Contents

1 Introduction

The following are "skeleton notes" for the course, giving the basic definitions and some simple proofs. During each section we will extend beyond what is covered here, but it will be helpful for you to read ahead in the current section so that you have some idea about what we are going to be covering next. Copies of these notes will be distributed to you all during the course. Use of Maple is encouraged and the relevant code for loading the number theory package is `with(NumberTheory);`.

As its name implies, we shall be studying numbers, in particular the natural numbers $1, 2, 3, \ldots$, normally denoted by $\mathbb{N}$. We shall use negatives of the natural numbers and zero as well, this set being the integers $\mathbb{Z}$ (from the German for "count", Zählen). We shall, hopefully, never have to use the real numbers $\mathbb{R}$ and only rarely use the rationals, $\mathbb{Q}$.

1.1 Arithmetic in Other Number Bases

We shall first divert to investigate what happens if we try to use other base systems than decimal in order to do mathematics. A base system is when, instead of representing a number written in base 10 as $n_3n_2n_1 = n_3 \times 100 + n_2 \times 10 + n_1$, we use a base b not equal to 10. So the number which is 35 in decimal is 43 in octal (base 8), since $35 = 4 \times 8 + 3$. Note that Maple usually gives the answer in the reverse order: the output from `convert(35,base,8);` is `[3,4]`.

For bases greater than 10 we have to introduce letters to stand for the numbers 10, 11,.... The most common base systems are bases 2, 8, 12 and 16, and in the latter two we use δ (called "dec") and ϵ (called "el") for 10 and 11 in base 12 (dozenal/duodecimal), and A, B, ..., F for 10 through to 15 in base 16 (hexadecimal).

For instance, this is the digits addition and multiplication tables for base 8:

+	1	2	3	4	5	6	7
1	2	3	4	5	6	7	10
2	3	4	5	6	7	10	11
3	4	5	6	7	10	11	12
4	5	6	7	10	11	12	13
5	6	7	10	11	12	13	14
6	7	10	11	12	13	14	15
7	10	11	12	13	14	15	16

×	1	2	3	4	5	6	7
1	1	2	3	4	5	6	7
2	2	4	6	10	12	14	16
3	3	6	11	14	17	22	25
4	4	10	14	20	24	30	34
5	5	12	17	24	31	36	43
6	6	14	22	30	36	44	52
7	7	16	25	34	43	52	61

In order to add, multiply and divide we can use the familiar carry rules from ordinary arithmetic, but we must be careful not to get confused. With practice one can easily learn your δ times table, which makes life much easier. For instance, in octal we can evaluate 263×154 as follows:

		2	6	3
	×	1	5	4
2	6	3	0	0
1	5	7	7	0
	1	3	1	4
4	5	6	0	4

Note that this product is equivalent in decimal to $(128+48+3)\times(64+40+4) = 179\times108 = 19332$ and 45604 converted from octal to decimal is $4\times8^4+5\times8^3+6\times8^2+4 = 16384+2560+384+4 = 19332$! Also, we can use familiar tricks from decimal, such as noting that $263 = 300 - 15$ in octal and so $263 \times 154 = 300 \times 154 - 15 \times 154 = 50400 - 2574 = 45604$ again.

Exercise 1 *What is 17+66 in base 8 ?* $4F \times 16$ *in base 16 ?* $56 \div 3$ *in base 12 ?*

In fact decimal as our base for arithmetic is completely arbritrary and everything we do in this course can work in any base, as we may see in future weeks. The main constraint is that Maple is designed to work most easily in decimal... We will now go back to decimal for most of our work, but algebra is the same in any base.

1.2 Divisibility

Theorem 1.1 *For any two integers* $a \neq 0$ *and* b *we can find unique integers* q *and* r*;* $0 \leq r < |a|$ *(the* **quotient** *and the* **remainder** *of* b *on division by* a*) such that* $b = aq + r$*.*

Proof: If $0 \leq b < a$ then we have the only possibility as $q = 0$ and $r = b$. For $b \geq a$ we can use induction as follows: consider $b - a$, which is a non-negative number: by the induction hypothesis it can be written uniquely as $b - a = aq' + r'$. We can then rearrange this and, using the fact that both r and r' are between 0 and $|a| - 1$ (so that $-a + 1 \leq r - r' \leq a - 1$). Since $r - r'$ is an integer multiple of a we get that $r = r'$ and so $q = 1 + q'$. It is possible to verify their uniqueness using similar logic. For $b < 0$ we know from the previous lines that

$$-b = -aq - r = \begin{cases} a \times (-q) & + & 0 & r = 0 \\ a \times (-q-1) & + & (a-r) & r \neq 0 \end{cases}$$

which values of the quotient and remainder satisfy the criteria. We can similarly cope with the cases in which a is negative. ◇

We shall say that a *divides* b (written $a|b$) if the remainder of b on division by a is zero (so that there exists an integer x such that $a \times x = b$). We shall usually suppress the multiplication sign so that the above equation would read $ax = b$. Thus, for example, $2|8$, $3|3$, $-5|15$ and $1|-3$. We shall also use the same symbol with a line through it ($\not|$) to signify *doesn't divide*, so that $3 \not| 1$, $5 \not| 9$ and $0 \not| 11$.

Lemma 1 *For all* $n \in \mathbb{Z}$ *we have* $1|n$ *and* $n|0$*.*

Proof: The x we require in the former case is n ($1 \times n = n$) and in the latter it is 0. ◇

Lemma 2 *If* $a|b$ *and* $b|c$ *then* $a|c$*.*

Proof: Since $a|b$ we have $ax = b$ and similarly $by = c$. Hence $axy = by = c$ and since xy is an integer too we have $a|c$ as required. ◇

Exercise 2 *Show that if* $d|b$ *and* $d|c$ *then* $d|(kb + lc)$*.*

1.3 GCD and LLC

In the same way as in exercise 2 we can show that if $a|b_1, \ldots, a|b_n$ then $a|k_1b_1 + \ldots + k_nb_n$ for any integers $k_1, \ldots k_n$. We define a *linear combination* of the integers to be a sum of the form $k_1b_1 + \ldots + k_nb_n$ and a *least linear combination* (usually abbreviated to l.l.c) to be a set of values for the k_i so that the value of the linear combination is as small a positive number as possible. The *greatest common divisor* (g.c.d.) of two integers a and b (usually written $\gcd(a, b)$ or (a, b) if there is no danger of confusion) is defined as the largest integer d such that $d|a$ and $d|b$.

Theorem 1.2 $\mathrm{llc(a, b)} = \mathrm{gcd(a, b)}$.

Proof: Let $c = \mathrm{llc(a, b)}$ and $d = \gcd(a, b)$. From exercise 2 and the definition of the gcd we see that $d|c$ and so $d \leq c$.

By theorem 1.1 we may write $a = sc + u$ and $b = tc + v$, with $0 \leq u, v < c$. But since c is a linear combination of a and b we see that both u and v are linear combinations of a and b, and, moreover, are less than c. Since c is the least linear combination of a and b we see that $u = v = 0$ and so $c|a$ and $c|b$. Hence c is a common divisor of a and b and so must divide d. Hence $c \leq d$ and so $c = d$. ◇

We note that if a and b are not both zero then $d := \gcd(a, b)$ is bounded above by $\min(a, b)$ (since $d \leq a$ and $d \leq b$) and below by 1 (since $1|a$ and $1|b$) and so is well defined. If $(a, b) = 1$ then we say that a and b are *relatively prime* or *co-prime*.

Theorem 1.3 *If* $d = (a, b)$ *then* $(\frac{a}{d}, \frac{b}{d}) = 1$.

Proof: Let $c := (\frac{a}{d}, \frac{b}{d})$. By the remarks before theorem 1.3 we see that $c \geq 1$. We now show that $c \leq 1$ to get the required result.
By definition there exists integers s and t such that $cs = \frac{a}{d}$ and $ct = \frac{b}{d}$. Hence $cd|a$ and $cd|b$ and so cd is a common divisor of a and b and so $cd \leq d$. Since $d \geq 1$ we see that $c = 1$, as required. ◇

For Maple, the command for integer greatest common divisors is `igcd`.

1.4 The Euclidean Algorithm

Lemma 3 *If* $b = aq + r$ *then* $\gcd(a, b) = \gcd(a, r)$.

Proof: Let $d := \gcd(a, b) := (a, b)$. Since $r = a - qb$ and $d|a$ and $d|b$ we have, by exercise 2, $d|r$. If c is any common divisor of r and a then $c|aq + r = b$ and so $c \leq d$. Thus d is the greatest common divisor of a and r. ◇

Putting lemma 3 together with theorem 1.1 we get the following algorithm for finding the greatest common divisor of any two integers b and a:
find the quotient (q) and remainder (r) of the two given numbers.

1. if $r \neq 0$ repeat the algorithm to find (a, r).
2. if $r = 0$ then the previous remainder is (b, a) and stop the algorithm.

We normally tabulate the working as follows:

$$\begin{array}{rcll} b &=& aq+r, & 0 \leq r < a, \\ a &=& rq_1+r_1, & 0 \leq r_1 < r, \\ r &=& r_1q_2+r_2, & 0 \leq r_2 < r_1, \\ & \vdots & & \vdots \\ r_{k-1} &=& r_kq_{k+1}+r_{k+1}, & 0 \leq r_{k+1} < r_k, \\ r_k &=& r_{k+1}q_{k+2}. & \end{array}$$

and so $(b,a) = r_{k+1}$. We notice that since $a > r > r_1 > r_2 > \ldots \geq 0$ has a finite number of terms in the sequence the process is guaranteed to terminate.

For example, we shall find $(96, 21)$ using this method:

$$\begin{aligned} 96 &= 21 \times 4 + 12, \\ 21 &= 12 \times 1 + 9, \\ 12 &= 9 \times 1 + 3, \\ 9 &= 3 \times 3 + 0. \end{aligned}$$

Hence $(96, 21) = 3$.

Exercise 3 *Evaluate* $\gcd(77, 21)$*,* $\gcd(8, 38)$ *and* $\gcd(30, 48)$ *using the Euclidean algorithm. How about* $(1026, 703)$ *?*

By lemma 3 we do not actually have to take q and r to be the quotient and divisor of b and a, just two numbers satisfying that equation: sometimes it is useful to be able to replace "$+r_i$" by "$-r_i$" if the new value for r_i is closer to zero.

For instance we can get $\gcd(96, 21)$ in one step fewer this way:

$$\begin{aligned} 96 &= 21 \times 5 - 9, \\ 21 &= 9 \times 2 + 3, \\ 9 &= 3 \times 3. \end{aligned}$$

Exercise 4 *Repeat exercise 3 using this refinement of the technique to see how it can sometimes speed things up.*

We can also use the Euclidean algorithm in reverse to find a least linear combination of a and b: using our worked example again we see that $12 = 96 - 21 \times 4$ and $9 = 21 - 12 \times 1$. Thus

$$\begin{aligned} \mathrm{llc}(96, 21) = \gcd(96, 21) = 3 &= 21 - 9 \times 2, \\ &= 21 - (21 - 12 \times 1) \times 2, \\ &= 12 \times 2 - 21 \times 1 = 24 - 21 = 3, \\ &= (96 - 21 \times 4) \times 2 - 21 \times 2, \\ &= 96 \times 2 - 21 \times 9 = 192 - 189 = 3. \end{aligned}$$

The Maple command to verify this is `igcdex(96,21,'s','t'); s; t;`.

