> conclusions were as infallible as so many propositions of Euclid. So startling would his results appear to the uninitiated that until they learned the processes by which he had arrived at them they might well consider him as a necromancer.[7]

1.2 Divisibility and greatest common divisors

Much of modern cryptography is built on the foundations of algebra and number theory. So before we explore the subject of cryptography, we need to develop some important tools. In the next four sections we begin this development by describing and proving fundamental results from algebra and number theory. If you have already studied number theory in another course, a brief review of this material will suffice. But if this material is new to you, then it is vital to study it closely and to work out the exercises provided at the end of the chapter.

At the most basic level, *Number Theory* is the study of the natural numbers

$$1, 2, 3, 4, 5, 6, \dots,$$

or slightly more generally, the study of the integers

$$\dots, -5, -4, -3, -2, -1, 0, 1, 2, 3, 4, 5, \dots.$$

The set of integers is denoted by the symbol $\mathbb{Z}$. Integers can be added, subtracted, and multiplied in the usual way, and they satisfy all the usual rules of arithmetic (commutative law, associative law, distributive law, etc.). The set of integers with their addition and multiplication rules are an example of a *ring*. See Section 2.10.1 for more about the theory of rings.

If a and b are integers, then we can add them $a + b$, subtract them $a - b$, and multiply them $a \cdot b$. In each case, we get an integer as the result. This property of staying inside of our original set after applying operations to a pair of elements is characteristic of a ring.

But if we want to stay within the integers, then we are not always able to divide one integer by another. For example, we cannot divide 3 by 2, since there is no integer that is equal to $\frac{3}{2}$. This leads to the fundamental concept of divisibility.

Definition. Let a and b be integers with $b \neq 0$. We say that b *divides* a, or that a *is divisible by* b, if there is an integer c such that

$$a = bc.$$

We write $b \mid a$ to indicate that b divides a. If b does not divide a, then we write $b \nmid a$.

[7] *A Study in Scarlet* (Chapter 2), Sir Arthur Conan Doyle.

Example 1.2. We have $847 \mid 485331$, since $485331 = 847 \cdot 573$. On the other hand, $355 \nmid 259943$, since when we try to divide 259943 by 355, we get a remainder of 83. More precisely, $259943 = 355 \cdot 732 + 83$, so 259943 is not an exact multiple of 355.

Remark 1.3. Notice that every integer is divisible by 1. The integers that are divisible by 2 are the *even integers*, and the integers that are not divisible by 2 are the *odd integers*.

There are a number of elementary divisibility properties, some of which we list in the following proposition.

Proposition 1.4. *Let $a, b, c \in \mathbb{Z}$ be integers.*

(a) *If $a \mid b$ and $b \mid c$, then $a \mid c$.*

(b) *If $a \mid b$ and $b \mid a$, then $a = \pm b$.*

(c) *If $a \mid b$ and $a \mid c$, then $a \mid (b + c)$ and $a \mid (b - c)$.*

Proof. We leave the proof as an exercise for the reader; see Exercise 1.6. □

Definition. A *common divisor* of two integers a and b is a positive integer d that divides both of them. The *greatest common divisor* of a and b is, as its name suggests, the largest positive integer d such that $d \mid a$ and $d \mid b$. The greatest common divisor of a and b is denoted $\gcd(a, b)$. If there is no possibility of confusion, it is also sometimes denoted by (a, b). (If a and b are both 0, then $\gcd(a, b)$ is not defined.)

It is a curious fact that a concept as simple as the greatest common divisor has many applications. We'll soon see that there is a fast and efficient method to compute the greatest common divisor of any two integers, a fact that has powerful and far-reaching consequences.

Example 1.5. The greatest common divisor of 12 and 18 is 6, since $6 \mid 12$ and $6 \mid 18$ and there is no larger number with this property. Similarly,

$$\gcd(748, 2024) = 44.$$

One way to check that this is correct is to make lists of all of the positive divisors of 748 and of 2024.

$$\begin{aligned} \text{Divisors of } 748 &= \{1, 2, 4, 11, 17, 22, 34, 44, 68, 187, 374, 748\}, \\ \text{Divisors of } 2024 &= \{1, 2, 4, 8, 11, 22, 23, 44, 46, 88, 92, 184, 253, \\ &\qquad\qquad 506, 1012, 2024\}. \end{aligned}$$

Examining the two lists, we see that the largest common entry is 44. Even from this small example, it is clear that this is not a very efficient method. If we ever need to compute greatest common divisors of large numbers, we will have to find a more efficient approach.

