is useful to display the effect of a permutation  $\sigma: X \rightarrow X$  by indicating where each element is sent by  $\sigma$  with the aid of arrows. To do this we display the elements of  $X$  in two similar rows with an arrow joining  $x_i$  in the first row to  $\sigma(x_i)$  in the second. For example,

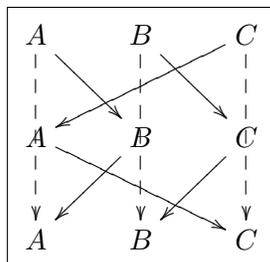
the permutation  $\sigma = \begin{pmatrix} A & B & C \\ B & C & A \end{pmatrix}$  acting on  $X = \{A, B, C\}$  can be displayed as



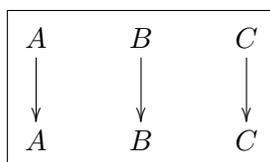
We can compose permutations by composing the arrows. Thus

$$\begin{pmatrix} A & B & C \\ C & A & B \end{pmatrix} \begin{pmatrix} A & B & C \\ B & C & A \end{pmatrix}$$

can be determined from the diagram



which gives the identity function whose diagram is



### 3. The sign of a permutation

Let  $\sigma \in S_n$  and consider the arrow diagram of  $\sigma$  as above. Let  $c_\sigma$  be the number of crossings of arrows. The *sign* of  $\sigma$  is the number

$$\operatorname{sgn} \sigma = (-1)^{c_\sigma} = \begin{cases} +1 & \text{if } c_\sigma \text{ is even,} \\ -1 & \text{if } c_\sigma \text{ is odd.} \end{cases}$$

Then  $\operatorname{sgn}: S_n \rightarrow \{+1, -1\}$ . Notice that  $\{+1, -1\}$  is actually a group under multiplication.

PROPOSITION 2.7. *The function  $\operatorname{sgn}: S_n \rightarrow \{+1, -1\}$  satisfies*

$$\operatorname{sgn}(\tau\sigma) = \operatorname{sgn}(\tau) \operatorname{sgn}(\sigma) \quad (\tau, \sigma \in S_n).$$

PROOF. By considering the arrow diagram for  $\tau\sigma$  obtained by joining the diagrams for  $\sigma$  and  $\tau$ , we see that the total number of crossings is  $c_\sigma + c_\tau$ . If we straighten out the paths starting at each number in the top row, so that we change the total number of crossings by 2 each time. So  $(-1)^{c_\sigma + c_\tau} = (-1)^{c_{\tau\sigma}}$ .  $\square$

A permutation  $\sigma$  is called *even* if  $\operatorname{sgn} \sigma = 1$ , otherwise it is *odd*. The set of all even permutations in  $S_n$  is denoted by  $A_n$ . Notice that  $\iota \in A_n$  and in fact the following result is true.

PROPOSITION 2.8. *The set  $A_n$  forms a group under composition.*

PROOF. By Proposition 2.7, if  $\sigma, \tau \in A_n$ , then

$$\operatorname{sgn}(\tau\sigma) = \operatorname{sgn}(\tau) \operatorname{sgn}(\sigma) = 1.$$

Note also that  $\iota \in A_n$ .

The arrow diagram for  $\sigma^{-1}$  is obtained from that for  $\sigma$  by interchanging the rows and reversing all the arrows, so  $\operatorname{sgn} \sigma^{-1} = \operatorname{sgn} \sigma$ . Thus if  $\sigma \in A_n$ , then  $\operatorname{sgn} \sigma^{-1} = 1$ .

Hence,  $A_n$  is a group under composition.  $\square$

$A_n$  is called the *n-th alternating group*.

EXAMPLE 2.9. The elements of  $A_3$  are

$$\iota = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}.$$

We will see later that  $|A_n| = |S_n|/2 = n!/2$ .

#### 4. The cycle type of a permutation

Suppose  $\sigma \in S_n$ . Now carry out the following steps.

- Form the sequence

$$1 \rightarrow \sigma(1) \rightarrow \sigma^2(1) \rightarrow \cdots \rightarrow \sigma^{r_1-1}(1) \rightarrow \sigma^{r_1}(1) = 1$$

where  $\sigma^k(j) = \sigma(\sigma^{k-1}(j))$  and  $r_1$  is the smallest positive power for which this is true.

- Take the smallest number  $k_2 = 1, 2, \dots, n$  for which  $k_2 \neq \sigma^t(1)$  for every  $t$ . Form the sequence

$$k_2 \rightarrow \sigma(k_2) \rightarrow \sigma^2(k_2) \rightarrow \cdots \rightarrow \sigma^{r_2-1}(k_2) \rightarrow \sigma^{r_2}(k_2) = k_2$$

where  $r_2$  is the smallest positive power for which this is true.

- Repeat this with  $k_3 = 1, 2, \dots, n$  being the smallest number for which  $k_3 \neq \sigma^t(k_2)$  for every  $t$ .
- $\vdots$

Writing  $k_1 = 1$ , we end up with a collection of *disjoint cycles*

$$\begin{aligned} k_1 &\rightarrow \sigma(k_1) \rightarrow \sigma^2(k_1) \rightarrow \cdots \rightarrow \sigma^{r_1-1}(k_1) \rightarrow \sigma^{r_1}(k_1) = k_1 \\ k_2 &\rightarrow \sigma(k_2) \rightarrow \sigma^2(k_2) \rightarrow \cdots \rightarrow \sigma^{r_2-1}(k_2) \rightarrow \sigma^{r_2}(k_2) = k_2 \\ &\vdots \\ k_d &\rightarrow \sigma(k_d) \rightarrow \sigma^2(k_d) \rightarrow \cdots \rightarrow \sigma^{r_d-1}(k_d) \rightarrow \sigma^{r_d}(k_d) = k_d \end{aligned}$$

in which every number  $k = 1, 2, \dots, n$  occurs in exactly one row.

The  $s$ -th one of these cycles can be viewed as corresponding to the permutation of  $\mathbf{n}$  which behaves according to the action of  $\sigma$  on the elements that appear as  $\sigma^t(k_s)$  and fix every other element. We indicate this permutation using the *cycle notation*

$$(k_s \ \sigma(k_s) \ \cdots \ \sigma^{r_s-1}(k_s)).$$

Then we have

$$\sigma = (k_1 \ \sigma(k_1) \ \cdots \ \sigma^{r_1-1}(k_1)) \cdots (k_d \ \sigma(k_d) \ \cdots \ \sigma^{r_d-1}(k_d)),$$

which is the *disjoint cycle decomposition* of  $\sigma$ . It is unique apart from the order of the factors and the order in which the numbers within each cycle occur.

For example, in  $S_4$ ,

$$\begin{aligned} (1\ 2)(3\ 4) &= (2\ 1)(4\ 3) = (3\ 4)(1\ 2) = (4\ 3)(2\ 1), \\ (1\ 2\ 3)(1) &= (3\ 1\ 2)(1) = (2\ 3\ 1)(1) = (1)(1\ 2\ 3) = (1)(3\ 1\ 2) = (1)(2\ 3\ 1). \end{aligned}$$

We usually leave out cycles of length 1, so for example  $(1\ 2\ 3)(1) = (1\ 2\ 3)$ .

Recall that when performing elementary row operations (ERO's) on  $n \times n$  matrices, one of the types involves interchanging a pair of rows, say rows  $r$  and  $s$ , this operation is denoted by  $R_r \leftrightarrow R_s$ . The corresponding elementary matrix  $E(R_r \leftrightarrow R_s)$  is obtained from the identity matrix  $I_n$  by performing this operation. In fact, we can do a sequence of such operations to obtain any *permutation matrix*  $P_\sigma = [p_{ij}]$ , whose rows are obtained by applying the permutation  $\sigma \in S_n$  to those of  $I_n$  so that

$$p_{ij} = \delta_{\sigma(i)j} = \begin{cases} 1 & \text{if } j = \sigma(i), \\ 0 & \text{if } j \neq \sigma(i). \end{cases}$$

For example, if  $n = 3$  and  $\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$ , then

$$P_\sigma = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}.$$

PROPOSITION 2.10. For  $\sigma \in S_n$ ,  $\det P_\sigma = \operatorname{sgn} \sigma$ .

A permutation  $\tau \in S_n$  which interchanges two elements of  $\mathbf{n}$  and leaves the rest fixed is called a *transposition*.

PROPOSITION 2.11. Let  $\sigma \in S_n$ . Then there are transpositions  $\tau_1, \dots, \tau_k$  such that  $\sigma = \tau_1 \cdots \tau_k$ .

One way to decompose a permutation  $\sigma$  into transpositions is to first decompose it into disjoint cycles then use the easily checked formula

$$(2.1) \quad (i_1 i_2 \dots i_r) = (i_1 i_r) \cdots (i_1 i_3)(i_1 i_2).$$

EXAMPLE 2.12. Decompose

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 5 & 3 & 1 & 4 \end{pmatrix} \in S_5$$

into a product of transpositions.

SOLUTION. We have

$$\sigma = (3)(1\ 2\ 5\ 4) = (1\ 2\ 5\ 4) = (1\ 4)(1\ 5)(1\ 2).$$

Some alternative decompositions are

$$\sigma = (2\ 1)(2\ 4)(2\ 5) = (5\ 2)(5\ 1)(5\ 4). \quad \square$$

## 5. Symmetry groups

Let  $S$  be a set of points in  $\mathbb{R}^n$ , where  $n = 1, 2, 3, \dots$ . A *symmetry* of  $S$  is a surjection  $\varphi: S \rightarrow S$  which preserves distances, *i.e.*,

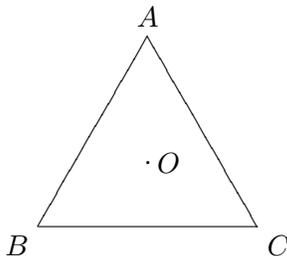
$$|\varphi(\mathbf{u}) - \varphi(\mathbf{v})| = |\mathbf{u} - \mathbf{v}| \quad (\mathbf{u}, \mathbf{v} \in S).$$

THEOREM 2.13. Let  $\varphi$  be a symmetry of  $S \subseteq \mathbb{R}^n$ . Then

- $\varphi$  is a bijection and  $\varphi^{-1}$  is also a symmetry of  $S$ ;
- $\varphi$  preserves distances between points and angles between lines joining points.

COROLLARY 2.14. Let  $S \subseteq \mathbb{R}^n$ . Then the set  $\operatorname{Sym}(S)$  of all symmetries of  $S$  is a group under composition.

EXAMPLE 2.15. Let  $T \subseteq \mathbb{R}^2$  be an equilateral triangle  $\triangle$  with vertices  $A, B, C$ .

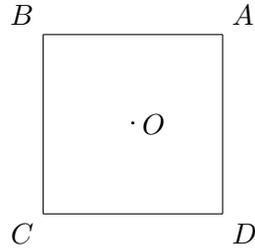


Then a symmetry is defined once we know where the vertices go, hence there are as many symmetries as permutations of the set  $\{A, B, C\}$ . Each symmetry can be described using permutation notation and we obtain the 6 symmetries

$$\iota = \begin{pmatrix} A & B & C \\ A & B & C \end{pmatrix}, \begin{pmatrix} A & B & C \\ B & C & A \end{pmatrix}, \begin{pmatrix} A & B & C \\ C & A & B \end{pmatrix}, \begin{pmatrix} A & B & C \\ A & C & B \end{pmatrix}, \begin{pmatrix} A & B & C \\ C & B & A \end{pmatrix}, \begin{pmatrix} A & B & C \\ B & A & C \end{pmatrix}.$$

Therefore we have  $|\text{Sym}(\Delta)| = 6$ .

EXAMPLE 2.16. Let  $S \subseteq \mathbb{R}^2$  be the square  $\square$  centred at the origin  $O$  with vertices at  $A(1, 1)$ ,  $B(-1, 1)$ ,  $C(-1, -1)$ ,  $D(1, -1)$ .

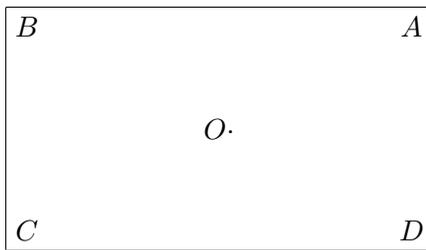


Then a symmetry is defined by sending  $A$  to any one of the 4 vertices then choosing how to send  $B$  to one of the 2 adjacent vertices. This gives a total of  $4 \times 2 = 8$  such symmetries, thus  $|\text{Sym}(\square)| = 8$ .