Note that there are an infinite number of pairs of coefficients which make the same least linear combination:

$$96 \times (-5) + 21 \times 23 = -480 + 483 = 3$$

Using the idea in this section we can prove the following:

Lemma 4 *If $d|ab$ and $(d,a)=1$ then $d|b$.*

Proof: We know there exist integers x and y so that

$$dx + ay = (d,a) = 1. \tag{1}$$

Multiplying both sides of (1) by b we get

$$bdx + aby = b$$

and we note that $d|bd$ and $d|ab$ so that $d|b$, as required. ◇

Exercise 5 *Find* $\mathrm{llc}(77,21)$, $\mathrm{llc}(8,38)$ *and* $\mathrm{llc}(30,48)$.

2 Primes and Factorisation

We say that a natural number p is *prime* if it has no positive divisors apart from itself and 1. A number which is not prime is called *composite* and 1, which is a special case, is neither; it is a *unit*.

2.1 Primes

Lemma 5 *Every natural number can be written as a product of primes.*

Proof: Any prime number can be said to be in this form (a product with one term) and 1 is often taken to be the product of no terms. If n is composite it must have some non-trivial number which divides it, and hence some prime p which divides it. We then consider $\frac{n}{p}$ which we can then check to see whether it is prime or composite and, if necessary, continue this process. ◇

Lemma 6 *If p is prime and $p|ab$ then $p|a$ or $p|b$ (or p divides both).*

Proof: Since p is prime its only positive divisors are 1 and p. Thus $(p,a)=1$ or p. In the former case we apply lemma 4 and in the latter we can easily see that $p|a$. ◇

Lemma 7 *If $p, q_1, q_2, \ldots, q_n$ are primes and $p|q_1q_2\ldots q_n$ then $p = q_k$ for some k.*

Proof: From lemma 6 we see that $p|q_1$ or $p|q_2\ldots q_n$. Since both p and all the q_is are prime we must have $p = q_1$ or $p|q_2\ldots q_n$, to which we can repeat the process if $n > 2$ and otherwise $p = q_2$. ◇

Theorem 2.1 (The Fundamental Theorem of Arithmetic) *Any positive integer can be written as a product of powers of distinct primes in a unique way.*

Proof: We saw in lemma 5 that any number has at least one way to express it as a product of (not necessarily distinct) primes. We shall show that, after collecting the primes together this representation is the only one possible. Consider

$$n = t_1 t_2 \dots t_r = q_1 q_2 \dots q_s \tag{2}$$

as two different prime representations of the same number. By lemma 7 we see that $t_1 = q_i$ for some i and divide both sides of (2) by their common prime. We repeat this r times and thus deduce that $r = s$ since we cannot run out of t_js before q_is, or vice versa, as that would leave a non-empty product of primes equal to 1. Hence every t_j is matched with some q_i and the representations differ only in the order of the primes. ◇

The most convenient notation for this product representing n is

$$\prod_{i=1}^{\infty} {p_i}^{\alpha_i},$$

where $p_1 = 2$, $p_2 = 3$, $p_3 = 5, \dots$ and α_i is the power of the i^{th} prime. The infinite product is acceptable to represent a finite number since only a finite number of terms in it are non-zero.

Exercise 6 *What are the prime power decompositions of 32, 6, 12, 77 and 79 ?*

Theorem 2.2 (Euclid) *There are an infinite number of primes.*

Proof: Suppose the set of primes is finite and is $\mathbb{P} := \{p_1, p_2, \dots, p_r\}$, say. Consider the (necessarily finite) number

$$n = (p_1 \times p_2 \times \dots \times p_r) + 1.$$

There must be a prime divisor p of n, but we see that if $p = p_k$ for some k $(1 \leq k \leq r)$ then $p | (n - p_1 \times p_2 \times \dots \times p_r) = 1$, a clear contradiction since all primes are greater than one. Hence p is a prime not in Π and so Π is not a finite set. ◇

2.2 Factorisation

Thus, given any set of primes it is always possible to find a new prime by forming the number n as in theorem 2.2. However, this number gets large very quickly and is an infeasible way to generate primes. In fact, if we plug the values into the formula in order starting with 2 we get this list:

$$2, 3, 7, 43, 13, 53, 5, 6221671, 38709183810571.$$

Just verifying that the last number is in fact prime would take a long time (without a computer ;) so we can chalk this method for generating prime numbers as a non-starter.

A second method, popularly known as the Sieve of Eratosthenes, is slightly more useful, in that it will give all the primes up to a certain value. We first need to prove the following result:

Lemma 8 *If n is composite then it has a prime divisor at most $\sqrt{n}$.*

Proof: Since n is composite it can be written as $n_1 n_2$, and one of these two integers is at most $\sqrt{n}$ (if not $n_1 > \sqrt{n}$ and $n_2 > \sqrt{n}$ and so $n > (\sqrt{n})^2 = n$, a contradiction). Say the smaller integer is n_1: if it is prime we are done; if it is composite then it has a prime divisor which is necessarily less than $n_1 \leq \sqrt{n}$. $\diamond$

We perform the Sieve of Eratosthenes as follows: given the number n we wish to find the primes up to we find the integer part of $\sqrt{n}$ and note it down as N, say. We then write the list of integers from 2 to n and proceed to cross some of them out according to this pattern:

1. find the first non-circled number in the list, call it a, and circle it,
2. erase all other multiples of a in the list,
3. repeat steps 1 and 2 while $a \leq N$.

We shall do a worked example for $n = 28$. We have $N = 5$:

2	3	5	7	9	11	13	15	17	19	21	23	25	27
2	**3**	5	7		11	13		17	19		23	25	
2	**3**	**5**	7		11	13		17	19		23		

and we can stop there since $7 > N$. As you can see it works fairly efficiently and, in modified forms, is still the main way to create large lists of prime numbers. Lemma 8 enables us to stop early since any numbers left in the list must be prime as otherwise they would have a prime factor less than $\sqrt{n}$ and we have removed all numbers for which this is so. However, it is not necessarily an efficient way to check whether a particular number is prime, once the numbers involved start to become large.

The idea behind it can be used to generate sets of prime numbers of a certain size though. If, for instance, we wish to find the first 10 primes after 300, we can list the numbers from 300 to 330, say, having first used the sieve as above to generate all primes from 2 to $\sqrt{330}$. We then identify which numbers in our set are multiples of these primes in turn, and cross them off the list. For instance, the final prime less than $\sqrt{330}$ is 17, and we evaluate $\frac{300}{17}$ on a calculator and note that it is $17.6470588235294117\ldots$. Hence the first multiple of 17 after 300 is $18 \times 17 = 306$, and we can cross this off the list (although it has already been crossed as it is divisible by both 2 and 3). We then add 17 to 306 to get 323 and we can cross this off also, and this one hasn't already been crossed since $323 = 17 \times 19$, and both are primes.

In practice I just write the odd numbers in the desired range since those ending in even numbers are obviously seen to be non prime. I then take coloured pencils and cross or circle the numbers divisible by 3, 5, 7 etc. in different colours, until all numbers are done. Those numbers which are left are the next primes after the start term, and if there are enough you are done. If there are too few, it is necessary to add another few numbers on the end and extend your markings from before to these numbers. Also remember that it is now necessary to check that the square root of the new largest number does not allow any more possible prime divisors than the previous number.

Other non-technical ways to test for primality include the use of tables (as in Appendix C of Dudley's book or the books which have nothing but lists of primes in) and computers (both Maple

and Matlab have a function called `isprime` which use probabilistic methods to check primality and Mathematica has a similar one called PrimeQ). Unfortunately, all are restricted, either by space or time constraints, to some extent, as there is no easy way to tell if any given number is or is not prime, as we shall see! Later in the course we shall use some of the techniques learned to enable us to study some of the better ways to test for primality.

3 Congruences

3.1 Introduction

In this section we shall develop a theory which will enable us to solve problems encountered in many different parts of the rest of the course. We say that a is *congruent* to b *modulo* m (written $a \equiv b \bmod m$) if and only if $m|(a-b)$, and we shall always suppose that m is a positive integer. a and b can be any integer but it is only important to which positive integer $< m$ it is congruent modulo m since, for any integer k, $km + a$ can be regarded as "the same number" as a.

We use the symbol $\equiv$ because congruence is very like equality, except it works with a finite set of elements (the *residues* $0, 1, \ldots, m-1$) rather than the integers. It is an equivalence relation like $=$ and algebra under it obeys most (but not all) of the same properties, the differences we shall discuss in detail. For example, in solving the equation $ax = b$ in the integers we know there is a (unique) solution if and only if $a|b$.

Exercise 7 *What are the solutions of* $3x \equiv 1 \bmod 4$, $2x \equiv 2 \bmod 4$ *and* $2x \equiv 1 \bmod 4$ *? (you can just test the values* $x = 0, \ldots, 3$ *in the equation).*

As you will have hopefully seen from the preceding exercise it is, at first sight, more difficult to tell whether an equation has solutions, and if so, how many. However, the following theorems will give us the ammunition to be able to find all the solutions to any linear congruence.

3.2 Solution of congruences

We first show that cancellation of common factors works both in the usual way, and in a more specialised way too.

Theorem 3.1 *If* $k \neq 0$ *then* $ka \equiv kb \bmod km$ *implies* $a \equiv b \bmod m$.

Proof: We have $k(a-b) \equiv 0 \bmod km$, and hence $km|k(a-b)$. Since $k \neq 0$, using the result which says if $c \neq 0$ then $ac|bc \Rightarrow a|b$, we know that $m|(a-b)$, as required. ◇

Theorem 3.2 *If* $(k, m) = 1$ *then* $ka \equiv kb \bmod m$ *if and only if* $a \equiv b \bmod m$.

Proof: Rearranging the left hand side of the theorem we get

$$k(a-b) \equiv 0 \bmod m,$$

which, by definition, is true if and only if $m|k(a-b)$. But, using lemma 4, $(k, m) = 1$ implies that $m|(a-b)$, in other words $a \equiv b \bmod m$. The reverse implication is true for all k, whether or not they are relatively prime to m, since multiplying both sides of an equation by a number cannot turn it from being true to false. ◇

Theorem 3.3 *If $d := \gcd(a, m) | b$ then $ax \equiv b \bmod m$ has d solutions, otherwise it has none.*

Proof: Suppose $x = x_0$ is a solution of $ax \equiv b \bmod m$ so that $m|(ax_0 - b)$, which implies that $ax_0 = b + km$ for some integer k. But since $d|a$ and $d|m$, $d|ax_0 - km = b$. Hence we have a contradiction if $d \not| b$ so that no such x_0 can exist, thus proving the second part of the theorem. If $d > 1$ we consider the equation $\frac{a}{d}x \equiv \frac{b}{d} \bmod \frac{m}{d}$ which is then in the same state as the original equation if $d = 1$ since $(\frac{a}{d}, \frac{m}{d}) = 1$ by lemma 1.3.

So, supposing that $d = 1$, we know there exist integers s and t so that $sa + tm = 1$ and we multiply this equation by b and get

$$a(sb) + m(tb) = b.$$

This clearly gives us a solution $x_0 = x \equiv sb \bmod m$ and we now have to verify it is the only one. Suppose there exists $y \not\equiv x \bmod m$ such that $ay \equiv b \bmod m$: then, taking the difference, $a(x-y) \equiv 0 \bmod m$ and by lemma 4 we have that $(x-y) \equiv 0 \bmod m$, contrary to our supposition. Thus there is exactly one solution if $d = 1$.