The key to an efficient algorithm for computing greatest common divisors is *division with remainder*, which is simply the method of "long division" that you learned in elementary school. Thus if a and b are positive integers and if you attempt to divide a by b, you will get a quotient q and a remainder r, where the remainder r is smaller than b. For example,

```
      13 R 9
17 ) 230
     17
      60
      51
       9
```

so 230 divided by 17 gives a quotient of 13 with a remainder of 9. What does this last statement really mean? It means that 230 can be written as

$$230 = 17 \cdot 13 + 9,$$

where the remainder 9 is strictly smaller than the divisor 17.

Definition. (Division Algorithm) Let a and b be positive integers. Then *a divided by b has quotient q and remainder r* means that

$$a = b \cdot q + r \qquad \text{with } 0 \le r < b.$$

The values of q and r are uniquely determined by a and b.

Suppose now that we want to find the greatest common divisor of a and b. We first divide a by b to get

$$a = b \cdot q + r \qquad \text{with } 0 \le r < b. \tag{1.1}$$

If d is any common divisor of a and b, then it is clear from equation (1.1) that d is also a divisor of r. (See Proposition 1.4(c).) Similarly, if e is a common divisor of b and r, then (1.1) shows that e is a divisor of a. In other words, the common divisors of a and b are the same as the common divisors of b and r; hence

$$\gcd(a, b) = \gcd(b, r).$$

We repeat the process, dividing b by r to get another quotient and remainder, say

$$b = r \cdot q' + r' \qquad \text{with } 0 \le r' < r.$$

Then the same reasoning shows that

$$\gcd(b, r) = \gcd(r, r').$$

Continuing this process, the remainders become smaller and smaller, until eventually we get a remainder of 0, at which point the final value $\gcd(s, 0) = s$ is equal to the gcd of a and b.

We illustrate with an example and then describe the general method, which goes by the name *Euclidean algorithm*.

Example 1.6. We compute $\gcd(2024, 748)$ using the Euclidean algorithm, which is nothing more than repeated division with remainder. Notice how the quotient and remainder on each line become the new a and b on the subsequent line:

$$\begin{aligned}
2024 &= 748 \cdot 2 + 528 \\
748 &= 528 \cdot 1 + 220 \\
528 &= 220 \cdot 2 + 88 \\
220 &= 88 \cdot 2 + 44 \qquad \leftarrow \boxed{\gcd = 44} \\
88 &= 44 \cdot 2 + 0
\end{aligned}$$

Theorem 1.7 (The Euclidean Algorithm). *Let a and b be positive integers with $a \geq b$. The following algorithm computes $\gcd(a, b)$ in a finite number of steps.*

(1) *Let $r_0 = a$ and $r_1 = b$.*

(2) *Set $i = 1$.*

(3) *Divide r_{i-1} by r_i to get a quotient q_i and remainder r_{i+1},*

$$r_{i-1} = r_i \cdot q_i + r_{i+1} \qquad \textit{with} \quad 0 \leq r_{i+1} < r_i.$$

(4) *If the remainder $r_{i+1} = 0$, then $r_i = \gcd(a, b)$ and the algorithm terminates.*

(5) *Otherwise, $r_{i+1} > 0$, so set $i = i + 1$ and go to Step 3.*

The division step (Step 3) is executed at most

$$2 \log_2(b) + 1 \quad \textit{times.}$$

Proof. The Euclidean algorithm consists of a sequence of divisions with remainder as illustrated in Figure 1.2 (remember that we set $r_0 = a$ and $r_1 = b$).

$$\boxed{\begin{array}{rcll}
a &=& b \cdot q_1 + r_2 & \text{with } 0 \leq r_2 < b, \\
b &=& r_2 \cdot q_2 + r_3 & \text{with } 0 \leq r_3 < r_2, \\
r_2 &=& r_3 \cdot q_3 + r_4 & \text{with } 0 \leq r_4 < r_3, \\
r_3 &=& r_4 \cdot q_4 + r_5 & \text{with } 0 \leq r_5 < r_4, \\
\vdots & & \vdots & \qquad\qquad \vdots \\
r_{t-2} &=& r_{t-1} \cdot q_{t-1} + r_t & \text{with } 0 \leq r_t < r_{t-1}, \\
r_{t-1} &=& r_t \cdot q_t & \\
\multicolumn{4}{c}{\text{Then } r_t = \gcd(a, b).}
\end{array}}$$