Again we can describe symmetries in terms of their effect on the vertices. Here are the 8 elements of  $\text{Sym}(\square)$  described in permutation notation.

$$\iota = \begin{pmatrix} A & B & C & D \\ A & B & C & D \end{pmatrix}, \begin{pmatrix} A & B & C & D \\ B & C & D & A \end{pmatrix}, \begin{pmatrix} A & B & C & D \\ C & D & A & B \end{pmatrix}, \begin{pmatrix} A & B & C & D \\ D & A & B & C \end{pmatrix}, \\ \begin{pmatrix} A & B & C & D \\ A & D & C & B \end{pmatrix}, \begin{pmatrix} A & B & C & D \\ D & C & B & A \end{pmatrix}, \begin{pmatrix} A & B & C & D \\ C & B & A & D \end{pmatrix}, \begin{pmatrix} A & B & C & D \\ B & A & D & C \end{pmatrix}.$$

EXAMPLE 2.17. Let  $R \subseteq \mathbb{R}^2$  be the rectangle centred at the origin  $O$  with vertices at  $A(2, 1)$ ,  $B(-2, 1)$ ,  $C(-2, -1)$ ,  $D(2, -1)$ .



A symmetry can send  $A$  to any of the vertices, and then the long edge  $AB$  must go to the longer of the adjacent edges. This gives a total of 4 such symmetries, thus  $|\text{Sym}(R)| = 4$ .

Again we can describe symmetries in terms of their effect on the vertices. Here are the 4 elements of  $\text{Sym}(R)$  described in permutation notation.

$$\iota = \begin{pmatrix} A & B & C & D \\ A & B & C & D \end{pmatrix}, \begin{pmatrix} A & B & C & D \\ B & A & D & C \end{pmatrix}, \begin{pmatrix} A & B & C & D \\ C & D & A & B \end{pmatrix}, \begin{pmatrix} A & B & C & D \\ D & C & B & A \end{pmatrix}$$

Given a regular  $n$ -gon (*i.e.*, a regular polygon with  $n$  sides all of the same length and  $n$  vertices  $V_1, V_2, \dots, V_n$ ) the symmetry group is the *dihedral group of order  $2n$*   $D_{2n}$ , with elements

$$\iota, \alpha, \alpha^2, \dots, \alpha^{n-1}, \tau, \alpha\tau, \alpha^2\tau, \dots, \alpha^{n-1}\tau$$

where  $\alpha^k$  is an anticlockwise rotation through  $2\pi k/n$  about the centre and  $\tau$  is a reflection in the line through  $V_1$  and the centre. Moreover we have

$$|\alpha| = n, |\tau| = 2, \tau\alpha\tau = \alpha^{n-1} = \alpha^{-1}.$$

In permutation notation this becomes

$$\alpha = (V_1 V_2 \cdots V_n),$$

but  $\tau$  is more complicated to describe.

For example, if  $n = 6$  we have

$$\alpha = (V_1 V_2 V_3 V_4 V_5 V_6), \quad \tau = (V_2 V_6)(V_3 V_5),$$

while if  $n = 7$

$$\alpha = (V_1 V_2 V_3 V_4 V_5 V_6 V_7), \quad \tau = (V_2 V_7)(V_3 V_6)(V_4 V_5).$$

We have seen that when  $n = 3$ ,  $\text{Sym}(\Delta)$  is the permutation group of the vertices and so  $D_6$  is essentially the same group as  $S_6$ .

## 6. Subgroups and Lagrange's Theorem

Let  $(G, *)$  be a group and  $H \subseteq G$ . Then  $H$  is a *subgroup* of  $G$  if  $(H, *)$  is a group. In detail this means that

- for  $x, y \in H$ ,  $x * y \in H$ ;
- $\iota \in H$ ;
- if  $z \in H$  then  $z^{-1} \in H$ .

We write  $H \leq G$  whenever  $H$  is a subgroup of  $G$  and  $H < G$  if  $H \neq G$ , *i.e.*,  $H$  is a *proper subgroup* of  $G$ .

EXAMPLE 2.18. For  $n \in \mathbb{Z}^+$ ,  $A_n$  is a subgroup of  $S_n$ , *i.e.*,  $A_n \leq S_n$ .

By Example 2.3, for each choice of  $R = \mathbb{Q}, \mathbb{R}, \mathbb{C}$ , there is a group  $(\text{GL}_2(R), *)$  with

$$\text{GL}_2(R) = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} : a, b, c, d \in R, ad - bc \neq 0 \right\}.$$

EXAMPLE 2.19. Let

$$\text{SL}_2(R) = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} : a, b, c, d \in R, ad - bc = 1 \right\} \subseteq \text{GL}_2(R).$$

Then  $\text{SL}_2(R)$  is a subgroup of  $\text{GL}_2(R)$ , *i.e.*,  $\text{SL}_2(R) \leq \text{GL}_2(R)$ .

SOLUTION. This follows easily with aid of the three identities

$$\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc; \quad \det(AB) = \det A \det B; \quad \begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}. \quad \square$$

Let  $(G, *)$  be a group. From now on, if  $x, y \in G$  we will write  $xy$  for  $x * y$ . Also, for  $n \in \mathbb{Z}$  we write

$$x^n = \begin{cases} x(x^{n-1}) & \text{if } n > 0, \\ \iota & \text{if } n = 0, \\ (x^{-1})^{-n} & \text{if } n < 0. \end{cases}$$

If  $g \in G$ ,

$$\langle g \rangle = \{g^n : n \in \mathbb{Z}\} \subseteq G$$

is a subgroup of  $G$  called the *subgroup generated by  $g$* . This follows from the three equations

$$g^m g^n = g^{m+n}; \quad \iota = g^0; \quad (g^n)^{-1} = g^{-n}.$$

If  $\langle g \rangle$  is finite and contains exactly  $n$  elements then  $g$  is said to have *finite order*  $|g| = n$ . If  $\langle g \rangle$  is infinite then  $g$  is said to have *infinite order*  $|g| = \infty$ .

PROPOSITION 2.20. *If  $g \in G$  has finite order  $|g|$  then*

$$|g| = \min\{m \in \mathbb{Z}^+ : m > 0, g^m = \iota\}.$$

EXAMPLE 2.21. In the group  $S_n$  the cyclic permutation  $(i_1 i_2 \cdots i_r)$  of length  $r$  has order

$$|(i_1 i_2 \cdots i_r)| = r.$$

SOLUTION. Setting  $\sigma = (i_1 i_2 \cdots i_r)$ , we have

$$\sigma^k(1) = \begin{cases} i_{k+1} & \text{if } k < r, \\ i_1 & \text{if } k = r, \end{cases}$$

hence  $|\sigma| \leq r$ . As  $i_k \neq 1$  for  $1 < k \leq r$ ,  $r$  is the smallest such power which is  $\iota$ , hence  $|\sigma| = r$ .  $\square$

So for example,  $|(1 2)| = 2$ ,  $|(1 2 3)| = 3$  and  $|(1 2 3 4)| = 4$ . But notice that the product  $(1 2)(3 4 5)$  satisfies

$$((1 2)(3 4 5))^2 = (1 2)(3 4 5)(1 2)(3 4 5) = (3 5 4),$$

hence  $|(1 2)(3 4 5)| = 6$ . On the other hand, the product  $(1 2)(2 3 4)$  satisfies

$$((1 2)(2 3 4))^2 = (1 2)(2 3 4)(1 2)(2 3 4) = (1 3)(2 4)$$

so  $|(1 2)(3 4 5)| = |(1 3)(2 4)| = 2$ .

A group  $(G, *)$  is called *cyclic* if there is an element  $c \in G$  such that  $G = \langle c \rangle$ ; such a  $c$  is called a *generator* of  $G$ . Notice that for such a group,  $|G| = |c|$ .

EXAMPLE 2.22. The group  $(\mathbb{Z}, +)$  is cyclic of infinite order with generators  $\pm 1$ .

EXAMPLE 2.23. If  $0 < n \in \mathbb{N}_0$ , then the group  $(\mathbb{Z}/n, +)$  is cyclic of finite order  $n$ . Two generators are  $\pm 1_n \in \mathbb{Z}/n$ . More generally,  $t_n$  is a generator if and only if  $\gcd(t, n) = 1$ .

SOLUTION. We have that for each  $k \in \mathbb{Z}$ ,  $k = \pm(1 + 1 + \cdots + 1)$  (with  $\pm k$  summands). From this we see that  $\pm 1_n$  are obvious generators and so  $\mathbb{Z}/n = \langle \pm 1_n \rangle$ .

If  $\gcd(t, n) = 1$ , then by Theorem 1.9, there is an integer  $u$  such that  $ut \equiv 1 \pmod{n}$ . Hence  $1_n \in \langle t_n \rangle$  and so  $\mathbb{Z}/n = \langle t_n \rangle$ .

Conversely, if  $\mathbb{Z}/n = \langle t_n \rangle$  then for some  $k \in \mathbb{N}_0$  we have  $1 \equiv 1 + \cdots + 1 \pmod{n}$  (with  $k$  summands) and so  $kt \equiv 1 \pmod{n}$ , hence  $kt + \ell n = 1$  for some  $\ell \in \mathbb{Z}$ . But this implies  $\gcd(t, n) \mid 1$ , hence  $\gcd(t, n) = 1$ .  $\square$

The *Euler  $\varphi$ -function*  $\varphi: \mathbb{Z}^+ \rightarrow \mathbb{N}_0$  is defined by

$$\begin{aligned} \varphi(n) &= \text{number of generators of } \mathbb{Z}/n \\ &= \text{number of elements } t_n \in \mathbb{Z}/n \text{ with } \gcd(t, n) = 1. \end{aligned}$$

In order to state some properties of  $\varphi$ , we need to introduce some notation. For a positive natural number  $n$  and a function  $f$  defined on the positive natural numbers, the symbol  $\sum_{d|n} f(d)$

denotes the sum of all the numbers  $f(d)$  where  $d$  ranges over all the positive integer divisors of  $n$ , including 1 and  $n$ . For example,

$$\sum_{d|6} f(d) = f(1) + f(2) + f(3) + f(6).$$

THEOREM 2.24. *The Euler function  $\varphi$  enjoys the following properties:*

- $\varphi(1) = 1$ ;
- if  $\gcd(m, n) = 1$  then  $\varphi(mn) = \varphi(m)\varphi(n)$ ;
- if  $p$  is a prime and  $r \geq 1$  then  $\varphi(p^r) = (p-1)p^{r-1}$ .
- for a non-zero natural number  $n$ ,  $\sum_{d|n} \varphi(d) = n$ .

For example,

$$\varphi(120) = \varphi(8 \cdot 3 \cdot 5) = \varphi(8)\varphi(3)\varphi(5) = \varphi(2^3)\varphi(3)\varphi(5) = 2^2 \cdot 2 \cdot 4 = 2^5 = 32.$$

The next result is actually a consequence of *Lagrange's Theorem* which follows immediately after it and is of great importance in the study of finite groups.

**PROPOSITION 2.25.** *Let  $G$  be a finite group and let  $g \in G$ . Then  $g$  has finite order and  $|g|$  divides  $|G|$ .*

**THEOREM 2.26** (Lagrange's Theorem). *Let  $(G, *)$  be a finite group and  $H \leq G$ . Then  $|H|$  divides  $|G|$ .*

**PROOF.** The idea is to divide up  $G$  into disjoint subsets of size  $H$ . We do this by defining for each  $x \in G$  the *left coset of  $x$  with respect to  $H$* ,

$$xH = \{g \in G : x^{-1}g \in H\} = \{g \in G : g = xh \text{ for some } h \in H\}.$$

We need the following facts.

i) For  $x, y \in G$ ,  $xH \cap yH \neq \emptyset \iff xH = yH$ .

This is seen as follows. If  $xH = yH$  then  $xH \cap yH \neq \emptyset$ . Conversely, suppose that  $xH \cap yH \neq \emptyset$ . If  $yh \in xH$  for some  $h \in H$ , then  $x^{-1}yh \in H$ . For  $k \in H$ ,

$$x^{-1}yk = (x^{-1}yh)(h^{-1}k),$$

which is in  $H$  since  $x^{-1}yh, h^{-1}k \in H$  and  $H$  is a subgroup of  $G$ . Hence  $yH \subseteq xH$ . Repeating this argument with  $x$  and  $y$  interchanged we also see that  $xH \subseteq yH$ . Combining these inclusions we obtain  $xH = yH$ .

ii) For each  $g \in G$ ,  $|gH| = |H|$ .