Finally, if $d > 1$ we can verify that $x \equiv x_0 + \frac{km}{d} \bmod m$ for $k = 0, 1, \ldots, d-1$ are the d solutions to the equation. ◇

Thus, using these two theorems we can now solve any linear congruence $ax \equiv b \bmod m$ as follows:

1. Determine whether there is any solution to $ax \equiv b \bmod m$ by finding $d := (a, m)$ and seeing whether it divides b.
2. If $d \geq 1$ then cancel d from all three terms in the congruence so you know there is now just one solution (under the new modulus).
3. Multiply both sides of the congruence by well chosen numbers which are relatively prime to m (to avoid introducing spurious solutions) until we have the equation in the form $x \equiv x' \bmod m'$.

Exercise 8 *Using the above method, solve (if possible)* $5x \equiv 9 \bmod 10$, $3x \equiv 5 \bmod 7$, $4x \equiv 1 \bmod 9$, $2x \equiv 37 \bmod 145$, $11x \equiv 17 \bmod 30$, $5x - 3 \equiv 8 \bmod 10$ *and* $15x \equiv 21 \bmod 42$.

3.3 Divisibility Testing

We are all familiar with the rules in base 10 for seeing whether a number is divisible by 2, 5 or 10 (testing whether the final digit is divisible by that number), and there is also an easy way to test for divisibility by 3 and 9:

Theorem 3.4 *If a number is written $n_k n_{k-1} \ldots n_2 n_1 n_0$ in base 10 then it is divisible by $d = 3$ or 9 if $\sum_{i=0}^{k} n_i \equiv 0 \bmod d$. It is divisible by 11 if $\sum_{i=0}^{k} (-1)^i n_i \equiv 0 \bmod 11$.*

Proof: Suppose we write our number in base d for one of the three divisors d.

$$10^k n_k + 10^{k-1} n_{k-1} + \ldots + 10 n_1 + n_0 \tag{3}$$

Since $10 \equiv 1 \bmod 3$ and $10 \equiv 1 \bmod 9$ we have that $10^i \equiv 1 \bmod 3$ and so equation (3) just becomes $\sum_{i=0}^{k} n_i$ in either modulus as required. Similarly $10 \equiv -1 \bmod 11$ and so we get the co-factors as

shown in the statement of the theorem. If and only if our number is divisible by the modulus then equation (3) is congruent to 0. ◇

As you will appreciate, in other bases the same rules as in base 10 won't apply, but there may be more, even. For instance, in base 12, we can immediately tell by the units digit whether our number is divisible by 2,3,4 6 or 12. For 8 and 9 we only need consider the last two digits of our number and similar rules to those in theorem 3.4 apply for divisibility by 5, 7, 11 and 13. Thus it would be easier to factorise numbers if we used base 12 to do all our arithmetic.

In fact, any of the work we have done and will do in this course is independent of the base in which it is done; we can apply the sieve of Eratosthenes or do number theoretic functions or cryptography in any base.

3.4 Linear Diophantine Equations

Diophantine equations are simply equations whose solutions are required to be integral. Such an equation is *linear* if all variables aren't raised to any power. An example of an LDE is $4x+5y=13$ which has the unique solution $x=2$, $y=1$ if we specify $x,y\in\mathbb{N}$. Otherwise we can also have $x=-3$, $y=5$, and an infinite number of others, all necessarily of the form $x=2+5k$, $y=1-4k$ for integer k. These problems usually arise in simulating real-life situations where, whole numbers are constrained to appear, such as numbers of pieces of fruit, people or stamps. It can be easily seen that equations with just two variables as above can be expressed in congruence terms, so we can derive $5y\equiv y\equiv 13\equiv 1 \bmod 4$ and do similarly for x and then find the positive solutions by trial substitution.

We now consider the situation of simultaneous congruences:

Theorem 3.5 (The Chinese Remainder Theorem)
If $m_1,\ldots,m_k$ are pair-wise relatively prime positive integers then for each set of a_i $(1\le i\le k)$ there is a unique x modulo $\prod m_i$ such that $x\equiv a_i \bmod m_i$.

Proof: If $k=2$ then x is a solution if and only if $x=a_1+m_1y\equiv a_2 \bmod m_2$. This has a unique solution since it can be rewritten $m_1y\equiv a_2-a_1 \bmod m_2$ and $(m_1,m_2)=1$, satisfying the conditions of lemma 3.3. If $y\equiv y_0 \bmod m_2$ is the solution then $y=y_0+zm_2$ and so $x=(a_1+m_1y_0)+zm_1m_2$, i.e. $x\equiv a_1+m_1y_0 \bmod m_1m_2$. The result for $k>2$ follows by repeating this process seeing as each time we produce a smaller set of equations which still has pair-wise relatively prime moduli. ◇

For example, to solve the following problem we proceed as follows:

A person goes into a post office with less than four dollars, wanting a set of identical stamps and no change. On being asked the teller replies, "3c stamps would mean 1c change, 8c stamps 5c but 17c stamps are fine!" How much money was involved ?

If we set the amount of money to be x cents, the set of congruences involved is $x\equiv 1 \bmod 3$, $x\equiv 5 \bmod 8$ and $x\equiv 0 \bmod 17$. The latter gives $x=17k$ which we substitute in the other two and get $2k\equiv 1 \bmod 3$ and $k\equiv 5 \bmod 8$. The second here gives $k=5+8l$ and substituting it into the first we get $10+16l\equiv 1 \bmod 3$ which gives $l=3m$. Back-substituting we get $k=5+24m$ and $x=85+408m$ and so $x=85$ cents.

Normally, the best way to attack such problems is the follows:

1. Make sure each equation is of the form given in the Chinese Remainder Theorem, $x \equiv a_i \bmod m_i$, and verify that all the m_is are relatively prime.

2. Take the *largest* two values of m_i and express the first as $x = a_i + km_i$, then substitute this expression for x into the second largest modulus congruence, getting an congruence in k which can be solved, giving an expression for k of the form $k = b + lm_j$, which can be substituted into the expression for x to give $x = c + lm_im_j$. This expression can then be substituted into the third largest congruence, and so on, until the final expression is equivalent to $x \equiv y \bmod m_1m_2 \dots m_k$.

Exercise 9 *Solve these two sets of congruences:*

$$\begin{array}{l} x \equiv 2 \bmod 7 \\ x \equiv 1 \bmod 5 \\ x \equiv 7 \bmod 8 \end{array} \qquad \textit{and} \qquad \begin{array}{r} 3x \equiv 1 \bmod 4 \\ 5x \equiv 4 \bmod 7 \\ x \equiv 2 \bmod 5 \\ 2x \equiv 0 \bmod 3 \end{array} .$$

3.5 Inverses in Congruences

In order to solve standard equations in congruence equations it would be nice to be able to just divide by the number which is the coefficient of the unknown. We cannot divide though, but we can sometimes multiply by a special number which will give us the number 1. For example, if we have $3x \equiv 4 \bmod 10$ then if we multiply by 7 then we have $21x \equiv 28 \bmod 10$ which can be simplified to give $x \equiv 8 \bmod 10$ which can be verified to be correct. As such, we can say that the inverse of 3 modulo 10 is 7, since $3 \times 7 \equiv 1 \bmod 10$. Note that this also implies that $7^{-1} = 3$ also, in a similar way to that in which $(\frac{2}{3})^{-1} = 1 \div \frac{2}{3} = \frac{3}{2}$. and $\frac{3}{2} \times \frac{2}{3} = 1$.

However, in congruences, such a number may or may not exist: For instance, we can see that $2x \equiv 3 \bmod 10$ can have no solution by theorem 3.3 since $\gcd(2, 10) = 2 \not| 3$. Thus 2^{-1} doesn't exist modulo 10, whereas modulo 11 it is 6.

In order to find inverses, it is simply a matter of solving congruence equations as we have done previously, using the equation

$$a^{-1}a \equiv 1 \bmod m$$

for the inverse of a modulo m

4 Simple Codes

We now use some of the mathematics we have developed so far in a practical application. If it is wished to keep a message secret it is necessary to store it in some special format that only you and the intended recipient can understand. The processes for doing this that we shall study are called *encryptions* and in encrypting we take the message we wish to send (the *plaintext*) and convert it to *ciphertext*.

The plaintext will be split into blocks, normally of one, two or three letters/characters (the message will normally be in English and so use the 26 character alphabet, but can also use numbers, spaces and other punctuation marks). The alphabets used for both texts, together with the encryption (and decryption) method, are called a *cryptosystem*.

4.1 Affine Cryptosystems

For instance, the simplest cryptosystem is the *shift* transformation in which the same alphabet is used for both plain and cipher texts and the transformation is just to convert the alphabet to a number (0 (for "A") to 25 (for "Z")) and add some constant to it, modulo 26. A table such as that in table 1 is often useful as a cross-reference aid.

Table 1: Numerical/Alphabetic Lookup Table

Letter	A	B	C	D	E	F	G	H	I	J	K	L	M
Value	0	1	2	3	4	5	6	7	8	9	10	11	12
Letter	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Value	13	14	15	16	17	18	19	20	21	22	23	24	25

The easiest constant to add is 13 since then the same function ($f(x) \equiv x + 13 \bmod 26$) can be used to both encrypt and decrypt messages; this is known as *rot13*.

Exercise 10 *Decode this word, given that it is in rot13:* `mvzonojr`

In general an *affine* transformation of a plaintext message $\mathcal{P}$ is defined by

$$\mathcal{C} \equiv a\mathcal{P} + b \bmod N,$$

where, in order for the encryption to be unique, we need $(a, N) = 1$. From our earlier work we can see that the decryption formula for $\mathcal{P}$ is $\mathcal{P} \equiv a^{-1}(\mathcal{C} - b) \bmod N$, where a^{-1} is the inverse of a modulo N

Exercise 11 *Generalising the idea of a reversible code, as in rot13, which values for a and b give us exactly the same equations for encoding and decoding ?*

If the only information that we know is that the cryptosystem is an affine transformation and we have a large sample of encrypted plaintext then to determine the parameters a and b (and hence the decryption formula) we study the sample texts and perform *frequency analysis* upon it, using tables compiled by other cryptographers.

For instance, in standard English the most common letters in people's vocabularies are usually "E" and then "T", with " " (the space character) being even more common if it is included in the alphabet. We then identify the most common characters in our ciphertext and, in order to try to identify the values of a and b, we assume that some pair of these is "E" and " ", in some order. With this assumption we should be able to find a and b, and we can then check that the other common letters are also sensible (so that "Q" is the third most common letter, for instance). Alternatively, we can start to decode our text, and if it doesn't make sense we know we have incorrect values for a and b. We then repeat the procedure, supposing a different pair of common letters in our ciphertext is "E" and " ", until either we decode the text, or are exhausted! In the latter case it is then sensible to look at the structure of the message, and use our knowledge of English to see that any one letter words are probably "A" or "I", common sequences are maybe "THE" or repeating letters are not "H" or "K", etc. As you can probably tell, frequency analysis is a bit of an art which can take some practice to become good at.