Figure 1.2: The Euclidean algorithm step by step

The r_i values are strictly decreasing, and as soon as they reach zero the algorithm terminates, which proves that the algorithm does finish in a finite number of steps. Further, at each iteration of Step 3 we have an equation of the form

$$r_{i-1} = r_i \cdot q_i + r_{i+1}.$$

This equation implies that any common divisor of r_{i-1} and r_i is also a divisor of r_{i+1}, and similarly it implies that any common divisor of r_i and r_{i+1} is also a divisor of r_{i-1}. Hence

$$\gcd(r_{i-1}, r_i) = \gcd(r_i, r_{i+1}) \qquad \text{for all } i = 1, 2, 3, \ldots. \tag{1.2}$$

However, as noted above, we eventually get to an r_i that is zero, say $r_{t+1} = 0$. Then $r_{t-1} = r_t \cdot q_t$, so

$$\gcd(r_{t-1}, r_t) = \gcd(r_t \cdot q_t, r_t) = r_t.$$

But equation (1.2) says that this is equal to $\gcd(r_0, r_1)$, i.e., to $\gcd(a, b)$, which completes the proof that the last nonzero remainder in the Euclidean algorithm is equal to the greatest common divisor of a and b.

It remains to estimate the efficiency of the algorithm. We noted above that since the r_i values are strictly decreasing, the algorithm terminates, and indeed since $r_1 = b$, it certainly terminates in at most b steps. However, this upper bound is far from the truth. We claim that after every two iterations of Step 3, the value of r_i is at least cut in half. In other words:

$$\textbf{Claim:} \quad r_{i+2} < \tfrac{1}{2} r_i \quad \text{for all } i = 0, 1, 2, \ldots.$$

We prove the claim by considering two cases.

Case I: $r_{i+1} \le \frac{1}{2} r_i$

We know that the r_i values are strictly decreasing, so

$$r_{i+2} < r_{i+1} \le \tfrac{1}{2} r_i.$$

Case II: $r_{i+1} > \frac{1}{2} r_i$

Consider what happens when we divide r_i by r_{i+1}. The value of r_{i+1} is so large that we get

$$r_i = r_{i+1} \cdot 1 + r_{i+2} \quad \text{with} \quad r_{i+2} = r_i - r_{i+1} < r_i - \tfrac{1}{2} r_i = \tfrac{1}{2} r_i.$$

We have now proven our claim that $r_{i+2} < \frac{1}{2} r_i$ for all i. Using this inequality repeatedly, we find that

$$r_{2k+1} < \frac{1}{2} r_{2k-1} < \frac{1}{4} r_{2k-3} < \frac{1}{8} r_{2k-5} < \frac{1}{16} r_{2k-7} < \cdots < \frac{1}{2^k} r_1 = \frac{1}{2^k} b.$$

Hence if $2^k \ge b$, then $r_{2k+1} < 1$, which forces r_{2k+1} to equal 0 and the algorithm to terminate. In terms of Figure 1.2, the value of r_{t+1} is 0, so we have $t+1 \le 2k+1$, and thus $t \le 2k$. Further, there are exactly t divisions performed in Figure 1.2, so the Euclidean algorithm terminates in at most $2k$ iterations. Choose the smallest such k, so $2^k \ge b > 2^{k-1}$. Then

$$\#\text{ of iterations} \le 2k = 2(k-1) + 2 < 2\log_2(b) + 2,$$

which completes the proof of Theorem 1.7. □

Remark 1.8. We proved that the Euclidean algorithm applied to a and b with $a \ge b$ requires no more than $2\log_2(b) + 1$ iterations to compute $\gcd(a, b)$. This estimate can be somewhat improved. It has been proven that the Euclidean algorithm takes no more than $1.45\log_2(b) + 1.68$ iterations, and that the average number of iterations for randomly chosen a and b is approximately $0.85\log_2(b) + 0.14$. (See [61].)

Remark 1.9. One way to compute quotients and remainders is by long division, as we did on page 12. You can speed up the process using a simple calculator. The first step is to divide a by b on your calculator, which will give a real number. Throw away the part after the decimal point to get the quotient q. Then the remainder r can be computed as

$$r = a - b \cdot q.$$

For example, let $a = 2387187$ and $b = 27573$. Then $a/b \approx 86.57697748$, so $q = 86$ and

$$r = a - b \cdot q = 2387187 - 27573 \cdot 86 = 15909.$$

If you need just the remainder, you can instead take the decimal part (also sometimes called the *fractional part*) of a/b and multiply it by b. Continuing with our example, the decimal part of $a/b \approx 86.57697748$ is 0.57697748, and multiplying by $b = 27573$ gives

$$27573 \cdot 0.57697748 = 15909.00005604.$$

Rounding this off gives $r = 15909$.