If  $gh = gk$  for  $h, k \in H$  then  $g^{-1}(gh) = g^{-1}(gk)$  and so  $h = k$ . Thus there is a bijection

$$\theta: H \longrightarrow gH; \quad \theta(h) = gh,$$

which implies that the sets  $H$  and  $gH$  have the same number of elements.

Thus every element  $g \in G$  lies in exactly one such coset  $gH$ . Thus  $G$  is the union of these disjoint cosets which all have size  $H$ . Denoting the number of these cosets by  $[G : H]$  we have  $|G| = |H|[G : H]$ .  $\square$

The number  $[G : H]$  of cosets of  $H$  in  $G$  is called the *index of  $H$  in  $G$* . The set of all cosets of  $H$  in  $G$  is denoted  $G/H$ , *i.e.*,

$$G/H = \{gH : g \in G\}.$$

**COROLLARY 2.27.** *If  $G$  is a finite group and  $H \leq G$ , then  $|G| = |H||G/H| = |H|[G : H]$ .*

Proposition 2.25 now follows easily by taking  $H = \langle g \rangle$  and using the fact that  $|g| = |H|$ .

This allows us to give a promised proof of a number theoretic result, the Primitive Element Theorem 1.27. Indeed the following generalisation is true.

**THEOREM 2.28.** *Let  $G$  be a group of finite order  $n = |G|$  and suppose that for each divisor  $d$  of  $n$  there are at most  $d$  elements of  $G$  satisfying  $x^d = \iota$ . Then  $G$  is cyclic and so abelian.*

**PROOF.** Let  $\theta(d)$  denote the number of elements in  $G$  of order  $d$ . By Proposition 2.25,  $\theta(d) = 0$  unless  $d$  divides  $|G|$ . Since

$$G = \bigcup_{d||G|} \{g \in G : |g| = d\},$$

we have

$$|G| = \sum_{d||G|} \theta(d).$$

By Theorem 2.24(d), we also have

$$|G| = \sum_{d||G|} \varphi(d).$$

Combining these we obtain

$$(2.2) \quad \sum_{d||G|} \varphi(d) = \sum_{d||G|} \theta(d).$$

We will show that for each divisor  $d$  of  $|G|$ ,  $\theta(d) \leq \varphi(d)$ . For each such  $d$  of  $|G|$ , we have  $\theta(d) \geq 0$ . If  $\theta(d) = 0$  then  $\theta(d) < \varphi(d)$ , since the latter is positive. So suppose that  $\theta(d) > 0$ , hence there is an element  $a \in G$  of order  $d$ . In fact, the distinct powers  $\iota = a^0, a, a^2, \dots, a^{d-1}$  are all solutions of the equation  $x^d = \iota$  and indeed, by assumption on  $G$ , they must be the only such solutions since there are  $d$  of them. But now an element  $a^k \in \langle a \rangle$  with  $k = 0, 1, 2, \dots, d-1$  has order  $d$  precisely if  $\gcd(d, k) = 1$  since this requires  $\langle a^k \rangle = \langle a \rangle$  and so for some  $u \in \mathbb{Z}$ ,  $uk \equiv 1 \pmod{d}$  which happens precisely when  $\gcd(d, k) = 1$  as we know from Theorem 1.9. By the definition of  $\varphi$ , there are  $\varphi(d)$  of such elements in  $\langle a \rangle$ , hence  $\theta(d) = \varphi(d)$ . Thus we have shown that in all cases  $\theta(d) \leq \varphi(d)$ .

Notice that if  $\theta(d) < \varphi(d)$  for some  $d$  dividing  $|G|$ , this would give a *strict* inequality in place of Equation (2.2). Hence we must always have  $\theta(d) = \varphi(d)$ . In particular, there are  $\varphi(n)$  elements of order  $n$ , hence there must be an element of order  $n$ , so  $G$  is cyclic.  $\square$

Taking  $G = U_p$ , the group of invertible elements of  $\mathbb{Z}/p$  under multiplication, we obtain Theorem 1.27.

## 7. Group actions

If  $X$  is a set and  $(G, *)$  then a (*group*) *action* of  $(G, *)$  on  $X$  is a rule which assigns to each  $g \in G$  and  $x \in X$  an element  $gx \in X$  so that the following conditions are satisfied.

**GpAct1** For all  $g_1, g_2 \in G$  and  $x \in X$ ,  $(g_1 * g_2)x = g_1(g_2x)$ .

**GpAct2** For  $x \in X$ ,  $\iota x = x$ .

Thus each  $g \in G$  can be viewed as acting as a permutation of  $X$ .

**EXAMPLE 2.29.** Let  $G \leq S_n$  and let  $X = \mathbf{n}$ . For  $\sigma \in G$  and  $k \in \mathbf{n}$  let  $\sigma k = \sigma(k)$ . This defines an action of  $(G, \circ)$  on  $\mathbf{n}$ .

**EXAMPLE 2.30.** Let  $X \subseteq \mathbb{R}^n$  and let  $G \leq \text{Sym}(X)$  be a subgroup of the symmetry group of  $X$ . For  $\varphi \in G$  and  $x \in X$ , let  $\varphi x = \varphi(x)$ . This defines an action of  $(G, \circ)$  on  $X$ .

Suppose we have an action of a group  $(G, *)$  on a set  $X$ . For  $x \in X$ , the *stabilizer* of  $x$  is

$$\text{Stab}_G(x) = \{g \in G : gx = x\} \subseteq G,$$

and the *orbit* of  $x$  is

$$\text{Orb}_G(x) = \{gx : g \in G\} \subseteq X.$$

Notice that  $x = \iota x$ , so  $x \in \text{Orb}_G(x)$  and  $\iota \in \text{Stab}_G(x)$ . Thus  $\text{Stab}_G(x) \neq \emptyset$  and  $\text{Orb}_G(x) \neq \emptyset$ .

**THEOREM 2.31.** For each  $x, y \in X$ ,

- a)  $\text{Stab}_G(x) \leq G$ ;
- b)  $y \in \text{Orb}_G(x)$  if and only if  $x \in \text{Orb}_G(y)$ ;
- c)  $y \in \text{Orb}_G(x)$  if and only if  $\text{Orb}_G(y) = \text{Orb}_G(x)$ .

**PROOF.**

a) If  $g_1, g_2 \in \text{Stab}_G(x)$  then by GpAct1,

$$(g_1 * g_2)x = g_1(g_2x) = g_1x = x.$$

By GpAct2,  $\iota x = x$ , hence  $\iota \in \text{Stab}_G(x)$ . Finally, if  $g \in \text{Stab}_G(x)$  then by GpAct1 and GpAct2,

$$g^{-1}x = g^{-1}(gx) = (g^{-1} * g)x = \iota x = x,$$

hence  $g^{-1} \in \text{Stab}_G(x)$ . So  $\text{Stab}_G(x) \leq G$ .

b) If  $y \in \text{Orb}_G(x)$ , then  $y = gx$  for some  $g \in G$ . Hence  $x = (g^{-1} * g)x = g^{-1}(gx) = g^{-1}y$  and so  $x \in \text{Orb}_G(y)$ . The converse is similar.

c) If  $y \in \text{Orb}_G(x)$  then by (b),  $x \in \text{Orb}_G(y)$  and so  $x = ky$  for some  $k \in G$ . Hence if  $g \in G$ ,  $gx = g(ky) = (g * k)y \in \text{Orb}_G(y)$  and so  $\text{Orb}_G(x) \subseteq \text{Orb}_G(y)$ . By (b),  $x \in \text{Orb}_G(y)$  and so we also have  $\text{Orb}_G(y) \subseteq \text{Orb}_G(x)$ . This gives  $\text{Orb}_G(y) = \text{Orb}_G(x)$ .

Conversely, if  $\text{Orb}_G(y) = \text{Orb}_G(x)$  then  $y \in \text{Orb}_G(y) = \text{Orb}_G(x)$ .  $\square$

EXAMPLE 2.32. Let  $X = \square$  be the square with vertices  $A, B, C, D$  and let  $G = \text{Sym}(\square)$ . Determine  $\text{Stab}_G(x)$  and  $\text{Orb}_G(x)$  where

- a)  $x$  is the vertex  $A$ ;
- b)  $x$  is the midpoint  $M$  of  $AB$ ;
- c)  $x$  is the point  $P$  on  $AB$  where  $AP : PB = 1 : 3$ .

SOLUTION. Recall Example 2.16. We will write permutations of the vertices in cycle notation.

a) We have

$$\text{Stab}_G(A) = \{\iota, (B D)\}.$$

Also, every vertex can be obtained from  $A$  by applying a suitable symmetry, hence

$$\text{Orb}_G(A) = \{A, B, C, D\}.$$

b) A symmetry  $\varphi$  fixes the midpoint of  $AB$  if and only if it maps this edge to itself. The symmetries doing this have one of the effects  $\varphi(A) = A, \varphi(B) = B$  or  $\varphi(A) = B, \varphi(B) = A$ . Thus

$$\text{Stab}_G(M) = \{\iota, (A B)(C D)\}.$$

Also, we can arrange to send  $A$  to any other vertex and  $B$  to either of the adjacent vertices of the image of  $A$ , hence the orbit of  $M$  consists of the set of 4 midpoints of edges.

c) A symmetry  $\varphi$  can only fix  $P$  if it sends  $A$  to a vertex  $A'$  say, and  $B$  to a vertex  $B'$  with  $A'P : PB' = 1 : 3$  and this is only possible if  $A' = A$  and  $B' = B$ , hence  $\varphi$  must also fix  $A, B$ . So  $\text{Stab}_G(P) = \{\iota\}$ . On the other hand, since we can select a symmetry to send  $A$  to any other vertex and  $B$  to either of the adjacent vertices to the image,  $P$  can be sent to any of the points  $Q$  which cut an edge in the ratio  $1 : 3$ . So the orbit of  $P$  is the set consisting of these 8 points.  $\square$

THEOREM 2.33 (Orbit-Stabilizer Theorem). *Let  $(G, *)$  act on  $X$ . Then for  $x \in X$  there is a bijection  $F: G/\text{Stab}_G(x) \rightarrow \text{Orb}_G(x)$  between the set of cosets of  $\text{Stab}_G(x)$  in  $G$  and the orbit of  $x$ , defined by  $F(g\text{Stab}_G(x)) = gx$ . Moreover we have*

$$F((t * g)\text{Stab}_G(x)) = tF(g\text{Stab}_G(x)) \quad (t \in G).$$

PROOF. We begin by checking that  $F$  is well defined. If  $g_1\text{Stab}_G(x) = g_2\text{Stab}_G(x)$ , then  $g_1^{-1}g_2 \in \text{Stab}_G(x)$  and

$$g_1x = g_1((g_1^{-1}g_2)x) = (g_1g_1^{-1}g_2)x = g_2x.$$

Hence  $F$  is well defined.

Notice that  $gx = kx$  if and only if  $(g^{-1}k)x = x$ , i.e.,  $g^{-1}k \in \text{Stab}_G(x)$  which means that

$$g\text{Stab}_G(x) = k\text{Stab}_G(x).$$

So  $F$  is an injection. Also, every  $y \in \text{Orb}_G(x)$  has the form  $tx = F(t\text{Stab}_G(x))$  for some  $t \in G$ , which shows that  $F$  is surjective.

The final equation property is a consequence of the definition of  $F$ . □

COROLLARY 2.34. *If  $G$  is finite then for each  $x \in X$ ,*

$$|\text{Orb}_G(x)| = \frac{|G|}{|\text{Stab}_G(x)|}.$$

PROOF. This follows from Corollary 2.27. □

The sizes of the orbits in Example 2.32 can be found using this result.

THEOREM 2.35. *The orbits of an action of  $(G, *)$  on  $X$  decompose  $X$  into a union of disjoint subsets,*

$$X = \bigcup_{U \text{ an orbit}} U.$$

COROLLARY 2.36. *If  $X$  is finite then*

$$|X| = \sum_{U \text{ an orbit}} |U|.$$

In these results, each orbit  $U$  has the form  $\text{Orb}_G(x_U)$  for some element  $x_U \in X$ . Moreover, if  $G$  is finite, then

$$|U| = [G : \text{Stab}_G(x_U)] = \frac{|G|}{|\text{Stab}_G(x_U)|}.$$

The formula in Corollary 2.36 becomes the *orbit-stabilizer equation*:

$$(2.3) \quad |X| = \sum_{U \text{ an orbit}} \frac{|G|}{|\text{Stab}_G(x_U)|}.$$

If there is only one orbit, then the action is said to be *transitive*, and in this case, for any  $x \in X$  we have  $X = \text{Orb}_G(x)$  and  $|X| = |G|/|\text{Stab}_G(x)|$ .