Exercise 12 *Take a page of a book or a magazine in another language (Shona, Ndebele, French) and see which are the 5 most common letters in that passage.*

Exercise 13 *Supposing we are using the 27 character alphabet (space included, with value 26) and from previous messages we see that the most common characters used are* `d` *and then* `b`*. Assuming the frequency distribution is standard, to what does* `ytbdwkydwqmbb` *decrypt ?*

4.2 Digraph and Matrix Cryptosystems

In order to make things a little harder for anyone wishing to decipher our codes against our wishes (even the affine transform has only about 700 different possibilities and so it is quite easy to find the possible plaintexts just by unleashing a computer upon it), we complicate our encoding process by taking larger subsets of our plaintext message. For instance, by adding an extra space (or other junk letter) at the end of the plaintext if necessary we can then split it into sets of two letters (*digraphs*). We encode by multiplying the first character's reference number by n (the number of letters in the alphabet) and adding the second number, we get a unique number between zero and $n^2 - 1$ for every different digraph. We then work modulo n^2 and hence there are a much larger number of possible values for the parameters of affine transformations.

For instance the word "CHELSEA!" would be split into four digraphs, "CH", "EL", "SE" and "A!", where, if we are using an n letter alphabet, the first digraph has value $2n + 7$, the second $4n + 11$, etc. Decoding a word encoded using digraphs is simply a matter of identifying which number has which quotient and remainder on division by n, as in theorem 1.1. For instance, if $n = 27$, the number 113 can be written as $27 \times 4 + 5$, so the two letters in the digraph have values 4 and 8, and so the digraph corresponding to 112 in the standard 26-letter alphabet is (by table 1) "EF".

Digraph encoding also requires a much larger database of previous transmissions since the frequency distribution of the digraphs will be much less pronounced than for single characters. However, it has been shown that in the 26 letter English alphabet the most common digraphs are usually "TH" and "HE", and with 27 characters they are "E ", "S " and " T", in that order.

Exercise 14 *If we are again using the 27 character alphabet and the two most common digraphs in the previous ciphertexts are* `fb` *and* `ab`*, try to determine the parameters of the affine transformations used to decode. You should find that there are 27 different possibilities for them. Rather than trying all these possibilities, you are given the extra information that "AC" codes to* `hh` *- to what does the message* `tnwf` *decode ?*

Assuming that you are all familiar with standard matrix operations; Digraphs can alternatively be represented as column vectors in an N dimensional space and we can then use matrix transformations to rearrange the points in this space as an encoding action. As with the real numbers it can be proved that the inverse of a matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ modulo N exists if $t :\equiv ad - bc \bmod N$ and $(t, N) = 1$ and is

$$t^{-1} \begin{pmatrix} d & -c \\ -b & a \end{pmatrix} = \begin{pmatrix} t^{-1}d & -t^{-1}c \\ -t^{-1}b & t^{-1}a \end{pmatrix}.$$

Exercise 15 *What is the inverse of* $\begin{pmatrix} 5 & 2 \\ 1 & 1 \end{pmatrix}$ *modulo 7 ? Check your answer.*

We represent a message consisting of k digraphs as a $2 \times k$ matrix in which the k^{th} column represents the k^{th} digraph. This can then be multiplied by the transforming matrix to give us another $2 \times k$ matrix which represents the coded message.

For example, the message "buy now!" can be represented in a 28 character alphabet and, split into digraphs, gives the 2×4 matrix $\begin{pmatrix} 1 & 24 & 13 & 22 \\ 20 & 26 & 14 & 27 \end{pmatrix}$, which, when multiplied by $\begin{pmatrix} 5 & 1 \\ 1 & 6 \end{pmatrix}$ we get $y := \begin{pmatrix} 25 & 6 & 23 & 25 \\ 9 & 12 & 13 & 16 \end{pmatrix}$, which is equivalent to the message `zjgmxnzq`. Obviously, to decode this, we just multiply y by $A^{-1} = 1^{-1} \begin{pmatrix} 6 & -1 \\ -1 & 5 \end{pmatrix}$ (since $t = 6 \times 5 - 1 \times 1 = 29 \equiv 1 \bmod 28$) and reconvert to our alphabet.

Exercise 16 *Code a six-letter word using the method above, with the transformation matrix below and the 26 letter alphabet. Swap your ciphertext with your neighbour and try to decode correctly before them!*

$$\begin{pmatrix} 14 & 3 \\ 7 & 11 \end{pmatrix}$$

4.3 Further Extensions

We have now seen the major types of simple codes we can use, but there are many improvements to the techniques in order to complicate matters even further, but at the expense of making the encoding more difficult.

- Shift Words: We can, in addition to using an affine transformation, also add a code word, repeated over and over, so that adjacent letters are not shifted using the same formula. The original cryptosystem based upon this idea is the wordshift transform, where one simply chooses a word, "dog", for example, then takes the value of each of the letters in the word (3, 14 and 6 in our case) and forms the following sum:

	m	y	d	o	g	i	s	b	l	u	e
	12	24	3	14	6	8	18	1	11	20	4
+	d	o	g	d	o	g	d	o	g	d	o
	3	14	6	3	14	6	3	14	6	3	14
=	p	m	j	r	u	o	v	p	r	x	s
	15	12	9	17	20	14	21	15	17	23	18

  This example involved just the shift transformation, but one can, with just a little bit more work, use the affine, or even digraph transformations, using the same rule

$$\mathcal{C} \equiv a\mathcal{P} + \mathcal{S} \bmod N. \tag{4}$$

  In order to decode such equations, if we know the values of a and $\mathcal{S}$, we simplify solve equation 4 to get $\mathcal{P} \equiv a^{-1}(\mathcal{C} - \mathcal{S}) \bmod N$. However, if the parameters are unknown, we have a harder task. We have to first guess the length l of the shift word, and then break our ciphertext up into l groups of letters, group G_i consisting of the letters in positions i, $l + i$, $2l + i$,.... We then perform frequency analysis separately on each of these groups, in order to find what the i^{th} letter of the shift word was. Since we are unable to check whether words are formed

without solving for all l groups we have to use the technique of checking whether the most common letters in any particular G_i make sense, before trying to put all l solutions together. Obviously there is a lot more work involved, and a larger margin for error in cracking the code, especially given that, if the length of the shift word is guessed incorrectly, the process can continue indefinitely.

- Tri- and Quad- graphs: There is no special reason to just consider digraphs, we can just as easily break our text up into groups of 3 or 4 letters, and then find the value of these trigraphs in just the same way. This way frequency analysis becomes almost useless since now very few groups of letters will appear more than once, even in a whole document. However, the trade-off is that enciphering the code will require a lot of work, since we will now be working with moduli in the ranges of 10 000 to 500 000! To even find the inverse of some number working in such a modulus is time-consuming.

5 Number-Theoretic Functions

We now study several functions which are defined just for the integers and see how they are used in many results. We say that a function is *multiplicative* if it is true that $f(mn) = f(m)f(n)$ if $(m, n) = 1$.

5.1 The divisors of an integer

We define $\tau(n)$ as the number of positive divisors of an integer n (including 1 and n) and $\sigma(n)$ as the sum of these same positive divisors. Thus, for example, the set of divisors for $n = 10$ is $\{1, 2, 5, 10\}$ and so $\tau(10) = 4$ and $\sigma(10) = 1 + 2 + 5 + 10 = 18$. Mathematically we usually write

$$\tau(n) := \sum_{d|n} 1 \quad \text{and } \sigma(n) := \sum_{d|n} d.$$

We can first notice that $\tau(p) = 2$ and $\sigma(p) = p + 1$ for any prime p, since, by definition, p has no other divisors apart from itself and 1.

Exercise 17 *What are $\tau(n)$ and $\sigma(n)$ for $n = pq$ (p and q both primes), $n = p^2$, $n = p^k$?*

Using the final result of exercise 17 we can now actually calculate the values of $\tau(n)$ and $\sigma(n)$ for any n as shown in the following theorems by proving both functions to be multiplicative.

Theorem 5.1 *If $n = \prod_{i=1}^{\infty} {p_i}^{\alpha_i}$ then*

$$\tau(n) = \prod_{i=1}^{\infty} \tau({p_i}^{\alpha_i}) = \prod_{i=1}^{\infty} (\alpha_i + 1)$$

and

$$\sigma(n) = \prod_{i=1}^{\infty} \sigma({p_i}^{\alpha_i}) = \prod_{i=1}^{\infty} (1 + p_i + \cdots + {p_i}^{\alpha_i}) = \prod_{i=1}^{\infty} \frac{{p_i}^{\alpha_i+1} - 1}{p_i - 1}$$

Proof: We consider the numbers of the form $\prod_{i=1}^{\infty} {p_i}^{\beta_i}$ where $0 \leq \beta_i \leq \alpha_i$ and aim to prove that these are *exactly* the divisors of n. Any of these numbers *is* a divisor, since, when multiplied by $\prod_{i=1}^{\infty} {p_i}^{\alpha_i - \beta_i}$ (which is an integer too), we get n.

By theorem 2.1 we have that every number is representable in prime power form and so we need only consider as possible divisors those integers with $\beta_i > \alpha_i$. But any integer with a β_i satisfying that cannot be a divisor since then it has more p_is in its decomposition than n.

Thus we just need to count and to add the numbers of the given form to find $\tau(n)$ and $\sigma(n)$ respectively. The formulae given are surely true for n a power of a prime, so we use induction on the number of prime factors in n. Consider the case when n has $k+1$ different prime divisors and suppose the induction hypothesis is true for less than this number, Choose one prime p which is to the power γ in n: the divisors of n can be split into $\gamma + 1$ groups, according to what power of p (from 0 to γ) divides them. Each of these groups has $\tau(\frac{n}{p^\gamma})$ members by the induction hypothesis and thus we can see that $\tau(n) = (1+\gamma)\tau(\frac{n}{p^\gamma})$ and hence that part of our result is proven.

Similarly, each group has sum $p^\beta \sigma(\frac{n}{p^\gamma})$, for some integer β, and so the total sum is $(1 + p + \cdots + p^\gamma)\sigma(\frac{n}{p^\gamma}) = \sigma(p^\gamma)\sigma(\frac{n}{p^\gamma})$. ◇

5.2 Perfect Numbers

We say that a number n is *perfect* if $\sigma(n) = 2n$, or, equivalently, if the sum of its proper divisors (those other than n) is n itself. For instance, $n = 6$ satisfies this relation ($6 = 1 + 2 + 3$) and is the smallest perfect number, and it is clear that since $2p > p + 1$ for all primes p no perfect number can be prime.

Exercise 18 *Find the next smallest perfect number by trial and error and then, by comparing the prime factorisations of these numbers predict the next, which is less than 500.*

Theorem 5.2 (Euclid) *If $n = 2^k - 1$ is prime, then $m := 2^{k-1}(2^k - 1)$ is perfect.*

Proof: Since $n = 2^k - 1$ is prime we have, by theorem 5.1 and since n and 2^{k-1} are relatively prime, ◇

$$\sigma(m) = \sigma(2^{k-1})\sigma(2^k - 1) = (2^{k-1+1} - 1)(1 + (2^k - 1)) = (2^k - 1)(2 \times 2^{k-1}) = 2m.$$

Exercise 19 *What are the divisors of 198 ? What, therefore, are $\tau(198)$ and $\sigma(198)$.*

Listing the divisors of a number n in increasing order we notice that they "pair up" from either end, each pair multiplying to n. However, there is one exception in which there is an odd number of divisors:

Lemma 9 *$\tau(n)$ is odd if and only if n is a square number.*

Proof: We give two proofs, to show that there is more than one way to skin a cat:

1. By the observation above we note that the number which won't pair up will be $\sqrt{n}$ since if $d|n$ then also $\frac{n}{d}|n$, and these two numbers are distinct if $\frac{n}{d} \neq d$, i.e. $n \neq d^2$.