After performing the Euclidean algorithm on two numbers, we can work our way back up the process to obtain an extremely interesting formula. Before giving the general result, we illustrate with an example.

Example 1.10. Recall that in Example 1.6 we used the Euclidean algorithm to compute $\gcd(2024, 748)$ as follows:

$$\begin{aligned}
2024 &= 748 \cdot 2 + 528 \\
748 &= 528 \cdot 1 + 220 \\
528 &= 220 \cdot 2 + 88 \\
220 &= 88 \cdot 2 + 44 \qquad \leftarrow \boxed{\gcd = 44} \\
88 &= 44 \cdot 2 + 0
\end{aligned}$$

We let $a = 2024$ and $b = 748$, so the first line says that

$$528 = a - 2b.$$

We substitute this into the second line to get

$$b = (a - 2b) \cdot 1 + 220, \qquad \text{so} \qquad 220 = -a + 3b.$$

We next substitute the expressions $528 = a - 2b$ and $220 = -a + 3b$ into the third line to get

$$a - 2b = (-a + 3b) \cdot 2 + 88, \qquad \text{so} \qquad 88 = 3a - 8b.$$

Finally, we substitute the expressions $220 = -a + 3b$ and $88 = 3a - 8b$ into the penultimate line to get

$$-a + 3b = (3a - 8b) \cdot 2 + 44, \qquad \text{so} \qquad 44 = -7a + 19b.$$

In other words,

$$-7 \cdot 2024 + 19 \cdot 748 = 44 = \gcd(2024, 748),$$

so we have found a way to write $\gcd(a, b)$ as a linear combination of a and b using integer coefficients.

In general, it is always possible to write $\gcd(a, b)$ as an integer linear combination of a and b, a simple sounding result with many important consequences.

Theorem 1.11 (Extended Euclidean Algorithm)**.** *Let a and b be positive integers. Then the equation*

$$au + bv = \gcd(a, b)$$

always has a solution in integers u and v. (See Exercise 1.12 for an efficient algorithm to find a solution.)

If (u_0, v_0) is any one solution, then every solution has the form

$$u = u_0 + \frac{b \cdot k}{\gcd(a, b)} \quad \textit{and} \quad v = v_0 - \frac{a \cdot k}{\gcd(a, b)} \qquad \textit{for some } k \in \mathbb{Z}.$$

Proof. Look back at Figure 1.2, which illustrates the Euclidean algorithm step by step. We can solve the first line for $r_2 = a - b \cdot q_1$ and substitute it into the second line to get

$$b = (a - b \cdot q_1) \cdot q_2 + r_3, \qquad \text{so} \qquad r_3 = -a \cdot q_2 + b \cdot (1 + q_1 q_2).$$

Next substitute the expressions for r_2 and r_3 into the third line to get

$$a - b \cdot q_1 = \bigl(-a \cdot q_2 + b \cdot (1 + q_1 q_2)\bigr) q_3 + r_4.$$

After rearranging the terms, this gives

$$r_4 = a \cdot (1 + q_2 q_3) - b \cdot (q_1 + q_3 + q_1 q_2 q_3).$$

The key point is that $r_4 = a \cdot u + b \cdot v$, where u and v are integers. It does not matter that the expressions for u and v in terms of q_1, q_2, q_3 are rather messy. Continuing in this fashion, at each stage we find that r_i is the sum of an integer multiple of a and an integer multiple of b. Eventually, we get to $r_t = a \cdot u + b \cdot v$ for some integers u and v. But $r_t = \gcd(a, b)$, which completes the proof of the first part of the theorem. We leave the second part as an exercise (Exercise 1.11). □

An especially important case of the extended Euclidean algorithm arises when the greatest common divisor of a and b is 1. In this case we give a and b a special name.

Definition. Let a and b be integers. We say that a and b are *relatively prime* if $\gcd(a, b) = 1$.

More generally, any equation

$$Au + Bv = \gcd(A, B)$$

can be reduced to the case of relatively prime numbers by dividing both sides by $\gcd(A, B)$. Thus

$$\frac{A}{\gcd(A, B)} u + \frac{B}{\gcd(A, B)} v = 1,$$

where $a = A/\gcd(A, B)$ and $b = B/\gcd(A, B)$ are relatively prime and satisfy $au + bv = 1$. For example, we found earlier that 2024 and 748 have greatest common divisor 44 and satisfy

$$-7 \cdot 2024 + 19 \cdot 748 = 44.$$

Dividing both sides by 44, we obtain

$$-7 \cdot 46 + 19 \cdot 17 = 1.$$

Thus $2024/44 = 46$ and $748/44 = 17$ are relatively prime, and $u = -7$ and $v = 19$ are the coefficients of a linear combination of 46 and 17 that equals 1.