Given an action of  $(G, *)$  on  $X$ , another useful idea is that of the *fixed point set* or *fixed set* of an element  $g \in G$ ,

$$\text{Fix}_G(g) = \{x \in X : gx = x\}.$$

$\text{Fix}_G(g)$  is also often denoted  $X^g$ .

THEOREM 2.37 (Burnside Formula). *If  $(G, *)$  acts on  $X$  with  $G$  and  $X$  finite, then*

$$\text{number of orbits} = \frac{1}{|G|} \sum_{g \in G} |\text{Fix}_G(g)|.$$

PROOF. The right hand side of the formula is

$$\begin{aligned}
 \frac{1}{|G|} \sum_{g \in G} |\text{Fix}_G(g)| &= \frac{1}{|G|} \sum_{g \in G} \sum_{x \in \text{Fix}_G(g)} 1 \\
 &= \frac{1}{|G|} \sum_{x \in X} \sum_{g \in \text{Stab}_G(x)} 1 \\
 &= \frac{1}{|G|} \sum_{x \in X} |\text{Stab}_G(x)| \\
 &= \frac{1}{|G|} \sum_{\substack{U = \text{Orb}_G(x) \\ \text{an orbit}}} |U| \cdot |\text{Stab}_G(x)| \quad (\text{by Corollary 2.34}) \\
 &= \frac{1}{|G|} \sum_{U \text{ an orbit}} |G| \\
 &= \sum_{U \text{ an orbit}} 1 \\
 &= \text{number of orbits.} \quad \square
 \end{aligned}$$

EXAMPLE 2.38. Let  $X = \{1, 2, 3, 4\}$  and let  $G \leq S_4$  be the subgroup

$$G = \{\iota, (12), (34), (12)(34)\}$$

acting on  $X$  in the obvious way. How many orbits does this action have?

SOLUTION. Here  $|G| = 4 = |X|$ . Furthermore we have

$$\text{Fix}_G(\iota) = X, \quad \text{Fix}_G((12)) = \{3, 4\}, \quad \text{Fix}_G((34)) = \{1, 2\}, \quad \text{Fix}_G((12)(34)) = \emptyset.$$

The Burnside Formula gives

$$\text{number of orbits} = \frac{1}{4} (4 + 2 + 2 + 0) = \frac{8}{4} = 2.$$

So there are 2 orbits, namely  $\{1, 2\}$  and  $\{3, 4\}$ .  $\square$

EXAMPLE 2.39. Let  $X = \{1, 2, 3, 4, 5, 6\}$  and let  $G = \langle (123)(45) \rangle \leq S_6$  be the cyclic subgroup acting on  $X$  in the obvious way. How many orbits does this action have?

SOLUTION. Here  $|G| = 6$  and  $|X| = 6$ . The elements of  $G$  are

$$\iota, (123)(45), (132), (45), (123), (132)(45).$$

The fixed sets of these are

$$\begin{aligned}
 \text{Fix}_G(\iota) &= X, & \text{Fix}_G((123)(45)) &= \text{Fix}_G((132)(45)) = \{6\}, \\
 \text{Fix}_G((45)) &= \{1, 2, 3, 6\}, & \text{Fix}_G((123)) &= \text{Fix}_G((132)) = \{4, 5, 6\}.
 \end{aligned}$$

By the Burnside Formula,

$$\text{number of orbits} = \frac{1}{6} (6 + 1 + 3 + 4 + 3 + 1) = \frac{18}{6} = 3.$$

So there are 3 orbits, namely  $\{1, 2, 3\}$ ,  $\{4, 5\}$  and  $\{6\}$ .  $\square$

EXAMPLE 2.40. A dinner party of seven people is to sit around a circular table with seven seats. How many distinguishable ways are there to do this if there is to be no ‘head of table’?

SOLUTION. View the seven places as numbered 1 to 7. There are  $7!$  ways to arrange the diners in these places. Take  $X$  to be the set of all possible such arrangements, so  $|X| = 7!$ . Regard two such arrangements as indistinguishable if one is obtained from the other by a rotation of the diners around the places. Clearly there are 7 such rotations, each involving everyone moving  $k$  seats to the right for some  $k = 0, 1, \dots, 6$ . Let  $\alpha$  denote the rotation corresponding to everyone moving one seat to the right. Then to get everyone to move  $k$  seats we repeatedly apply  $\alpha$   $k$  times in all, *i.e.*,  $\alpha^k$ . This suggests we should consider the group

$$G = \{\iota, \alpha, \alpha^2, \alpha^3, \alpha^4, \alpha^5, \alpha^6\}$$

consisting of all of these operations, with composition as the binary operation. This provides an action of  $G$  on  $X$ .

The number of indistinguishable seating plans is the number of orbits under this action, *i.e.*,

$$\frac{1}{|G|} \sum_{g \in G} |\text{Fix}_G(g)|.$$

Notice that apart from the identity element, no rotation can fix any arrangement, so when  $g \neq \iota$ ,  $\text{Fix}_G(g) = \emptyset$ , while  $\text{Fix}_G(\iota) = X$ . Hence the number of indistinguishable seating plans is  $7!/7 = 6! = 720$ .  $\square$

EXAMPLE 2.41. Find the number of distinguishable ways there are to colour the edges of an equilateral triangle using four different colours, where each colour can be used on more than one edge.

SOLUTION. Let  $X$  be the set of all possible such colourings of the equilateral triangle  $ABC$  whose symmetry group is  $G = S_3$ , which we view as the permutation group of  $\{A, B, C\}$ ; hence  $|G| = 6$ . Also  $|X| = 4^3 = 64$  since each edge can be coloured in 4 ways.  $G$  acts on  $X$  in the obvious way. A pair of colourings is indistinguishable precisely if they are in the same orbit.

By the Burnside formula, the number of distinguishable colourings is given by

$$\text{number of orbits} = \frac{1}{6} \sum_{\sigma \in G} |\text{Fix}_G(\sigma)|.$$

The fixed sets of elements of the various cycle types in  $G$  are as follows.

Identity element  $\iota$ :  $\text{Fix}_G(\iota) = X$ ,  $|\text{Fix}_G(\iota)| = 64$ .

3-cycles (*i.e.*,  $\sigma = (ABC), (ACB)$ ): these give rotations and can only fix a colouring that has all sides the same colour, hence  $|\text{Fix}_G(\sigma)| = 4$ .

2-cycles (*i.e.*,  $\sigma = (AB), (AC), (BC)$ ): each of these gives a reflection in a line through a vertex and the midpoint of the opposite edge. For example,  $(AB)$  fixes  $C$  and interchanges the edges  $AC, BC$ , it will therefore fix any colouring that has these edges the same colour. There are  $4 \times 4 = 16$  of these, so  $|\text{Fix}_G((AB))| = 16$ . Similarly for the other 2-cycles.

By the Burnside formula,

$$\text{number of distinguishable colourings} = \frac{1}{6} (64 + 2 \times 4 + 3 \times 16) = \frac{120}{6} = 20. \quad \square$$

### Problem Set 2

2-1. Which of the following pairs  $(G, *)$  forms a group?

- (a)  $G = \{x \in \mathbb{Z} : x \neq 0\}$ ,  $* = \times$ ;
- (b)  $G = \{x \in \mathbb{Q} : x \neq 0\}$ ,  $* = \times$ ;
- (c)  $G = \left\{ \begin{bmatrix} a & b \\ -b & a \end{bmatrix} : a, b \in \mathbb{R}, a^2 + b^2 = 1 \right\}$ ,  $* = \text{multiplication of matrices}$ ;
- (d)  $G = \left\{ \begin{bmatrix} z & w \\ -\bar{w} & \bar{z} \end{bmatrix} : z, w \in \mathbb{C}, |z|^2 + |w|^2 = 1 \right\}$ ,  $* = \text{multiplication of matrices}$ ;
- (e)  $G = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} : a, b, c, d \in \mathbb{Z}, ad - bc \neq 0 \right\}$ ,  $* = \text{multiplication of matrices}$ ;
- (f)  $G = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} : a, b, c, d \in \mathbb{Z}, ad - bc = 1 \right\}$ ,  $* = \text{multiplication of matrices}$ ;
- (g)  $G = \{\varphi \in S_n : \varphi(n) = n\}$ ,  $* = \text{composition of functions}$ .

2-2. For each of the following permutations in  $S_6$ , determine its sign and decompose it into disjoint cycles:

$$\alpha = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 4 & 2 & 5 & 6 & 1 \end{pmatrix}, \quad \beta = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 6 & 4 & 1 & 5 & 2 \end{pmatrix}, \quad \gamma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 3 & 4 & 1 & 6 & 2 & 5 \end{pmatrix}.$$

2-3. Find the orders of the symmetry groups of the following geometric objects, and in each case try to describe the symmetry groups as groups of permutations:

- a regular pentagon;
- a regular hexagon;
- a regular hexagon with vertices alternately coloured red and green;
- a regular hexagon with edges alternately coloured red and green;
- a cube;
- a cube with the pairs of opposite faces coloured red, green and blue respectively.

2-4. [*Challenge question.*] Suppose Tet is a regular tetrahedron with vertices  $A, B, C, D$ .

- Show that the symmetry group  $\text{Sym}(\text{Tet})$  of Tet can be identified with the symmetric group  $S_4$  which acts by permuting the vertices.
- For each pair of distinct vertices  $P, Q$ , how many symmetries map the edge  $PQ$  into itself? Show that these symmetries form a group.
- Find a geometric interpretation of the alternating group  $A_4$  acting as symmetries of Tet.

2-5. In each of the following groups  $(G, *)$  decide whether the subset  $H$  is a subgroup of  $G$  and when it is, decide whether it is cyclic.

- $G = \{x \in \mathbb{Q} : x \neq 0\}$ ,  $H = \{x \in G : x > 0\}$ ,  $* = \times$ ;
- $G = \{x \in \mathbb{Q} : x \neq 0\}$ ,  $H = \{x \in G : x < 0\}$ ,  $* = \times$ ;
- $G = \{x \in \mathbb{Q} : x \neq 0\}$ ,  $H = \{x \in G : x^2 = 1\}$ ,  $* = \times$ ;
- $G = \{x \in \mathbb{C} : x \neq 0\}$ ,  $H = \{x \in G : x^d = 1\}$ ,  $* = \times$ ;
- $G = \{z \in \mathbb{C} : z \neq 0\}$ ,  $H = \{z \in G : |z| < \infty\}$ ,  $* = \times$ ;
- $G = \left\{ \begin{bmatrix} z & w \\ -\bar{w} & \bar{z} \end{bmatrix} : z, w \in \mathbb{C}, |z|^2 + |w|^2 = 1 \right\}$ ,  $H = \{A \in G : |A| < \infty\}$ ,  
 $* = \text{matrix multiplication}$ ;
- $G = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} : a, b, c, d \in \mathbb{R}, ad - bc \neq 0 \right\}$ ,  $H = \left\{ A \in G : A = \begin{bmatrix} a & b \\ 0 & d \end{bmatrix} \right\}$ ,

\* = matrix multiplication;

h)  $G = \text{Sym}(\square)$ ,  $H$  = the subset of rotations in  $G$ , \* = composition of functions.

2-6. Using Lagrange's Theorem, find all possible orders of elements of each of the following groups and decide whether there are indeed elements of those orders:

$$\mathbb{Z}/6, S_3, A_3, S_4, A_4, D_8, D_{10}.$$

2-7. [*Challenge question*] Let  $G$  be a group. Show that each of the following subsets of  $G$  is a subgroup:

- $C_G(x) = \{c \in G : cx = xc\}$ , where  $x \in G$  is any element;
- $Z(G) = \{c \in G : cg = gc \text{ for all } g \in G\}$ ;
- $N_G(H) = \{n \in G : \text{for every } h \in H, nhn^{-1} \in H, \text{ and } n^{-1}hn \in H\}$ , where  $H \leq G$  is any subgroup.

2-8. Using Lagrange's Theorem, find all subgroups of each of the groups

$$\mathbb{Z}/6, S_3, A_3, S_4, A_4, D_8, D_{10}.$$

2-9. Let  $G = S_4$  and let  $X$  denote the set consisting of all subsets of  $\mathbf{4} = \{1, 2, 3, 4\}$ . For  $\sigma \in S_4$  and  $U \in X$ , let

$$\sigma U = \{\sigma(u) \in X : u \in U\}.$$

- Show that this defines an action of  $G$  on  $X$ .
- For each of the following elements  $U$  of  $X$ , find  $\text{Orb}_G(U)$  and  $\text{Stab}_G(U)$ :

$$\emptyset, \{1\}, \{1, 2\}, \{1, 2, 3\}, \{1, 2, 3, 4\}.$$

- For each of the following elements of  $G$  find  $\text{Fix}_G(g)$ :

$$\iota, (12), (123), (1234), (12)(34).$$

2-10. Let  $G = \text{GL}_2(\mathbb{R})$  be the group of  $2 \times 2$  invertible real matrices under matrix multiplication and let  $X = \mathbb{R}^2$  be the set of all real column vectors of length 2. For  $A \in G$  and  $\mathbf{x} \in X$  let  $A\mathbf{x}$  be the usual product.