2. By theorem 5.1 we see that $\tau(n) = \prod_{i=1}^{\infty}(1+\alpha_i)$ and so $\tau(n)$ is odd if and only α_i is even for all i. This is equivalent to n being a square.

$\diamond$

Exercise 20 *For which n is $\sigma(n)$ odd ?*

5.3 Euler's totient function

We now consider a slightly more complex number theoretic function: we define $\phi(n)$ as the number of integers less than n which are relatively prime to n. We note for primes p that $\phi(p) = p-1$ since every integer $n < p$ satisfies $(n,p) = 1$.

Exercise 21 *By considering cases find a formula for $\phi(p^2)$. How about $\phi(p^3)$?*
[Hint: it may be easier to consider those numbers which **aren't** *relatively prime to a number. This is $n - \phi(n)$.]*

Theorem 5.3 $\phi(p^l) = p^{l-1}(p-1)$.

Proof: Those integers which aren't relatively prime to p^l, and hence p, are numbers of the form kp, where $1 \leq k \leq p^{l-1}$. There are therefore p^{l-1} of these and so $\phi(p^l) = p^l - p^{l-1}$. $\diamond$

We, of course, now wish to find a formula for $\phi(n)$ in general and, it turns out that ϕ is a multiplicative function as we would have hoped:
Suppose $(m,n) = 1$. The number $\phi(m)\phi(n)$ is the number of pairs of integers (i,j) such that $(i,m) = 1$ and $(j,n) = 1$. By theorem 3.5 (the Chinese remainder theorem) we know that there is a unique solution x modulo mn to $x \equiv i \bmod m$ and $x \equiv j \bmod n$ for each pair of i and j and so it remains to show that this x is relatively prime to mn and every such residue modulo mn can be generated in this way so that $\phi(mn) = \phi(m)\phi(n)$ as required.

"$\Rightarrow$" Suppose $(i,m) = (j,n) = 1$ but $p|(x,mn)$ so that $p|m$ or $p|n$ by lemma 6. If $p|m$ then $p|x$ also and so, since $i = km - x$ for some integer k, $p|i$, contravening $(i,m) = 1$. But then $p|n$ and we can deduce $p|ln - x = j$, another contradiction. Thus $(x,mn) = 1$.

"$\Leftarrow$" Suppose $(x,mn) = 1$. If there exists a $p|(i,m)$ then, as before, $p|km - i = x$ and $p|mn$ implying $p|1$, a contradiction, and similarly we can prove $(j,n) = 1$.

Thus there is a one-to-one correspondence between the residues x and the pairs (i,j) and so $\phi(mn) = \phi(m)\phi(n)$.

We can thus compute the value of $\phi(n)$ for any integer n by finding the prime factorisation of n and multiplying the invidual terms, or we can also see that

$$\phi(n) = n \times \prod_{i=1}^{\infty}\left(\frac{p_i - 1}{p_i}\right).$$

6 Higher Power Congruences

6.1 More complicated congruences

Since $\equiv$ is working just like $=$ we define powers of the residues in exactly the same way - $a^0 = 1$ and $a^n = aa^{n-1}$ if $n \geq 1$. The various familiar rules still work ($a^m a^n = a^{m+n}$ and $(a^m)^n = a^{mn}$) and so we shall explore this relation for small integers to see how it works. Firstly, in order to find $a^n \bmod m$ for some large n it is not necessary to first evaluate the large integer a^n and then calculate its remainder on division by m - we just repeatedly square or cube, recording our intermediate results and then multiply these values together. For instance, to solve $x \equiv 3^5 \bmod 7$ we first note that $3^2 \equiv 2 \bmod 7$ and so $3^4 \equiv 2^2 = 4 \bmod 7$. Finally, $3^5 = 3^4 3 \equiv 4 \times 3 \equiv 5 \bmod 7$.

Exercise 22 *Find the values of* $2^7 \bmod 5$ *and* $6^3 \bmod 11$.

Table 2: Powers of the residues modulo 5

a	a^2	a^3	a^4	a^5
0	0	0	0	0
1	1	1	1	1
2	4	3	1	2
3	4	2	1	3
4	1	4	1	4

We can make tables of powers too as in table 2. What do you notice about it ?
Construct a table of powers for moduli 3 and 7. What about 4 or other composite numbers ?

What you may have noticed is the result known as Fermat's Theorem. This isn't the famous "last" one, the one that has only just been proved, but this is much more useful anyway and it goes as follows: If p is prime and $(a, p) = 1$ then

$$a^{p-1} \equiv 1 \bmod p. \tag{5}$$

We can remove the gcd condition and make it true for all a if we restate it, as: if p is prime, then

$$a^p \equiv a \bmod p$$

since if $(a, p) = 1$ it is valid to multiply both sides of the congruence by a and if $(a, p) = p$ (so $a = kp$ for some integer k) then $0 \equiv 0 \bmod p$.

Lemma 10 *If* $(a, m) = 1$ *then the least residues of*

$$L_1 := a, 2a, 3a, \ldots, (m-1)a \bmod m$$

are

$$L_2 := 1, 2, 3, \ldots, (m-1) \bmod m$$

in some order.

Proof: No number in L_1 is congruent to 0 mod m since $(a, m) = 1$. Thus we just need to show that no residue appears twice in L_1: suppose this is the case and we have $ra \equiv sa \bmod m$. But then we can again cancel a from both sides to get $r \equiv s \bmod m$, a contradiction of the choice of r and s as different residues. ◇

Theorem 6.1 (Fermat's Theorem) *If $(a, p) = 1$ and p is prime then $a^{p-1} \equiv 1 \bmod p$.*

Proof: We consider L_1 from lemma 10 again. We showed it was the rearrangement of L_2 so we consider both of the products when all the terms in each list are multiplied together. Because they are intrinsically the same set we must have that

$$a \times 2a \times 3a \times \cdots \times (p-1)a \equiv 1 \times 2 \times 3 \cdots \times (p-1) \bmod p.$$

Cancelling the common factors of 2, 3, $\ldots, p-1$, all of which are relatively prime to p (since p is prime), we get the required result. ◇

We can use a similar technique to prove another famous result in congruence arithmetic, but this time involving factorials. Investigate them yourself if you like, but for our proof we shall first need a lemma.

Lemma 11 *If p is an odd prime then $x^2 \equiv 1 \bmod p$ has exactly two solutions.*

Proof: Let r be any solution. Thus $p|(r^2 - 1) = (r-1)(r+1)$. By lemma 6 we have from this that $p|(r-1)$ or $p|(r+1)$, which can be stated as $r \equiv 1 \bmod p$ or $r \equiv p-1 \bmod p$, and both of these can be easily seen to be the solutions $r \equiv \pm 1 \bmod p$. We note that if $p \neq 2$ then these two solutions are different, since $1 < p-1$ in this case. ◇

Exercise 23 *Find a composite number n which doesn't have exactly two solutions to $x^2 \equiv 1 \bmod n$.*

We now consider the notion of inverses of the residues, restricting ourselves to prime modulus - the inverse is the residue a^{-1} so that $aa^{-1} \equiv 1 \bmod p$. If $a \equiv 0 \bmod p$ then there is obviously no solution to $aa^{-1} \equiv 1 \bmod p$. From lemma 11, we see that if and only if $a \equiv 1$ or $p-1 \bmod p$ then a is its own inverse. Otherwise we shall prove that the remaining residues "pair up" in that all the inverses are distinct and the inverse of the inverse of a is a itself. First suppose $a^{-1} \equiv b' \bmod p$: then $b \equiv aa^{-1}b \equiv ab'b \equiv a \bmod p$ and so the inverses are distinct and also every residue has a unique inverse.

Theorem 6.2 (Wilson's Theorem) *If p is prime then $(p-1)! \equiv p-1 \equiv -1 \bmod p$.*

Proof: From the above observations we see that when we take the product of $1, 2, 3, \ldots$ and $(p-1) \bmod m$ the first and last terms we leave alone but the remaining $p-3$ terms consist of $\frac{p-3}{2}$ pairs of residues and their inverses, which when multiplied obviously give 1. Thus $(p-1)! = 1 \times 2 \times 3 \times \cdots \times (p-1) \equiv p-1 \equiv -1 \bmod m$. ◇

Referring back to equation 5 we now extend it to be valid for all the integers. Noting that in that equation the exponent was $p - 1 = \phi(p)$ and after experimenting with the non-prime moduli one would notice that the smallest number for which $a^k \equiv 1 \bmod m$ is $k = \phi(m)$. This we now prove.

Theorem 6.3 *If* $(a, m) = 1$ *then* $a^{\phi(m)} \equiv 1 \bmod m$

Proof: By lemma 3.3 we know that $ax \equiv 1 \bmod m$ has a solution if and only if $(a, m) = 1$, and thus the number of residues modulo m which have inverses is the same as the number which are relatively prime, $\phi(m)$. We call the set of inverses S. For some a such that $(a, m) = 1$ we consider the set of residues $S_a := \{aa_i\}$, where each a_i is an invertible residue modulo m. No element of S_a is repeated since $aa_i \equiv aa_j \bmod m \iff a_i \equiv a_j \bmod m$ and all elements of S_a are in S since $(aa_i)^{-1} = (a_i)^{-1}a^{-1}$. Taking the product of both S and S_a we get the same result and since they each have $\phi(m)$ elements we have

$$\begin{aligned} \textstyle\prod_{i=1}^{n} aa_i - \prod_{i=1}^{n} a_i &\equiv 0 \bmod m \\ (a^{\phi(m)} - 1) \textstyle\prod_{i=1}^{n} a_i &\equiv 0 \bmod m. \end{aligned} \tag{6}$$

Since each a_i is invertible we can multiply equation (6) by each of their inverses and so we get $a^{\phi(m)} \equiv 1 \bmod m$ as required. ◇

Corollary 1 *If* $(a, m) = 1$ *and* $k \equiv n \bmod \phi(m)$ *then* $a^n \equiv a^k \bmod m$.

Proof: By theorem 6.3 we have that, for any integer j,

$$(a^{\phi(m)})^j \equiv 1^j \equiv 1 \bmod m$$

and so, since $n = k + j\phi(m)$ we have the stated result. ◇

Exercise 24 *Solve* $x \equiv 2^{777} \bmod 75$ *using the corollary.*

Exercise 25 *Evaluate* $\sum_{d|n} \phi(d)$ *for* $n = 3, 6, 7, 18, 21$ *and* $n = p$, *a prime.*

Theorem 6.4 *If* $n \geq 1$ *then* $\sum_{d|n} \phi(d) = n$.

Proof: We shall use a method first devised by Gauss. Group the integers $1, \ldots, n$ depending upon their gcd with n, thus for $n = 12$ we get the following:

$$\begin{aligned} C_1 &= \{1, 5, 7, 11\}, & C_4 &= \{4, 8\}, \\ C_2 &= \{2, 10\}, & C_6 &= \{6\}, \\ C_3 &= \{3, 9\}, & C_{12} &= \{12\}. \end{aligned}$$

Clearly all n integers are in exactly one of these classes. But $(m, n) = d$ if and only if $(m/d, n/d) = 1$ and so $m \in C_d \iff (m/d, n/d) = 1$. Thus the number of elements in C_d is $\phi(n/d)$. Hence the total number of elements in all the classes is $n = \sum_{d|n} \phi(n/d)$. But $\sum_{d|n} \phi(n/d)$ is the same number as $\sum_{d|n} \phi(d)$ since, for each $d|n$ there exists an integer $e := n/d|n$ also, and hence all divisors are counted in both sums. Hence $n = \sum_{d|n} \phi(d)$ as required. ◇

6.2 Completing the Square

We now study how to solve more complicated algebraic congruence equations - for instance, we have seen in lemma 11 that, modulo p, that quadratic equation has two solutions. But how about $x^2 \equiv 2 \bmod 5$ and $x^2 \equiv 2 \bmod 7$?