In Example 1.10 we explained how to substitute the values from the Euclidean algorithm in order to solve $au + bv = \gcd(a, b)$. Exercise 1.12 describes an efficient computer-oriented algorithm for computing u and v. If a and b are relatively prime, we now describe a more conceptual version of this substitution procedure. We first illustrate with the example $a = 73$ and $b = 25$. The Euclidean algorithm gives

$$\begin{aligned} 73 &= 25 \cdot 2 + 23 \\ 25 &= 23 \cdot 1 + 2 \\ 23 &= 2 \cdot 11 + 1 \\ 2 &= 1 \cdot 2 + 0. \end{aligned}$$

We set up a box, using the sequence of quotients 2, 1, 11, and 2, as follows:

		2	1	11	2
0	1	$*$	$*$	$*$	$*$
1	0	$*$	$*$	$*$	$*$

Then the rule to fill in the remaining entries is as follows:

$$\begin{aligned} \text{New Entry} = (\text{Number at Top}) \cdot (\text{Number to the Left}) \\ + (\text{Number Two Spaces to the Left}). \end{aligned}$$

Thus the two leftmost $*$'s are

$$2 \cdot 1 + 0 = 2 \qquad \text{and} \qquad 2 \cdot 0 + 1 = 1,$$

so now our box looks like this:

		2	1	11	2
0	1	2	$*$	$*$	$*$
1	0	1	$*$	$*$	$*$

Then the next two leftmost $*$'s are

$$1 \cdot 2 + 1 = 3 \qquad \text{and} \qquad 1 \cdot 1 + 0 = 1,$$

and then the next two are

$$11 \cdot 3 + 2 = 35 \qquad \text{and} \qquad 11 \cdot 1 + 1 = 12,$$

and the final entries are

$$2 \cdot 35 + 3 = 73 \qquad \text{and} \qquad 2 \cdot 12 + 1 = 25.$$

The completed box is

		2	1	11	2
0	1	2	3	35	73
1	0	1	1	12	25

Notice that the last column repeats a and b. More importantly, the next to last column gives the values of $-v$ and u (in that order). Thus in this example we find that $73 \cdot 12 - 25 \cdot 35 = 1$. The general algorithm is given in Figure 1.3.

In general, if a and b are relatively prime and if $q_1, q_2, \ldots, q_t$ is the sequence of quotients obtained from applying the Euclidean algorithm to a and b as in Figure 1.2 on page 13, then the box has the form

		q_1	q_2	$\cdots$	q_{t-1}	q_t
0	1	P_1	P_2	$\ldots$	P_{t-1}	a
1	0	Q_1	Q_2	$\ldots$	Q_{t-1}	b

The entries in the box are calculated using the initial values

$$P_1 = q_1, \qquad Q_1 = 1, \qquad P_2 = q_2 \cdot P_1 + 1, \qquad Q_2 = q_2 \cdot Q_1,$$

and then, for $i \geq 3$, using the formulas

$$P_i = q_i \cdot P_{i-1} + P_{i-2} \qquad \text{and} \qquad Q_i = q_i \cdot Q_{i-1} + Q_{i-2}.$$

The final four entries in the box satisfy

$$a \cdot Q_{t-1} - b \cdot P_{t-1} = (-1)^t.$$

Multiplying both sides by $(-1)^t$ gives the solution $u = (-1)^t Q_{t-1}$ and $v = (-1)^{t+1} P_{t-1}$ to the equation $au + bv = 1$.

Figure 1.3: Solving $au + bv = 1$ using the Euclidean algorithm

1.3 Modular arithmetic

You may have encountered "clock arithmetic" in grade school, where after you get to 12, the next number is 1. This leads to odd-looking equations such as

$$6 + 9 = 3 \qquad \text{and} \qquad 2 - 3 = 11.$$

These look strange, but they are true using clock arithmetic, since for example 11 o'clock is 3 hours before 2 o'clock. So what we are really doing is first computing $2 - 3 = -1$ and then adding 12 to the answer. Similarly, 9 hours after 6 o'clock is 3 o'clock, since $6 + 9 - 12 = 3$.

The theory of *congruences* is a powerful method in number theory that is based on the simple idea of clock arithmetic.

Definition. Let $m \geq 1$ be an integer. We say that the integers a and b are *congruent modulo m* if their difference $a - b$ is divisible by m. We write

$$a \equiv b \pmod{m}$$

to indicate that a and b are congruent modulo m. The number m is called the *modulus*.