- Show that this defines an action of  $G$  on  $X$ .
- Find the orbit and stabilizer of each the following vectors:

$$\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

- For each of the following matrices  $A$  find  $\text{Fix}_G(A)$ :

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 5 \end{bmatrix}, \begin{bmatrix} 2 & 0 \\ 0 & -3 \end{bmatrix}, \begin{bmatrix} \sin \theta & \cos \theta \\ \cos \theta & -\sin \theta \end{bmatrix}, \begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} u & 0 \\ 0 & u \end{bmatrix},$$

where  $\theta, u \in \mathbb{R}$  with  $u \neq 0$ .

2-11. [*Challenge question*] Using the same group  $G = \text{GL}_2(\mathbb{R})$  and notation as in the previous question, let  $Y$  denote the set of all lines through the origin in  $\mathbb{R}^2$ . For  $A \in G$  and  $L \in Y$ , let

$$AL = \{A\mathbf{x} \in \mathbb{R}^2 : \mathbf{x} \in L\}.$$

- Show that  $AL$  is always a line and that this defines an action of  $G$  on  $Y$ .
- For each of the following vectors  $\mathbf{v}$  find the line  $L_{\mathbf{v}}$  through the origin containing it and find the orbit and stabilizer of  $L_{\mathbf{v}}$ :

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

- For each of the matrices  $A$  in (c) of the previous question, find  $\text{Fix}_G(A)$  for this action.

2-12. Let  $G = \text{Sym}(\text{Tet})$  be the symmetry group of the regular tetrahedron Tet with vertices  $A, B, C, D$ . Let  $X$  denote the set of edges of Tet. For  $\varphi \in G$  and  $E \in X$  let

$$\varphi E = \{\varphi(P) \in \text{Tet} : P \in E\}.$$

- a) Show that  $\varphi E$  is an edge and that this defines an action of  $G$  on  $X$ .
- b) Find  $\text{Orb}_G(E)$  and  $\text{Stab}_G(E)$  for the edge  $AB$ .
- c) For each of the following elements of  $G$  find  $\text{Fix}_G(g)$ :

$$\iota, (AB), (ABC), (ABCD), (AB)(CD).$$

2-13. Let  $X = \{1, 2, 3, 4, 5, 6, 7, 8\}$  and  $G = \langle (123456)(78) \rangle$  be the cyclic subgroup of  $S_7$  acting on  $X$  in the obvious way. How many orbits does this action of  $G$  have?

2-14. How many distinguishable 5-bead circular necklaces can be made where each bead has to be a different colour chosen from 5 colours? Here two such necklaces are deemed to be indistinguishable if one can be obtained from the other by a combination of rotations and flips. What if the number of colours used is 6? 7? 8?

What if we only allow rotations between indistinguishable necklaces?

2-15. How many distinguishable regular tetrahedral dice can be made where each face has one of the numbers 1,2,3,4 on it? Here two such dice are deemed to be indistinguishable if one can be obtained from the other by a rotation.

What about if we allow arbitrary symmetries between indistinguishable such dice?



## CHAPTER 3

# Arithmetic functions

### 1. Definition and examples of arithmetic functions

Let  $\mathbb{Z}^+ = \mathbb{N}_0 - \{0\}$  be the set of positive integers. A function  $\psi: \mathbb{Z}^+ \rightarrow \mathbb{R}$  (or  $\psi: \mathbb{Z}^+ \rightarrow \mathbb{C}$ ) is called a real (or complex) *arithmetic function* if  $\psi(1) = 1$ . There are many important and interesting examples.

EXAMPLE 3.1. The following are all real arithmetic functions:

a) The ‘identity’ function

$$\text{id}: \mathbb{Z}^+ \rightarrow \mathbb{R}; \quad \text{id}(n) = n.$$

b) The Euler function  $\varphi: \mathbb{Z}^+ \rightarrow \mathbb{R}$  of Theorem 2.24.

c) For each positive natural number  $r$ ,

$$\sigma_r: \mathbb{Z}^+ \rightarrow \mathbb{R}; \quad \sigma_r(n) = \sum_{d|n} d^r.$$

$\sigma_1$  is often denoted  $\sigma$ ;  $\sigma(n)$  is equal to the sum of the (positive) divisors of  $n$ .

d) The function given by

$$\delta: \mathbb{Z}^+ \rightarrow \mathbb{R}; \quad \delta(n) = \begin{cases} 1 & \text{if } n = 1, \\ 0 & \text{otherwise.} \end{cases}$$

e) The function given by

$$\eta: \mathbb{Z}^+ \rightarrow \mathbb{R}; \quad \eta(n) = 1.$$

The set of all real (or complex) arithmetic functions will be denoted by  $\text{AF}_{\mathbb{R}}$  (or  $\text{AF}_{\mathbb{C}}$ ).

An arithmetic function  $\psi$  is called (*strictly*) *multiplicative* if

$$\psi(mn) = \psi(m)\psi(n) \quad \text{whenever } \gcd(m, n) = 1.$$

By Theorem 2.24(b), the Euler function is strictly multiplicative. In fact, each of the functions in Example 3.1 is strictly multiplicative.

An important example is the *Möbius function*  $\mu: \mathbb{Z}^+ \rightarrow \mathbb{R}$  defined as follows. If  $n \in \mathbb{Z}^+$  then by the Fundamental Theorem of Arithmetic and Corollary 1.19, we have the prime power factorization  $n = p_1^{r_1} p_2^{r_2} \cdots p_t^{r_t}$ , where for each  $j$ ,  $p_j$  is a prime,  $1 \leq r_j$  and  $2 \leq p_1 < p_2 < \cdots < p_t$ . We set

$$\mu(n) = \mu(p_1^{r_1} p_2^{r_2} \cdots p_t^{r_t}) = \begin{cases} 0 & \text{if any } r_j > 1, \\ (-1)^t & \text{if all } r_j = 1. \end{cases}$$

So for example, if  $n = p$  is a prime,  $\mu(p) = -1$ , while  $\mu(p^2) = 0$ . Also,  $\mu(60) = \mu(2^2 \times 3 \times 5) = 0$ .

PROPOSITION 3.2. *The Möbius function  $\mu$  is multiplicative.*

PROOF. This follows from the definition and the fact that the prime power factorizations of two coprime natural numbers  $m, n$  have no common prime factors.  $\square$

So for example,

$$\mu(105) = \mu(3)\mu(5)\mu(7) = (-1)^3 = -1.$$

PROPOSITION 3.3. *The Möbius function  $\mu$  satisfies*

$$\begin{aligned}\mu(1) &= 1, \\ \sum_{d|n} \mu(d) &= 0 \quad \text{if } n \geq 2.\end{aligned}$$

PROOF. By Induction on  $r$ , the number of prime factors in the prime power factorization of  $n = p_1^{r_1} \cdots p_t^{r_t}$ , so  $r = r_1 + \cdots + r_t$ .

If  $r = 1$ , then  $n = p$  is prime and  $\mu(p) = -1$ , hence

$$\sum_{d|p} \mu(d) = 1 - 1 = 0.$$

Assume that whenever  $r < k$ . Then if  $r = k$ , let  $n = mp_t^{r_t}$  where  $p_t$  is a prime factor of  $n$ . Then  $\mu(n) = \mu(m)\mu(p_t^{r_t})$  and so

$$\sum_{d|n} \mu(d) = \sum_{d|m} (\mu(d) + \mu(dp_t)) = \sum_{d|m} (\mu(d) + \mu(d)\mu(p_t)) = \sum_{d|m} \mu(d)(1 - 1) = 0.$$

This gives the Inductive Step. □

## 2. Convolution and Möbius Inversion

Let  $\theta, \psi: \mathbb{Z}^+ \rightarrow \mathbb{R}$  (or  $\mathbb{C}$ ) be arithmetic functions. The *convolution* of  $\theta$  and  $\psi$  is the function  $\theta * \psi$  for which

$$\theta * \psi(n) = \sum_{d|n} \theta(d)\psi(n/d).$$

PROPOSITION 3.4. *The convolution of two arithmetic functions is an arithmetic function. Moreover,  $*$  satisfies*

a) for arithmetic functions  $\alpha, \beta, \gamma$ ,

$$(\alpha * \beta) * \gamma = \alpha * (\beta * \gamma);$$

b) for an arithmetic function  $\theta$ ,

$$\delta * \theta = \theta = \theta * \delta;$$

c) for an arithmetic function  $\theta$ , there is a unique arithmetic function  $\tilde{\theta}$  for which

$$\theta * \tilde{\theta} = \delta = \tilde{\theta} * \theta;$$

d) For two arithmetic functions  $\theta, \psi$ ,

$$\theta * \psi = \psi * \theta.$$

Hence  $(\text{AF}_{\mathbb{R}}, *)$  and  $(\text{AF}_{\mathbb{C}}, *)$  are commutative groups.

PROOF.

(a) For  $n \in \mathbb{Z}^+$ ,

$$\begin{aligned}(\alpha * \beta) * \gamma(n) &= \sum_{d|n} \alpha * \beta(d)\gamma(n/d) \\ &= \sum_{k|d} \sum_{d|n} \alpha(k)\beta(d/k)\gamma(n/d) \\ &= \sum_{k\ell m=n} \alpha(k)\beta(\ell)\gamma(m),\end{aligned}$$

and similarly

$$\alpha * (\beta * \gamma)(n) = \sum_{k\ell m=n} \alpha(k)\beta(\ell)\gamma(m).$$

Hence  $(\alpha * \beta) * \gamma(n) = \alpha * (\beta * \gamma)(n)$  for all  $n$ , so  $(\alpha * \beta) * \gamma = \alpha * (\beta * \gamma)$ .

(b) We have

$$\delta * \theta(n) = \sum_{d|n} \delta(d)\theta(n/d) = \theta(n),$$

and similarly  $\theta * \delta(n) = \theta(n)$ .

(c) Take  $t_1 = 1$ . We will show by Induction that there are numbers  $t_n$  for which

$$\sum_{d|n} t_d \theta(n/d) = \delta(n).$$

Suppose that for some  $k > 1$  we have such numbers  $t_n$  for  $n < k$ . Consider the equation

$$\sum_{d|k} t_d \theta(k/d) = \delta(k) = 0.$$

Rewriting this as

$$t_k = - \sum_{\substack{d|k \\ d \neq k}} t_d \theta(k/d),$$

we see that  $t_k$  is uniquely determined from this equation. Now define  $\tilde{\theta}$  by  $\tilde{\theta}(n) = t_n$ . By construction,

$$\theta * \tilde{\theta}(n) = \sum_{d|n} \theta(n/d)\tilde{\theta}(d) = \delta(n).$$

By (d) we also have  $\tilde{\theta} * \theta = \theta * \tilde{\theta}$ .

(d) We have

$$\theta * \psi(n) = \sum_{d|n} \theta(d)\psi(n/d) = \sum_{d|n} \psi(n/d)\theta(d) = \sum_{k|n} \psi(k)\theta(n/k) = \psi * \theta(n). \quad \square$$

In each of the groups  $(\text{AF}_{\mathbb{R}}, *)$  and  $(\text{AF}_{\mathbb{C}}, *)$ , the inverse of an arithmetic function  $\theta$  is  $\tilde{\theta}$ . Here is an important example.