For small values of the modulus we can just try all the different values of x and see which, if any, work. However, this isn't practical for larger moduli, so we need a new scheme:

We initially start with just prime moduli, and work up to general solution: Suppose that the equation we are considering is

$$ax^2 + bx + c \equiv 0 \bmod m. \tag{7}$$

$m = 2$: 2 is a prime, and, although not odd, it certainly behaves in an odd way! To solve any congruence modulo 2 it suffices to substitute the values zero and one in the equation, or note that no units term implies that 0 is a solution and an even number of terms implies that 1 is. Thus $x^2 + x + 1 \equiv 0 \bmod 2$, for instance, has no solution.

$m = p$ an odd prime: Multiplying equation (7) by $4a$ (which is not congruent to zero modulo p, since neither 2 nor a was) we get

$$(2ax)^2 + 4abx + 4ac \equiv (2ax + b)^2 - b^2 + 4ac \equiv 0 \bmod p.$$

This equation has a solution if and only if $z^2 \equiv b^2 - 4ac \bmod p$ has a solution, i.e. $d := b^2 - 4ac$ is a square number. Thus our situation is exactly analogous to solution of real quadratic equations in that we need to verify whether d exists and then the roots of the equation are

$$\frac{-b \pm \sqrt{b^2 - 4ac}}{2a.}$$

Exercise 26 *What are the solutions of* $2x^2 - 2x + 3 \equiv 0 \bmod 7$ *?*

all other m: We break down m into its prime factors and then solve an equation for each of these until we have a solution, in much the same way as the Chinese remainder theorem. We shall show how to do this with an example, firstly with $m = p^\alpha$:

$$x^2 + 2x + 6 \equiv 0 \bmod 49$$

We know that if this holds then certainly $x^2 + 2x + 6 \equiv 0 \bmod 7$ and so we first find the solutions of this:

$$\begin{aligned} & x^2 + 2x + 6 \equiv 0 \bmod 7 \\ \Rightarrow \quad & (x+1)^2 + 5 \equiv 0 \bmod 7 \\ \Rightarrow \quad & (x+1)^2 \equiv 2 \bmod 7 \\ \Rightarrow \quad & x + 1 \equiv 3, 4 \bmod 7 \\ \Rightarrow \quad & x \equiv 2, 3 \bmod 7. \end{aligned}$$

We can verify that this is correct by verifying that $4 + 4 + 6 = 14$ and $9 + 6 + 6 = 21$, both divisible by 7. Thus we can write $x = 2 + 7y$ or $3 + 7y$ and substitute into the original equation to get, respectively,

$$(2 + 7y)^2 + 2(2 + 7y) + 6 \equiv 49y^2 + 42y + 14 \equiv 0 \bmod 49$$

and

$$(3+7y)^2 + 2(3+7y) + 6 \equiv 49y^2 + 56y + 21 \equiv 0 \bmod 49.$$

These can both be simplified to linear congruences $6y + 2 \equiv 0 \bmod 7$ and $y + 3 \equiv 0 \bmod 7$ (since we have an equation of the form $7a \equiv 7b \bmod 7c$, and we know we can cancel the 7s by lemma 3.1 (this will, amazingly, always be the case with powers of primes). Thus we get the solutions $y \equiv 2 \bmod 7$ and $y \equiv 4 \bmod 7$, which can then be resubstituted to give the required values for $x \equiv 16, 31 \bmod 49$.

If m is not a power of a prime then we just have to break the working down into the prime powers which constitute m and then put them together using the Chinese remainder theorem as below:

If the equation is $x^2 + 2x + 6 \equiv 0 \bmod 147$ then we know that we must have $x^2 + 2x + 6 \equiv 0 \bmod 49$ (which has the solution worked out above) as well as $x^2 + 2x + 6 \equiv 0 \bmod 3$, which can be easily seen to have solutions $x \equiv 0, 1 \bmod 3$, and so the solution to the whole is $x \equiv 16, 31, 114, 129 \bmod 147$.

Exercise 27 *Find the solutions of* $7x^2 - 13x + 31 \equiv 0 \bmod 75$.

6.3 Indices and Primitive Roots

The method outlined above will work for any quadratic, cubic,... and so the only problem left for us to solve is whether any particular integer is a root of $x^k \equiv a \bmod p$ for any prime p. To do this we shall use a counterpart to logarithms in modular arithmetic: The question to ask when forming a table of $\operatorname{ind}_k n$ is "What power of k is equal to n?"

Table 3: Power and Index table for 11

n	0	1	2	3	4	5	6	7	8	9	10
2^n	1	2	4	8	5	10	9	7	3	6	1
3^n	1	3	9	5	4	1	3	9	5	4	1
$\operatorname{ind}_2 n$		0	1	8	2	4	9	7	3	6	5

Primitive Roots

As we can see in table 3, if $a \not\equiv 0 \bmod 11$ then there exists some residue r modulo 10 so that $a \equiv 2^r \bmod 11$. We say that the *index (base b)* $\operatorname{ind}_b a$ is the residue r modulo $p-1$ such that $a \equiv b^r \bmod p$. We note, also from table 3, that 3 cannot be used as an index as, for instance, $7 \not\equiv 3^k \bmod 11$ for any integer k. If b is such that all invertible residues have representations of the form b^k then we say that b is a *primitive root*. The *order* of b modulo p is the smallest positive integer k such that $b^k \equiv 1 \bmod m$. Thus b is primitive if it has order $\phi(m)$.

Lemma 12 *If* $(a, m) = 1$ *and* a *has order* t *mod* m *then* $t | \phi(m)$.

Proof: By theorem 1.1 we can write $\phi(m) = qt + r$, with $0 \leq r < t$ and lemma 6.3 tells us that

$$1 \equiv a^{\phi(m)} \equiv a^{qt}a^r \equiv a^r \bmod m.$$

But since t is the smallest positive integer satisfying $a^t \equiv 1 \bmod m$ we must have $r = 0$, which implies our result. ◇

Corollary 2 *If p and q are odd primes and $q|(a^p - 1)$ then either $q|a - 1$ or $q = 2kp + 1$ for some integer k. In particular, any divisor of $2^p - 1$ is of the form $2kp + 1$.*

Proof: We have $a^p \equiv 1 \bmod q$ and so the order of a divides p and so is either 1 or p. In the former case we have $a^1 \equiv 1 \bmod q$ and so $q|a-1$, and in the latter $p|\phi(q) = q-1$ so that $q = 1+pr$. Since q and p are odd r must be even and so $q = 2kp + 1$. If $a = 2$ then the former case cannot occur since $q > 1$. ◇

Exercise 28 *Factorise $4^7 - 1$. Prove that $2^{13} - 1$ is prime.*

We know that $b^{\phi(m)} \equiv 1 \bmod m$ by theorem 5, so in order to check whether b is a primitive root we simply need, by lemma 12, to evaluate b^d for all divisors d of $\phi(m)$: if $b^d \not\equiv 1 \bmod m$ for every $d < \phi(m) - 1$ then b must be primitive. If $(a, m) \neq 1$ then a cannot be a primitive root since no positive power of a can be congruent to 1 modulo m.

Exercise 29 *What are the primitive roots of 8,10,17 ?*
[Hint: use the repeated squaring method, reducing modulo 17 each time, and use any repeating patterns you notice]

It can be proved that the only numbers which have primitive roots are those of the form p^k or $2p^k$, where p is an odd prime, in addition to 2 and 4. Also, p has $\phi(p - 1)$ primitive roots, but the actual distribution of them is hard to predict; For most primes, one or both of 2 and 3 are primitive roots and so one need only check a couple of powers in practice. But for 71, 7 is the smallest primitive root, for example.

Practical Application

To actually use the tables of indices such as table 3 we proceed as for logarithms as follows:

1. to solve $4x + 3 \equiv 2 \bmod 11$ we subtract 3 from both sides and take indices, so that from $4x \equiv -1 \bmod 11$ we get $\text{ind } 4 + \text{ ind } x \equiv \text{ ind } 10 \bmod 10$, i.e. $\text{ind } x \equiv 5 - 2 \equiv 3 \bmod 10$, and so $x \equiv 8 \bmod 11$. We can check that indeed $4 \times 8 + 3 = 35 \equiv 2 \bmod 11$.

2. to solve $x^2 + 3x \equiv 9 \bmod 11$ we complete the square as before, getting $(x - 4)^2 \equiv 3 \bmod 11$, and then take indices of both sides so we have $2 \text{ ind } (x - 4) \equiv 2w \equiv \text{ ind } 3 \equiv 8 \bmod 10$, We solve this to get $w \equiv 4 \bmod 5 \equiv 4, 9 \bmod 10$ which corresponds to $x - 4 \equiv 5, 6 \bmod 11$ and so $x \equiv 9, 10 \bmod 11$.

This method takes the guesswork out of solving congruences, but the work involved is usually greater unless one has a lot of congruences in one modulus to work out. Once you are practiced at solving by the method described earlier this way becomes only useful for solving equations of the form $5^x \equiv 9 \bmod 11$ or $x^5 \equiv 3 \bmod 11$. The former translates to solving $4x \equiv 6 \bmod 10$, which has solution $x \equiv 4 \bmod 5$.

7 Quadratic Residues

We still have the problem of determining whether or not there exists a solution to a congruence of the form

$$x^2 \equiv a \bmod m. \tag{8}$$

7.1 Euler's Criterion

If $(a, m) = 1$ and $m = p$, an odd prime, then one way to determine whether (8) has a solution is by using *Euler's criterion* which is the following:

$$a^{\frac{p-1}{2}} \equiv \begin{cases} 1 & \text{if there is a solution} \\ -1 & \text{if there is no solution} \end{cases} \bmod p$$

This result we shall verify in the problem classes, by using primitive roots. But we now move on and do a worked example of the criterion: does $x^2 \equiv 8 \bmod 23$ have a solution ?
we calculate $8^{11} \bmod 23$, and obviously use the repeated squaring method. $8^2 = 64 \equiv -5$, $8^4 \equiv (-5)^2 \equiv 2$, $8^8 \equiv 4$ and so $8^{11} \equiv 8 \times (-5) \times 4 \equiv 8 \times 3 \equiv 1$. Thus $8^{11} \equiv 1 \bmod 23$ and so there is a solution. But note that we get no clues to what the solution is, despite the work done. One way to do this, slightly more sophisticated than just trying all possible answers in the equation, is to proceed like this:

$$8 = 2^2 \times 2 \equiv 2^2 \times 25 = 2^2 \times 5^2 \equiv 10^2.$$

So the solution here came out easily to be ± 10, so $x \equiv 10, 13 \bmod 23$. However it is not always this easy:

Exercise 30 *Show that* $x^2 \equiv 41 \bmod 61$ *has a solution using Euler's criterion and then find it using the method demonstrated above.*

7.2 The Legendre Symbol

We now incorporate some new definitions which will enable us to determine whether or not equation 8 has a solution in the case $m = p$. We say that a is a *quadratic residue* if there is a solution and a *quadratic non-residue* if not. We define the Legendre symbol as follows:

$$\left(\frac{a}{p}\right) := \begin{cases} 1 & \text{if } a \text{ is a quadratic residue of } p \\ -1 & \text{if } a \text{ is a quadratic non residue of } p \end{cases}$$

Note that the bottom number in the Legendre symbol has to be an odd prime and not a divisor of the top number so that neither $\left(\frac{2}{9}\right)$ nor $\left(\frac{6}{3}\right)$ is defined. From before we see that $\left(\frac{8}{23}\right) = 1$ and $\left(\frac{41}{61}\right) = 1$ and note that this also implies that, say, $\left(\frac{31}{23}\right) = 1$ too, since $x^2 \equiv 8 \equiv 31 \bmod 23$ has a solution. This observation is the basis for one of the three basic rules for evaluating Legendre symbols.