**PROPOSITION 3.5.** *The inverse of  $\eta$  is  $\tilde{\eta} = \mu$ , the Möbius function.*

**PROOF.** Recall that  $\eta(n) = 1$  for all  $n$ . By Proposition 3.3 we have

$$\sum_{d|n} \mu(d)\eta(n/d) = \sum_{d|n} \mu(d) = \begin{cases} 1 & n = 1, \\ 0 & n > 1. \end{cases}$$

Hence  $\mu = \tilde{\eta}$  is the inverse of  $\eta$  by the proof of Proposition 3.4(c).  $\square$

**THEOREM 3.6 (Möbius Inversion).** *Let  $f, g: \mathbb{Z}^+ \rightarrow \mathbb{R}$  (or  $f, g: \mathbb{Z}^+ \rightarrow \mathbb{C}$ ) be arithmetic functions satisfying*

$$f(n) = \sum_{d|n} g(d) \quad (n \in \mathbb{Z}^+).$$

*Then*

$$g(n) = \sum_{d|n} f(d)\mu(n/d) \quad (n \in \mathbb{Z}^+).$$

PROOF. Notice that  $f = g * \eta$  from which we have

$$g = g * \delta = g * (\eta * \mu) = (g * \eta) * \mu = f * \mu.$$

Hence for  $n \in \mathbb{Z}^+$ ,

$$g(n) = \sum_{d|n} f(d)\mu(n/d). \quad \square$$

EXAMPLE 3.7. Use Möbius Inversion to find a formula for  $\varphi(n)$ , where  $\varphi$  is the Euler function.

SOLUTION. By Theorem 2.24(d),

$$\sum_{d|n} \varphi(d) = n.$$

This can be rewritten as the equation  $\varphi * \eta = \text{id}$  where  $\text{id}(n) = n$ . Applying Möbius Inversion gives  $\varphi = \text{id} * \mu$ , i.e.,

$$\varphi(n) = \sum_{d|n} \mu(d) \frac{n}{d} = \sum_{d|n} \mu(n/d)d. \quad \square$$

So for example, if  $n = p^r$  where  $p$  is a prime and  $r \geq 1$ ,

$$\varphi(p^r) = \sum_{d|p^r} \mu(d) \frac{p^r}{d} = \sum_{0 \leq s \leq r} \mu(p^s) p^{r-s} = \mu(1)p^r + \mu(p)p^{r-1} = p^r - p^{r-1} = (p-1)p^{r-1}.$$

EXAMPLE 3.8. Show that the function  $\sigma = \sigma_1$  satisfies

$$\sum_{d|n} \mu(d)\sigma(n/d) = n \quad (n \in \mathbb{Z}^+).$$

SOLUTION. By definition,

$$\sigma(n) = \sum_{d|n} d,$$

hence  $\sigma = \text{id} * \eta$ . By Möbius Inversion,

$$\text{id} = \text{id} * \delta = \text{id} * (\eta * \mu) = (\text{id} * \eta) * \mu = \sigma * \mu = \mu * \sigma,$$

so for  $n \in \mathbb{Z}^+$ ,

$$n = \sum_{d|n} \mu(d)\sigma(n/d) = \sum_{d|n} \sigma(d)\mu(n/d). \quad \square$$

PROPOSITION 3.9. If  $\theta, \psi$  are multiplicative arithmetic functions, then  $\theta * \psi$  is multiplicative.

PROOF. If  $m, n$  be coprime positive integers,

$$\begin{aligned} \theta * \psi(mn) &= \sum_{d|mn} \theta(d)\psi(mn/d) \\ &= \sum_{\substack{r|m \\ s|n}} \theta(rs)\psi(mn/rs) \\ &= \sum_{\substack{r|m \\ s|n}} \theta(r)\theta(s)\psi((m/r)(n/s)) \\ &= \sum_{\substack{r|m \\ s|n}} \theta(r)\theta(s)\psi(m/r)\psi(n/s) \\ &= \sum_{r|m} \theta(r)\psi(m/r) \sum_{s|n} \theta(s)\psi(n/s) \\ &= \theta * \psi(m)\theta * \psi(n). \end{aligned}$$

Hence  $\theta * \psi$  is multiplicative.  $\square$

COROLLARY 3.10. *Suppose that  $\theta$  is a multiplicative arithmetic function, and  $\psi$  is the arithmetic function satisfying*

$$\theta(n) = \sum_{d|n} \psi(d) \quad (n \in \mathbb{Z}^+).$$

*Then  $\psi$  is multiplicative.*

PROOF.  $\theta = \psi * \eta$ , so by Möbius Inversion,  $\psi = \theta * \mu$ , implying that  $\psi$  is multiplicative.  $\square$

**Problem Set 3**

3-1. Let  $\tau: \mathbb{Z}^+ \rightarrow \mathbb{R}$  be the function for which  $\tau(n)$  is the number of positive divisors of  $n$ .

a) Show that  $\tau$  is an arithmetic function.

b) Suppose that  $n = p_1^{r_1} p_2^{r_2} \cdots p_t^{r_t}$  is the prime power factorization of  $n$ , where  $2 \leq p_1 < p_2 < \cdots < p_t$  and  $r_j > 0$ . Show that

$$\tau(p_1^{r_1} p_2^{r_2} \cdots p_t^{r_t}) = (r_1 + 1)(r_2 + 1) \cdots (r_t + 1).$$

c) Is  $\tau$  multiplicative?

d) Show that  $\eta * \eta = \tau$ .

3-2. Show that each of the functions  $\sigma_r$  ( $r \geq 1$ ) of Example 3.1 are multiplicative.

3-3. For each  $r \in \mathbb{N}_0$  define the arithmetic function  $[r]: \mathbb{Z}^+ \rightarrow \mathbb{R}$  by

$$[r](n) = n^r.$$

In particular,  $[0] = \eta$  and  $[1] = \text{id}$ .

a) Show that  $[r]$  is multiplicative.

b) If  $r > 0$ , show that  $\sigma_r = [r] * \eta$ . Deduce that  $\sigma_r$  is multiplicative.

c) Show that  $[r] * [r]$  satisfies  $[r] * [r](n) = n^r \tau(n)$ .

d) Find a general formula for  $[r] * [s](n)$  when  $s < r$ .

3-4. For  $n \in \mathbb{Z}^+$ , prove the following formulæ, where the functions are defined in the text or in earlier questions.

$$(a) \sum_{d|n} \mu(d) \sigma(n/d) = n; \quad (b) \sum_{d|n} \mu(d) \tau(n/d) = 1; \quad (c) \sum_{d|n} \sigma_r(d) \mu(n/d) = n^r.$$

## Finite and infinite sets, cardinality and countability

The natural numbers originally arose from counting elements in sets. There are two very different possible ‘sizes’ for sets, namely *finite* and *infinite*, and in this section we discuss these concepts in detail.

### 1. Finite sets and cardinality

For a positive natural number  $n \geq 1$ , set

$$\mathbf{n} = \{1, 2, 3, \dots, n\}.$$

If  $n = 0$ , let  $\mathbf{0} = \emptyset$ . Then the set  $\mathbf{n}$  has  $n$  elements and we can think of it as the standard set of that size.

DEFINITION 4.1. Let  $f: X \rightarrow Y$  be a function.

- $f$  is an *injection* or *one-one (1-1)* if for  $x_1, x_2 \in X$ ,

$$f(x_1) = f(x_2) \implies x_1 = x_2.$$

- $f$  is a *surjection* or *onto* if for each  $y \in Y$ , there is an  $x \in X$  such that  $y = f(x)$ .
- $f$  is a *bijection* or *1-1 correspondence* if  $f$  is both injective and surjective. Equivalently,  $f$  is a bijection if and only if it has an inverse  $f^{-1}: Y \rightarrow X$ .

DEFINITION 4.2. A set  $X$  is *finite* if for some  $n \in \mathbb{N}_0$  there is a bijection  $\mathbf{n} \rightarrow X$ .  $X$  is *infinite* if it is not finite.

The next result is a formal version of what is usually called the *Pigeonhole Principle*.

THEOREM 4.3 (Pigeonhole Principle: first version).

- If there is an injection  $\mathbf{m} \rightarrow \mathbf{n}$  then  $m \leq n$ .
- If there is a surjection  $\mathbf{m} \rightarrow \mathbf{n}$  then  $m \geq n$ .
- If there is a bijection  $\mathbf{m} \rightarrow \mathbf{n}$  then  $m = n$ .

PROOF.

a) We will prove this by Induction on  $n$ . Consider the statement

$$P(n): \text{ For } m \in \mathbb{N}_0, \text{ if there is a injection } \mathbf{m} \rightarrow \mathbf{n} \text{ then } m \leq n.$$

When  $n = 0$ , there is exactly one function  $\emptyset \rightarrow \emptyset$  (the identity function) and this is a bijection; if  $m > 0$  then there are no functions  $\mathbf{m} \rightarrow \emptyset$ . So  $P(0)$  is true.

Suppose that  $P(k)$  is true for some  $k \in \mathbb{N}_0$  and let  $f: \mathbf{m} \rightarrow \mathbf{k} + \mathbf{1}$  be an injection. We have two cases to consider: (i)  $k + 1 \in \text{im } f$ , (ii)  $k + 1 \notin \text{im } f$ .

(i) For some  $r \in \mathbf{m}$  we have  $f(r) = k + 1$ . Consider the function  $g: \mathbf{m} - \mathbf{1} \rightarrow \mathbf{k}$  given by

$$g(j) = \begin{cases} f(j) & \text{if } 0 \leq j < r, \\ f(j + 1) & \text{if } r < j \leq m. \end{cases}$$

Then  $g$  is an injection, so by the assumption that  $m - 1 \leq k$ , hence  $m \leq k + 1$ .

(ii) Consider the function  $h: \mathbf{m} \rightarrow \mathbf{k}$  given by  $h(j) = f(j)$ . Then  $h$  is an injection, and by the

assumption that  $P(k)$  is true,  $m \leq k$  and so  $m \leq k + 1$ .

In either case we have established that  $P(k) \rightarrow P(k + 1)$ .

By PMI,  $P(n)$  is true for all  $n \in \mathbb{N}_0$ .

b) This time we proceed by Induction on  $m$ . Consider the statement

$Q(m)$ : For  $n \in \mathbb{N}_0$ , if there is a surjection  $\mathbf{m} \rightarrow \mathbf{n}$  then  $m \geq n$ .

When  $m = 0$ , there is exactly one function  $\emptyset \rightarrow \emptyset$  (the identity function) and this is a bijection; if  $n > 0$  there are no surjections  $\emptyset \rightarrow \mathbf{n}$ . So  $Q(0)$  is true.

Suppose that  $Q(k)$  is true for some  $k \in \mathbb{N}_0$  and let  $f: \mathbf{k} + \mathbf{1} \rightarrow \mathbf{n}$  be a surjection. Let  $f': \mathbf{k} \rightarrow \mathbf{n}$  be the restriction of  $f$  to  $\mathbf{k}$ , i.e.,  $f'(j) = f(j)$  for  $j \in \mathbf{k}$ . There are two cases to deal with: (i)  $f'$  is a surjection, (ii)  $f'$  is not a surjection.

(i) By the assumption that  $Q(k)$  is true,  $k \geq n$  which implies that  $k + 1 \geq n$ .

(ii) There must be exactly one  $s \in \mathbf{n}$  not in  $\text{im } f'$ . Define  $g: \mathbf{k} \rightarrow \mathbf{n} - \mathbf{1}$  by

$$g(j) = \begin{cases} f'(j) & \text{if } 0 \leq f'(j) < s, \\ f'(j) - 1 & \text{if } s < f'(j) \leq n. \end{cases}$$

Then  $g$  is a surjection, so by the assumption that  $Q(k)$  is true,  $k \leq n - 1$ , hence  $k + 1 \leq n$ .

In either case, we have established that  $Q(k) \rightarrow Q(k + 1)$ .

By PMI,  $Q(n)$  is true for all  $n \in \mathbb{N}_0$ .

c) This follows from (a) and (b) since a bijection is both injective and surjective.  $\square$

**COROLLARY 4.4.** *Suppose that  $X$  is a finite set and suppose that there are bijections  $\mathbf{m} \rightarrow X$  and  $\mathbf{n} \rightarrow X$ . Then  $m = n$ .*

**PROOF.** Let  $f: \mathbf{m} \rightarrow X$  and  $g: \mathbf{n} \rightarrow X$  be bijections. Using the inverse  $g^{-1}: X \rightarrow \mathbf{n}$  which is also a bijection, we can form a bijection  $h = g^{-1} \circ f: \mathbf{m} \rightarrow \mathbf{n}$ . By part (c),  $m = n$ .  $\square$

For a finite set  $X$ , the unique  $n \in \mathbb{N}_0$  for which there is a bijection  $\mathbf{n} \rightarrow X$  is called the *cardinality* of  $X$ , denoted  $|X|$ . If  $X$  is infinite then we sometimes write  $|X| = \infty$ , while if  $X$  is finite we write  $|X| < \infty$ .

We reformulate Theorem 4.3 without proof to give some important facts about cardinalities of finite sets.