Theorem 7.1 *The Legendre symbol has the following properties:*

1. *If* $a \equiv b \bmod p$ *then* $\left(\frac{a}{p}\right) = \left(\frac{b}{p}\right)$.

2. *If* $p \not| a$ *then* $\left(\frac{a^2}{p}\right) = 1$.

3. *If* $p \not| a$ *and* $p \not| b$ *then* $\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right)$.

Proof:

1. The solution (if it exists) of $x^2 \equiv a \bmod p$ is necessarily the same as that of $x^2 \equiv b \bmod p$ and so $\left(\frac{a}{p}\right) = \left(\frac{b}{p}\right)$.

2. There always exists a solution of $x^2 \equiv a^2 \bmod p$, namely $x \equiv a, -a \bmod p$.

3. We use Euler's criterion to prove this:
since we note that
$$\left(\frac{a}{p}\right) \equiv a^{\frac{p-1}{2}} \bmod p \tag{9}$$
we have
$$\left(\frac{ab}{p}\right) \equiv ab^{\frac{p-1}{2}} \equiv a^{\frac{p-1}{2}} b^{\frac{p-1}{2}} \equiv \left(\frac{a}{p}\right)\left(\frac{b}{p}\right) \bmod p.$$
But since both sides of this congruence can be only 1 or -1 we can replace the $\equiv$ by $=$ and thus we have the result.

$\diamond$

Exercise 31 *Find* $\left(\frac{19}{5}\right)$ *and* $\left(\frac{-9}{13}\right)$ *using theorem 7.1*

The final rules we need in order to work with Legendre symbols are the following:

Theorem 7.2 (The Quadratic Reciprocity Law) *If* p *and* q *are odd primes and* $p \equiv q \equiv 3 \bmod 4$ *then* $\left(\frac{p}{q}\right) = -\left(\frac{q}{p}\right)$, *otherwise* $\left(\frac{p}{q}\right) = \left(\frac{q}{p}\right)$.

There are many proofs of these but most are quite complicated. If you are interested, please consult any number theory textbook.

Lemma 13

$$\left(\frac{1}{p}\right) = 1, \left(\frac{-1}{p}\right) = \begin{cases} 1 & \text{if } p \equiv 1 \bmod 4 \\ -1 & \text{if } p \equiv 3 \bmod 4 \end{cases} \quad \textit{and} \quad \left(\frac{2}{p}\right) = \begin{cases} 1 & \text{if } p \equiv 1, 7 \bmod 8 \\ -1 & \text{if } p \equiv 3, 5 \bmod 8 \end{cases}$$

Proof:

- We have seen in lemma 11 that $x^2 \equiv 1 \bmod p$ always has the solutions $x \equiv 1, p-1 \bmod p$ so the first statement is true.

- Using Euler's criterion and by equation (9) we have $\left(\frac{-1}{p}\right) \equiv (-1)^{\frac{p-1}{2}} \bmod p$, which is easily seen to be equal to 1 or -1 depending upon the value of p as stated.

- Finally, we consider the sequence of even residues modulo p, $\{2, 4, \ldots, p-3, p-1\} = \{2, 4, \ldots -3, -1\}$. Taking the product of the first representation we get $2^{\frac{p-1}{2}}(\frac{p-1}{2})!$, and the second we have, taking terms alternately from both ends of the sequence starting at the right, $(-1)^{1+2+3+\frac{p-1}{2}}(\frac{p-1}{2})! = (-1)^{\frac{p^2-1}{8}}(\frac{p-1}{2})!$. Cancelling the factorial these two equivalent expressions have in common we get $2^{\frac{p-1}{2}} \equiv (-1)^{\frac{p^2-1}{8}} \mod p$ and the result follows by equation (9) again.

$\diamond$

For example, to calculate $\left(\frac{263}{631}\right)$ we can proceed as follows, after verifying that 631 is indeed prime:

$$\begin{aligned}
\left(\frac{261}{631}\right) &= \left(\frac{3^2 \times 29}{631}\right) = \left(\frac{3^2}{631}\right)\left(\frac{29}{631}\right) = \left(\frac{29}{631}\right) \\
&= \left(\frac{631}{29}\right) = \left(\frac{22}{29}\right) \\
&= \left(\frac{2}{29}\right)\left(\frac{11}{29}\right) = -\left(\frac{11}{29}\right) \\
&= -\left(\frac{29}{11}\right) = -\left(\frac{7}{11}\right) = \left(\frac{11}{7}\right) \\
&= \left(\frac{4}{7}\right) = \left(\frac{2^2}{7}\right) = 1.
\end{aligned}$$

Thus there are solutions to $x^2 \equiv 261 \mod 631$ and they are, in fact, $x \equiv 92, 539 \mod 631$, but I used a computer to find that out!

Exercise 32 *Evaluate* $\left(\frac{14}{17}\right)$ *and* $\left(\frac{667}{421}\right)$.

8 Further Uses of the Euclidean Algorithm

As we saw in section 1.4 we can apply the Euclidean algorithm to any two integers to find their greatest common divisor. Alternatively we can also use it to find all solutions of linear Diophantine equations and also new representation for rational numbers and even some transcendental numbers.

8.1 Solution of Linear Diophantine Equations

As described above we consider $\frac{67}{24}$:

$$\begin{aligned}
67 &= 2 \times 24 + 19, & k_0 &= 2 \\
24 &= 1 \times 19 + 5, & k_1 &= 1 \\
19 &= 3 \times 5 + 4, & k_2 &= 3 \\
5 &= 1 \times 4 + 1, & k_3 &= 1 \\
4 &= 4 \times 1 + 0, & k_4 &= 4
\end{aligned}$$

and

r	-2	-1	0	1	2	3	4
k_r			2	1	3	1	4
α_r	0	1	2	3	11	14	67
β_r	1	0	1	1	4	5	24

We form this table using the k_i and the relations

$$\alpha_r = k_r\alpha_{r-1} + \alpha_{r-2} \qquad\qquad \beta_r = k_r\beta_{r-1} + \beta_{r-2}$$

and notice that the final column contains our original a and b. We shall show that the penultimate column gives a particular solution of the linear Diophantine equation $ax - by = (a, b)$.

Lemma 14 *The equation $ax + by = c$ is solvable if and only if $(a, b)|c$ and if $x = x_0$ and $y = y_0$ is a solution then all solutions are of the form $x = x_0 + \frac{bt}{(a,b)}$ and $y = y_0 - \frac{at}{(a,b)}$, $t \in \mathbb{Z}$.*

Proof: The condition for solvability is necessary and sufficient by noting that we can restate the equation as $ax \equiv c \bmod b$, which, by theorem 3.3, is solvable if and only if $(a, b)|c$.

We note that the given x and y are solutions since

$$ax + by = ax_0 + \frac{abt}{(a,b)} + by_0 - \frac{bat}{(a,b)} = ax_0 + by_0 = c.$$

Suppose $x = u$, $y = v$ is a solution which is not of the given form so that

$$\begin{aligned} au + bv &= c = ax_0 + by_0 \\ \Rightarrow a(u - x_0) &= b(y_0 - v) \\ \Rightarrow \frac{a}{(a,b)}(u - x_0) &= \frac{b}{(a,b)}(y_0 - v). \end{aligned} \tag{10}$$

Thus $\frac{a}{(a,b)}$ divides the right hand of equation (11) and, since it is relatively prime to $\frac{b}{(a,b)}$ by lemma 1.3 we have $\frac{a}{(a,b)}|y_0 - v$ and so $v = y_0 + t\frac{a}{(a,b)}$, contrary to our assumption. ◇

8.2 Matrix Representation

We can rewrite the working for the Euclidean algorithm in terms of matrices as follows:

$$\begin{aligned} \begin{pmatrix} 67 \\ 24 \end{pmatrix} &= \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 24 \\ 19 \end{pmatrix} \\ \begin{pmatrix} 24 \\ 19 \end{pmatrix} &= \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 19 \\ 5 \end{pmatrix} \\ \begin{pmatrix} 19 \\ 5 \end{pmatrix} &= \begin{pmatrix} 3 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 5 \\ 4 \end{pmatrix} \\ \begin{pmatrix} 5 \\ 4 \end{pmatrix} &= \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 4 \\ 1 \end{pmatrix} \\ \begin{pmatrix} 4 \\ 1 \end{pmatrix} &= \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \end{aligned}$$

Putting these together we get

$$\begin{pmatrix} 67 \\ 24 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 3 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} (67, 24) \\ 0 \end{pmatrix},$$

and this is true in general so that

$$\begin{pmatrix} a \\ b \end{pmatrix} = \prod_{i=0}^{n} \begin{pmatrix} k_i & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} (a,b) \\ 0 \end{pmatrix}.$$

The partial products of the matrices give us the parameters α_i and β_i as

$$\begin{pmatrix} \alpha_r & \alpha_{r-1} \\ \beta_r & \beta_{r-1} \end{pmatrix} := \prod_{i=0}^{r} \begin{pmatrix} k_i & 1 \\ 1 & 0 \end{pmatrix} \tag{11}$$

and we note that if we define $\alpha_{-1} := 1$, $\alpha_{-2} := 0$, $\beta_{-1} := 0$ and $\beta_{-2} := 1$ to initialise the recurrence relation then

$$\begin{pmatrix} \alpha_r & \alpha_{r-1} \\ \beta_r & \beta_{r-1} \end{pmatrix} := \begin{pmatrix} \alpha_{r-1} & \alpha_{r-2} \\ \beta_{r-1} & \beta_{r-2} \end{pmatrix} \begin{pmatrix} k_r & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} \alpha_{r-1}k_r + \alpha_{r-2} & \alpha_{r-1} \\ \beta_{r-1}k_r + \beta_{r-2} & \beta_{r-1} \end{pmatrix},$$

and we see why the previously stated recurrence relation holds.

Since the product of matrices in equation (11) has determinant $(-1)^{r+1}$ we can invert any of the matrices we have used and, in particular, when $r = n$ (so that $\alpha_r = a$ and $\beta_r = b$) we have

$$\begin{pmatrix} (a,b) \\ 0 \end{pmatrix} = (-1)^{n-1} \begin{pmatrix} \beta_{n-1} & -\alpha_{n-1} \\ -\beta_n & \alpha_n \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix}$$

and so $(a,b) = (-1)^{n-1}\beta_{n-1}a - (-1)^{n-1}\alpha_{n-1}b$, giving us a solution to our Diophantine equation from which we can generate all others as shown above.