**THEOREM 4.5 (Pigeonhole Principle).** *Let  $X, Y$  be two finite sets.*

- a) *If there is an injection  $X \rightarrow Y$  then  $|X| \leq |Y|$ .*
- b) *If there is a surjection  $X \rightarrow Y$  then  $|X| \geq |Y|$ .*
- c) *If there is a bijection  $X \rightarrow Y$  then  $|X| = |Y|$ .*

The name Pigeonhole Principle comes from the use of this when distributing  $m$  letters into  $n$  pigeonholes. If each pigeonhole is to receive at most one letter,  $m \leq n$ ; if each pigeonhole is to receive at least one letter,  $m \geq n$ .

Let  $X$  be a set and  $P \subseteq X$ . Then  $P$  is a *proper subset* of  $X$  if  $P \neq X$ , i.e., there is an element  $x \in X$  with  $x \notin P$ .

Notice that if  $X$  is a finite set and  $S$  a subset, then the inclusion function  $\text{inc}: S \rightarrow X$  given by  $\text{inc}(j) = j$  is an injection. So we must have  $|S| \leq |X|$ . If  $P$  is a proper subset then we have  $|P| < |X|$  and this implies that there can be no injection  $X \rightarrow P$  nor a surjection  $P \rightarrow X$ . These conditions actually characterise finite sets. In the next section we investigate how to recognise infinite sets.

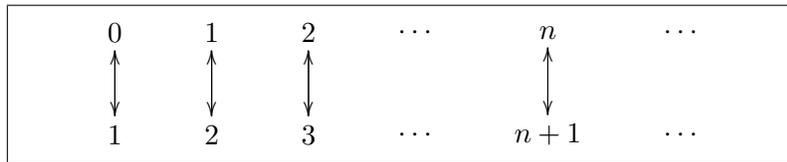
## 2. Infinite sets

THEOREM 4.6. Let  $X$  be a set.

- $X$  is infinite if and only if there is an injection  $X \rightarrow P$  where  $P \subseteq X$  is a proper subset.
- $X$  is infinite if and only if there is a surjection  $Q \rightarrow X$  where  $Q \subseteq X$  is a proper subset.
- $X$  is infinite if and only if there is an injection  $\mathbb{N}_0 \rightarrow X$ .
- $X$  is infinite if and only if there is a subset  $T \subseteq X$  and an injection  $\mathbb{N}_0 \rightarrow T$ .

EXAMPLE 4.7. The set of all natural numbers  $\mathbb{N}_0 = \{0, 1, 2, \dots\}$  is infinite.

SOLUTION. Let us take the subset  $P = \{1, 2, 3, \dots\}$  and define a function  $f: \mathbb{N}_0 \rightarrow P$  by  $f(n) = n + 1$ .



If  $f(m) = f(n)$  then  $m + 1 = n + 1$  so  $m = n$ , hence  $f$  is injective. If  $k \in P$  then  $k \geq 1$  and so  $(k - 1) \geq 0$ , implying  $(k - 1) \in \mathbb{N}_0$  whence  $f(k - 1) = k$ . Thus  $f$  is also surjective, hence bijective.  $\square$

EXAMPLE 4.8. Show that there are bijections between the set of all natural numbers  $\mathbb{N}_0$  and each of the sets

$$S_1 = \{2n : n \in \mathbb{N}_0\}, \quad S_2 = \{2n + 1 : n \in \mathbb{N}_0\}, \quad S_3 = \{3n : n \in \mathbb{N}_0\}.$$

In each case find a bijection and its inverse.

SOLUTION. For  $S_1$ , let  $f_1: \mathbb{N}_0 \rightarrow S_1$  be given by  $f_1(n) = 2n$ . Then  $f_1$  is a bijection: it is injective since  $2n_1 = 2n_2$  implies  $n_1 = n_2$ , and surjective since given  $2m \in S_1$ ,  $f_1(m) = 2m$ . The inverse function is given by  $f_1^{-1}(k) = k/2$ .

For  $S_2$ , let  $f_2: \mathbb{N}_0 \rightarrow S_2$  be given by  $f_2(n) = 2n + 1$ . Then  $f_2$  is a bijection: it is injective ( $2n_1 + 1 = 2n_2 + 1$  implies  $n_1 = n_2$ ) and surjective since given  $2m + 1 \in S_2$ ,  $f_2(m) = 2m + 1$ . The inverse function is given by  $f_2^{-1}(k) = (k - 1)/2$ .

For  $S_3$ , let  $f_3: \mathbb{N}_0 \rightarrow S_3$  be given by  $f_3(n) = 3n$ . Then  $f_3$  is a bijection: it is injective since  $3n_1 = 3n_2$  implies  $n_1 = n_2$ , and surjective since given  $3m \in S_3$ ,  $f_3(m) = 3m$ . The inverse function is given by  $f_3^{-1}(k) = k/3$ .  $\square$

Notice that each of the sets  $S_1, S_2, S_3$  is a proper subset of  $\mathbb{N}_0$ , yet each is in 1-1 correspondence with  $\mathbb{N}_0$  itself.

## 3. Countable sets

DEFINITION 4.9. A set  $X$  is *countable* if there is a bijection  $S \rightarrow X$  where either  $S = \mathbf{n}$  for some  $n \in \mathbb{N}_0$  or  $S = \mathbb{N}_0$ . A countable infinite set is said to be *countably infinite* or *of cardinality*  $\aleph_0$ . An infinite set which is not countable is said to be *uncountable*.

EXAMPLE 4.10. The following sets are countably infinite.

- Any infinite subset  $S \subseteq \mathbb{N}_0$ .
- $X \cup Y$  where  $X, Y$  are countably infinite.
- $X \cup Y$  where  $X$  is countably infinite and  $Y$  is finite.
- The set of all ordered pairs of natural numbers

$$\mathbb{N}_0 \times \mathbb{N}_0 = \{(m, n) : m, n \in \mathbb{N}_0\}.$$

e) The set of all positive rational numbers

$$\mathbb{Q}^+ = \left\{ \frac{a}{b} : a, b \in \mathbb{N}_0, a, b > 0 \right\}.$$

SOLUTION.

a) Since  $S$  is infinite it cannot be empty. Let  $S_0 = S$ . By WOP,  $S_0$  has a least element  $s_0$  say. Now consider the set  $S_1 = S - \{s_0\}$ ; this is not empty since otherwise  $S$  would be finite. Again WOP ensures that there is a least element  $s_1 \in S_1$ . Continuing, we can construct a sequence  $s_0, s_1, \dots, s_n, \dots$  of elements in  $S$  with  $s_n$  the least element of  $S_n = S - \{s_0, s_1, \dots, s_{n-1}\}$  which is never empty. Notice in particular that

$$s_0 < s_1 < \dots < s_n < \dots,$$

from which it easily follows that  $s_n \geq n$ . If  $s \in S$ , then for some  $m \in \mathbb{N}_0$  must satisfy  $m \geq s$ , so by construction of the  $s_n$  we must have  $s = s_{m_0}$  for some  $m_0$ . Hence

$$S = \{s_n : n \in \mathbb{N}_0\}.$$

Now define a function  $f: \mathbb{N}_0 \rightarrow S$  by  $f(n) = s_n$ ; this is easily seen to be a bijection.

b) The simplest case is where  $X \cap Y = \emptyset$ . Then given bijections  $f: \mathbb{N}_0 \rightarrow X$  and  $g: \mathbb{N}_0 \rightarrow Y$  we construct a function  $h: \mathbb{N}_0 \rightarrow X \cup Y$  by

$$h(n) = \begin{cases} f\left(\frac{n}{2}\right) & \text{if } n \text{ is even,} \\ g\left(\frac{n-1}{2}\right) & \text{if } n \text{ is odd.} \end{cases}$$

Then  $h$  is a bijection.

If  $Z = X \cap Y$  and  $Y - Z$  are both countably infinite, let  $f: \mathbb{N}_0 \rightarrow X$  and  $g: \mathbb{N}_0 \rightarrow Y - Z$  be bijections. Then we define  $h: \mathbb{N}_0 \rightarrow X \cup Y$  by

$$h(n) = \begin{cases} f\left(\frac{n}{2}\right) & \text{if } n \text{ is even,} \\ g\left(\frac{n-1}{2}\right) & \text{if } n \text{ is odd.} \end{cases}$$

This is again a bijection.

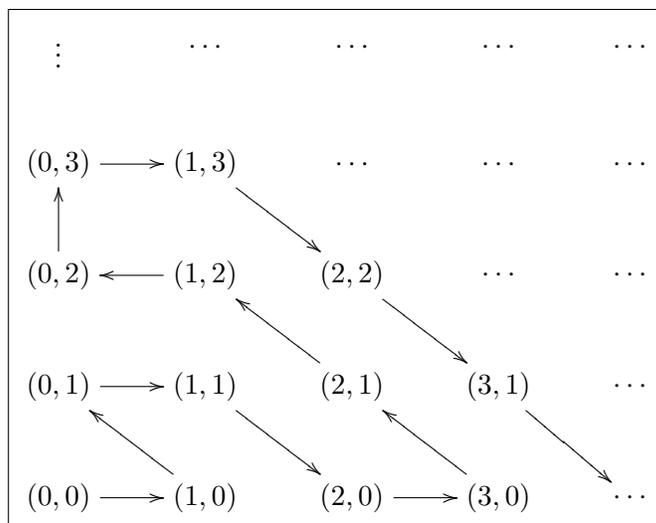
The case where one of  $X - X \cap Y$  and  $Y - X \cap Y$  is finite is easy to deal with by the method used for (c).

c) Since  $Y$  is finite so is  $Y - X \cap Y \subseteq Y$ . Let  $f: \mathbb{N}_0 \rightarrow X$  and  $g: \mathbf{m} \rightarrow Y - X \cap Y$  be bijections. Define  $h: \mathbb{N}_0 \rightarrow X \cup Y$  by

$$h(n) = \begin{cases} g(n-1) & \text{if } 1 \leq n \leq m, \\ f(n-m-1) & \text{if } m < n. \end{cases}$$

Then  $h$  is a bijection.

d) Plot each pair  $(a, b)$  as the point in the  $xy$ -plane with coordinates  $(a, b)$ ; such points are all those with natural number coordinates. Starting at  $(0, 0)$  we can now trace out a path passing through all of these points and we can arrange to do this without ever recrossing such a point.

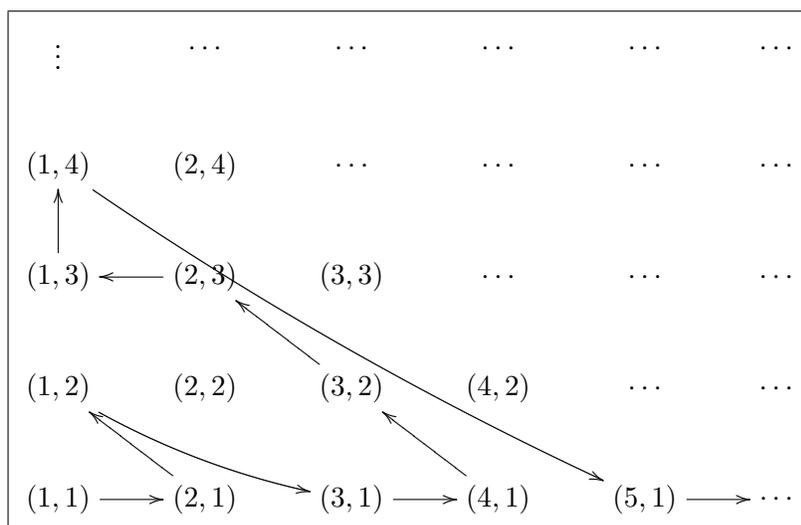


This gives a sequence  $\{(r_n, s_n)\}_{0 \leq n}$  of elements of  $\mathbb{N}_0 \times \mathbb{N}_0$  which contains every natural number exactly once. The function

$$f: \mathbb{N}_0 \longrightarrow \mathbb{N}_0 \times \mathbb{N}_0; \quad f(n) = (r_n, s_n),$$

is a bijection.

e) This is demonstrated in a similar way to (d) but is slightly more involved. For each  $a/b \in \mathbb{Q}^+$ , we can assume that  $a, b$  are coprime (*i.e.*, have no common factors) and plot it as the point in the  $xy$ -plane with coordinates  $(a, b)$ . Starting at  $(1, 1)$  we can now trace out a path passing through all of these points with coprime positive natural number coordinates and can even arrange to do this without ever recrossing such a point.