8.3 Continued Fractions

The parameters k_i have other significance too - by dividing each line of the Euclidean algorithm by the quotient we get equations like this:

$$\begin{array}{rcl} \frac{67}{24} & = & 2 + \frac{19}{24} \\ \frac{24}{19} & = & 1 + \frac{5}{19} \\ \frac{19}{5} & = & 3 + \frac{4}{5} \\ \frac{5}{4} & = & 1 + \frac{1}{4} \\ \frac{4}{1} & = & 4 \end{array}$$

so that $\frac{67}{24} = 2 + \frac{1}{1+\frac{5}{19}} = 2 + \frac{1}{1+\frac{1}{3+\frac{4}{5}}} = 2 + \frac{1}{1+\frac{1}{3+\frac{1}{1+\frac{1}{4}}}}$. This is the *continued fraction* representation of a number and is usually written in the more convenient form $(2;1,3,1,4)$.

Any rational number can be written in the form $(k_0;k_1,k_2,\ldots,k_n)$, and this representation is unique if we insist that $k_n > 1$ (otherwise we can also have $(k_0;k_1,k_2,\ldots,k_n-1,1)$ as a valid representation). Each continued fraction is also unique (given $k_n > 1$) and we can find the value of it using the recurrence relations developed in the previous section as follows:
What is the value of $(0;3,1,1,5)$?
We form the table as before:

r	-2	-1	0	1	2	3	4
k_r			0	3	1	1	5
α_r	0	1	0	1	1	2	11
β_r	1	0	1	3	4	7	39

and thus the fraction is $\frac{11}{39}$. We can easily check this with a calculator by entering $\frac{11}{39}$ as a decimal and repeatedly doing the following: subtract the integer part, note it down and then take the reciprocal.

Exercise 33 *Find the fractional representation of* $(1;4,1,2,3)$ *and the continued fraction representation of* $\frac{44}{31}$.

You may have considered the fact that the method for finding the continued fraction representation of the fraction will actually work for any decimal, and indeed it will, although one must be careful of rounding errors on some calculators. For instance, we can investigate π, and find that its continued fraction representation starts $(3;7,15,1)$. If we just consider the first two figures we see that $(3;7)$ is equivalent to $\frac{22}{7}$, that well known approximation to π.

Exercise 34 *What fraction do the first three figures give you ? To how many decimal places is it accurate ?*

We shall now move back to simpler numbers which do have a pattern in their continued fraction representation, although they are still non-terminating. Consider $\sqrt{3}$, which has the form $(1;\overline{1,2})$, where the bar represents repeating digits, as we do with non-terminating decimals. We can see that it repeats as follows:

$$\begin{aligned}
\sqrt{3} &= 1+\sqrt{3}-1 \\
\frac{1}{\sqrt{3}-1} &= \frac{\sqrt{3}+1}{2} = 1+\frac{\sqrt{3}-1}{2} \\
\frac{2}{\sqrt{3}-1} &= \frac{2(\sqrt{3}+1)}{2} = 2+(\sqrt{3}-1) \\
\frac{1}{\sqrt{3}-1} &= \dots
\end{aligned}$$

We can do this analysis for any expression involving square roots, with more or less work. I find it easier to try to find the pattern on the calculator first and then do the arithmetic to check that there hasn't been a rounding or any other kind of error.

To reverse this procedure, we continue as follows: for $x=(0;1,\overline{4,2})$, we reorganise this to get an expression for the repeating part alone, so $y=1-\frac{1}{x}=(0;\overline{4,2})$. We then express y in terms of itself, and get the long form expression of the relation $y=(0;\overline{4,2,y})$

$$\begin{aligned}
y &= \frac{1}{4+\frac{1}{2+y}} \\
&= \frac{2+y}{4(2+y)+1} \\
&= \frac{2+y}{9+4y},
\end{aligned}$$

so that $y(9+4y)=2+y$. We solve $4y^2+8y-2=0$ to get $y=\frac{\sqrt{6}-2}{2}$ (we take the positive solution since we know y must be an positive number). We can then substitute this into our expression for x and, simplifying, we get $x=\frac{\sqrt{6}}{3}$. We can check this has the required continued fraction using a calculator, and also I recommend that you check the expression for y too, as it is easy to make an arithmetical error.

Exercise 35 *Find the continued fraction expression for* $3-\sqrt{7}$ *and which surd expression has the continued fraction* $(1;2,\overline{5,2})$.

9 Advanced Techniques

We now look at ways to apply the knowledge we have learnt so far in the course in more complicated ways:

9.1 Primality Testing

Currently, the only way we know to prove a number n is prime is to try to divide it by all the prime numbers up to $\sqrt{n}$. This soon becomes prohibitively time consuming once we get larger values of n. Thus we need to find a new method, or at least to refine our techniques. We shall see that it is easier to prove a number is *not* prime, but this at least helps us to dismiss candidates for primality more easily than finding a large divisor. The basis for the technique is equation (5), which can be rephrased to say:

$$\text{if } n \text{ is prime then } b^{n-1} \equiv 1 \bmod n \text{ for all } b \tag{12}$$

Thus, if we can find a b such that equation (12) is not satisfied then we can deduce that our particular n is not prime. It is possible to prove that, so long as equation (12) is not satisfied by all b (such a value for n is called a Carmichael number), it is satisfied by at most half, and so if we choose i values for b at random and get the answer 1 each time we know that either n is prime or Carmichael or there is a 1 in 2^i chance that it is composite. Thus, in practice we choose $i = 3$ or 4 and if we find that our n is composite we are happy, otherwise, we forget that n and choose another, hopefully easier, one.

Note that this method doesn't tell us any of the prime factors, just that the number is not prime. Note that since Carmichael numbers do exist, it is impossible to use this method to prove that any number is prime. The following theorem gives us some information about Carmichael numbers:

Theorem 9.1 *Let* n *be an odd composite integer.*

1. n *is not a Carmichael number if there is any prime* p *such that* $p^2|n$.
2. *if* n *is a Carmichael number then* $(p-1)|(n-1)$ *for all* $p|n$.

Exercise 36 *Prove that 561 satisfies both the conditions in theorem 9.1.*

9.2 Factorisation

Now we know a good way to prove numbers composite but one which does not give us the factors directly we need a new technique. Of course we can still use trial division for the smallest primes, say 2-31 (Maple uses 2-200), but once beyond there we need a more efficient idea. We shall look at Fermat factorisation, which is useful when there are no small factors in our number and the number breaks into two numbers which are close together. We note that if $n = ab$ and we can write $a = t + s$ and $b = t - s$ then $n = t^2 - s^2$ and so $s^2 = t^2 - n$. Thus, we take values of t just greater then $\sqrt{n}$ and subtract n and take the square root, which will be s. If this number is an integer we

can easily find a and b and hence factorise n. However, it is sometimes easier to instead start near $\sqrt{kn}$ (for small odd values of k), since we can then find not only factorisations in which two factors are approximately equal, but ones in which there exist factors which are twice, or three times the other. If we use even values of k it is necessary to consider half-integers, which complicates matters greatly.

For example, to factorise 9523, we can calculate $\sqrt{9523} > 97$. Taking $t = 98$ we get $s = \sqrt{(}9604 - 9523) = 9$. Hence $a = 98 + 9 = 107$ and $b = 98 - 9 = 89$. If we try to factorise 18881 by the same method we would start with $t = 137$ and would need 22 tries before we got $18881 = (159+80)(159-80)$. However, if we use $\sqrt{3 \times 18881} > 237$ we get the answer immediately, since $238^2 - 3 \times 18881 = 1$ and so we can write

$$3 \times 18881 = 237 \times 239,$$

and deduce that 239 is a factor of 18881, and so the other factor is 79.

Exercise 37 *Factorise 8265, 16157 and 19637.*

Pollard $p-1$ Factorisation

We have a number n which has been proven to be composite, but no factors are known. We form a sequence of numbers a_i, starting with a randomly chosen a_0. Given a_i, we calculate $\gcd(a_i - 1, n)$ and if this is greater than 1 then it is a factor of n. If not we generate $a_{i+1} \equiv a_i^{i+1} \mod n$ by the repeated squaring method if necessary and repeat the step until a factor is found, or i becomes too large to handle. If no factor is found then the process can be repeated with a different choice of a_0.

For example, we try to factor 323: Choosing the start number $a_0 = 26$, we first check gcd(26-1,323) = 1 since $25 = 5^2$ and 5 doesn't divide 323. We then calculate $a_1 := 26^2 \equiv 30 \mod 323$ and gcd(30-1,323)=1 since 29 is prime and doesn't divide 323. Next we calculate $a_2 := 30^3 \equiv 191 \mod 323$ and then check gcd(191-1,323). Using the Euclidean algorithm we get $323 = 2 \times 190 - 57$ and $190 = 3 \times 57 + 19$ and $57 = 3 \times 19$ so 19 is a divisor of 323, which we can check, and see that the factorisation of 323 is 19×17. Note that this algorithm doesn't necessarily find the smallest factor, and the factor found and the time taken depends on the starting value chosen.

Pollard Rho Factorisation

This time, to factorise n, we choose two numbers, a and b, and use the recursive formula $x_{i+1} \equiv {x_i}^2 + b \mod n$ with $x_1 := a$. We want to check when $\gcd(|x_i - x_j|, n) > 1$ in which case this will again be a factor. However, it is a lot of work to do this for each possible pair of values for x_i and x_j so we use a method which won't always find the first available pair of values, but will overall find a pair quickly. We keep track of special values of x_i, those when i is a power of 2. We let the last such x_i be y. We then utilise each new x_i generated by calculating $\gcd(—x_i - y|, n)$ each time.

The procedure would go as follows for $n = 713$ would go as follows, supposing we choose $a = 19$ and $b = 10$:

1. $x_1 := a = 19$, let $y := x_1 = 19$
2. $x_2 := 19^2 + 10 = 371$, check gcd(371-19,713)=1, set $y := 371$
3. $x_3 := 371^2 + 10 \equiv 42 \mod 713$, check gcd(371-42,713)=1

4. $x_4 := 42^2 + 10 \equiv 348 \bmod 713$, check gcd(371-348,713)=23

Therefore 23 is a factor of 713, as we required. If this step hadn't worked we would have set $y := 348$ and generated x_5 through to x_8.

9.3 Public Key Cryptography

In all the cryptosystems we currently know, once we know the parameters used to encrypt a message is is quite easy to reverse this transformation and hence the decryption method. The idea behind public key cryptography is that even if we know the encryption parameters it is still mathematically difficult (perhaps even impossible) to calculate the decoding formula.

The most widespread method of public key encryption is called RSA, and it gets its security from the difficulty of factorising large integers. Each user chooses two large (say 100 decimal places) prime numbers p and q and calculates $\phi(pq) = (p-1)(q-1) = pq - p - q + 1$, and selects a number e less than this relatively prime to both p and q. The encryption key then uses pq and e, and the encoding method is to take the value of each letter in our word and raise it to the power e mod pq. In order to decode, we need to find the inverse of e mod $\phi(pq)$, since $a^{\phi(pq)} \equiv 1 \bmod pq$. This is easy if we know p and q, but almost impossible with current computer power and techniques if we don't.

Thus, in practice, everyone who wishes to use RSA to communicate with anyone just publishes their values of pq and e, and the person will encrypt the message for them using these numbers. However, no-one else can work out e^{-1} mod $\phi(pq)$, and so only the intended recipient can read the message.