This gives us a sequence  $\{r_n\}_{0 \leq n}$  of elements of  $\mathbb{Q}^+$  which contains every element exactly once. The function

$$f: \mathbb{N}_0 \longrightarrow \mathbb{Q}^+; \quad f(n) = r_n,$$

is a bijection. □

#### 4. Power sets and their cardinality

For two sets  $X$  and  $Y$ , let

$$Y^X = \{f : f: X \longrightarrow Y \text{ is a function}\}.$$

EXAMPLE 4.11. Let  $X$  and  $Y$  be finite sets. Then  $Y^X$  is finite and has cardinality

$$|Y^X| = |Y|^{|X|}.$$

SOLUTION. Suppose that the distinct elements of  $X$  are  $x_1, \dots, x_m$  where  $m = |X|$  and those of  $Y$  are  $y_1, \dots, y_n$  where  $n = |Y|$ . A function  $f: X \rightarrow Y$  is determined by specifying the values of the  $m$  elements  $f(x_1), \dots, f(x_m)$  of  $Y$ . Each  $f(x_k)$  can be chosen in  $n$  ways so the total number of choices is  $n^m$ . Hence  $|Y^X| = n^m$ .  $\square$

A particular case of this occurs when  $Y$  has two elements, e.g.,  $Y = \{0, 1\}$ . The set  $\{0, 1\}^X$  is called the *power set* of  $X$ , and has  $2^{|X|}$  elements and indeed it is often denoted  $2^{|X|}$ . It has another important interpretation.

For any set  $X$ , we can consider the set of all its subsets

$$\mathcal{P}(X) = \{U : U \subseteq X \text{ is a subset}\}.$$

Before stating and proving our next result we introduce the *characteristic* or *indicator function* of a subset  $U \subseteq X$ ,

$$\chi_U: X \rightarrow \{0, 1\}; \quad \chi_U(x) = \begin{cases} 1 & \text{if } x \in U, \\ 0 & \text{if } x \notin U. \end{cases}$$

THEOREM 4.12. For a set  $X$ , the function

$$\Theta: \mathcal{P}(X) \rightarrow \{0, 1\}^X; \quad \Theta(U) = \chi_U,$$

is a bijection.

PROOF. The indicator function of a subset  $U \subseteq X$  is clearly determined by  $U$ , so  $\Theta$  is well defined. Also, a function  $f \in \{0, 1\}^X$  determines a corresponding subset of  $X$

$$U_f = \{x \in X : f(x) = 1\}$$

with  $\chi_{U_f} = f$ . This shows that  $\Theta$  is a bijection whose inverse function satisfies

$$\Theta^{-1}(f) = U_f.$$

$\square$

EXAMPLE 4.13. If  $X$  is finite then  $\mathcal{P}(X)$  is finite with cardinality  $|\mathcal{P}(X)| = 2^{|X|}$ .

PROOF. This follows from Example 4.11.  $\square$

Using the standard finite sets  $\mathbf{n} = \{1, \dots, n\}$  ( $n \in \mathbb{N}_0$ ) we have

$$|\mathcal{P}(\mathbf{0})| = 2^0 = 1, \quad |\mathcal{P}(\mathbf{1})| = 2^1 = 2, \quad |\mathcal{P}(\mathbf{2})| = 2^2 = 4, \quad |\mathcal{P}(\mathbf{3})| = 2^3 = 8, \quad \dots$$

where

$$\begin{aligned} \mathcal{P}(\mathbf{0}) &= \{\emptyset\}, \\ \mathcal{P}(\mathbf{1}) &= \{\emptyset, \{1\}\}, \\ \mathcal{P}(\mathbf{2}) &= \{\emptyset, \{1\}, \{2\}, \{1, 2\}\}, \\ \mathcal{P}(\mathbf{3}) &= \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}, \\ &\vdots \end{aligned}$$

We will now see that for any set  $X$  the power set  $\mathcal{P}(X)$  is always ‘bigger’ than  $X$ .

THEOREM 4.14 (Russell’s Paradox). For a set  $X$ , there is no surjection  $X \rightarrow \mathcal{P}(X)$ .

PROOF. Suppose that  $g: X \rightarrow \mathcal{P}(X)$  is a surjection.

Consider the subset

$$W = \{x \in X : x \notin g(x)\} \subseteq X.$$

Then by surjectivity of  $g$  there is a  $w \in X$  such that  $g(w) = W$ . If  $w \in W$ , then by definition of  $W$  we must have  $w \notin g(w) = W$ , which is impossible. On the other hand, if  $w \notin W$ , then  $w \in g(w) = W$  and again this is impossible. But then  $w$  cannot be in  $W$  or the complement  $X - W$ , contradicting the fact that every element of  $X$  has to be in one or other of these subsets since  $X = W \cup (X - W)$ . Thus no such surjection can exist.  $\square$

When  $X$  is finite, this result is not surprising since  $2^n > n$  for  $n \in \mathbb{N}_0$ . For  $X$  an infinite set, it leads to the idea that there are different ‘sizes’ of infinity. Before showing how this result allows us to determine some concrete examples, we give a generalization.

COROLLARY 4.15. *Let  $X$  and  $Y$  be sets and suppose that  $Y$  has a subset  $Z \subseteq Y$  which admits a bijection  $g: Z \rightarrow \mathcal{P}(X)$ . Then there is no surjection  $X \rightarrow Y$ .*

PROOF. Suppose that  $f: X \rightarrow Y$  is a surjection. Choosing any element  $p \in \mathcal{P}(X)$  and defining the function

$$H: X \rightarrow \mathcal{P}(X); \quad h(x) = \begin{cases} g(f(x)) & \text{if } f(x) \in Z, \\ p & \text{if } f(x) \notin Z, \end{cases}$$

we easily see that  $h$  is a surjection, contradicting Russell’s Paradox. Thus no such surjection can exist.  $\square$

## 5. The real numbers are uncountable

THEOREM 4.16 (Cantor). *The set of real numbers  $\mathbb{R}$  is uncountable, i.e., there is no bijection  $\mathbb{N}_0 \rightarrow \mathbb{R}$ .*

PROOF. Suppose that  $\mathbb{R}$  is countable and therefore the obviously infinite subset  $(0, 1] \subseteq \mathbb{R}$  is countable. Then we can list the elements of  $(0, 1]$ :

$$q_0, q_1, \dots, q_n, \dots$$

For each  $n$  we can uniquely express  $q_n$  as a non-terminating expansion infinite decimal

$$q_n = 0.q_{n,1}q_{n,2} \cdots q_{n,k} \cdots,$$

where for each  $k$ ,  $q_{n,k} = 0, 1, \dots, 9$  and for every  $k_0$  there is always a  $k > k_0$  for which  $q_{n,k} \neq 0$ .

Now define a real number  $p \in (0, 1]$  by requiring its decimal expansion

$$p = 0.p_1p_2 \cdots p_k \cdots$$

to have the property that for each  $k \geq 1$ ,

$$p_k = \begin{cases} 1 & \text{if } q_{k-1,k} \neq 1, \\ 2 & \text{if } q_{k-1,k} = 1. \end{cases}$$

Notice that this is also non-terminating. Then  $p \neq q_1$  since  $p_1 \neq q_{0,1}$ ,  $p \neq q_2$  since  $p_2 \neq q_{1,2}$ , etc. So  $p$  cannot be in the list of  $q_n$ ’s, contradicting the assumption that  $(0, 1]$  is countable.  $\square$

The method of proof used here is often referred to as *Cantor’s diagonalization argument*. In particular this shows that  $\mathbb{R}$  is much bigger than the familiar subset  $\mathbb{Q} \subseteq \mathbb{R}$ , however it can be hard to identify particular elements of the complement  $\mathbb{R} - \mathbb{Q}$ . In fact the subset of all *real algebraic numbers* is countable, where such a real number is a root of a monic polynomial of positive degree,

$$X^n + a_{n-1}X^{n-1} + \cdots + a_0 \in \mathbb{Q}[X].$$

**Problem Set 4**

4-1. Show that each of the following sets is countable:

- a)  $\mathbb{Z}$ , the set of all integers;
- b)  $\{n^2 : n \in \mathbb{Z}\}$ , the set of all integers which are squares of integers;
- c)  $\{n \in \mathbb{Z} : n \neq 0\}$ , the set of all non-zero integers;
- d)  $\mathbb{Q}$ , the set of all rational numbers;
- e)  $\{x \in \mathbb{R} : x^2 \in \mathbb{Q}\}$ , the set of all real numbers which are square roots of rational numbers.

4-2. Show that a subset of a countable set is countable.

4-3. Let  $X$  be a countable set. If  $Y$  is a finite set, show that the cartesian product

$$X \times Y = \{(x, y) : x \in X, y \in Y\}$$

is countable.

Use Example 4.10(d) or a modification of its proof to show that this is still true if  $Y$  is countably infinite.

# Index

- 1-1, 53
  - correspondence, 53
- action, 30
  - group, 38
- alternating group, 31
- arithmetic function, 47
  
- back-substitution, 5
- bijection, 53
- binary operation, 29
- Burnside Formula, 40
  
- Cantor's diagonalization argument, 59
- cardinality, 54
  - $\aleph_0$ , 55
- characteristic function, 58
- common
  - divisor, 3
  - factor, 3
- commutative ring, 3, 9
- composite, 11
- congruence class, 9
- congruent, 8
- continued fraction
  - expansion, 16
  - finite, 16, 17
  - generalized finite, 18
  - infinite, 17
- convergent, 17, 18
- convolution, 48
- coprime, 3
- coset
  - left, 37
- countable
  - set, 55
- countably infinite set, 55
- cycle
  - decomposition, disjoint, 32
  - notation, 32
  - type, 32
  
- degree, 13
- dihedral group, 34
- Diophantine problem, 22
- disjoint cycle decomposition, 32
- divides, 3
  
- divisor, 3
  - common, 3
  - greatest common, 3
  
- element
  - greatest, 1
  - least, 1
  - maximal, 1
  - minimal, 1
- equivalence relation, 8
- Euclid's Lemma, 11
- Euclidean Algorithm, 4
- Euler  $\varphi$ -function, 36
- even permutation, 31
  
- factor
  - common, 3
  - highest common, 3
- factorization
  - prime, 12
  - prime power, 12
- Fermat's Little Theorem, 13
- Fibonacci sequence, 18
- finite
  - continued fraction, 17
  - set, 53
- fixed
  - point set, 40
  - set, 40
  
- function
  - arithmetic, 47
  - characteristic, 58
  - indicator, 58
- fundamental solution, 24
- Fundamental Theorem of Arithmetic, 11
  
- generator, 36
- greatest
  - common divisor, 3
  - element, 1
  
- group, 29
  - action, 38
  - action, transitive, 40
  - alternating, 31
  - dihedral, 34
  - permutation, 29, 30
  - symmetric, 30

- highest
  - common factor, 3
- Idiot's Binomial Theorem, 14
- index, 37
- indicator function, 58
- infinite
  - continued fraction, 17
  - set, 1, 53
- injection, 53
- integer, 3
- inverse, 9
- irrational, 13
  
- Lagrange's Theorem, 37
- least
  - element, 1
- left coset, 37
- Long Division Property, 3
  
- Möbius function, 47
- maximal
  - element, 1
- Maximal Principle (MP), 2
- minimal
  - element, 1
- multiplicative, 47
  - strictly, 47
  
- natural numbers, 1
- numbers
  - natural, 1
  
- odd permutation, 31
- one-one, 53
- onto, 53
- orbit, 38
- orbit-stabilizer equation, 40
- order, 14, 29
  
- Pell's Equation, 23
- period, 22
- permutation
  - even, 31
  - group, 29, 30
  - matrix, 32
  - odd, 31
  - sign of  $a$ , 31
- Pigeonhole Principle, 53
- power set, 58
- prime, 11
  - factorization, 12
  - power factorization, 12
- primitive root, 15
- Principle of Mathematical Induction (PMI), 1
- proper subgroup, 35
- proper subset, 54
  
- real algebraic numbers, 59
- recurrence relation, 18
  
- represents, 16, 18
- residue class, 9
- root, 13
  - primitive, 15
  
- set
  - countable, 55
  - countably infinite, 55
  - finite, 53
  - fixed, 40
  - fixed point, 40
  - infinite, 1, 53
  - of cardinality  $\aleph_0$ , 55
  - power, 58
  - standard, 30
  - uncountable, 55
- sign of a permutation, 31
- solution
  - fundamental, 24
- stabilizer, 38
- standard set, 30
- strictly multiplicative, 47
- subgroup, 35
  - generated by  $g$ , 35
  - proper, 35
- subset
  - proper, 54
- surjection, 53
- symmetric group, 30
- symmetry, 33
  
- tabular method, 6
- transitive, 1
  - group action, 40
- transposition, 33
  
- uncountable
  - set, 55
  
- Well Ordering Principle (WOP), 1
- Wilson's Theorem, 15