

1 A Refresher on the Integers

A good place to start might be the set of *integers*:

$$\mathbb{Z} = \{0, \pm 1, \pm 2, \pm 3, \ldots\}.$$

We often encounter the notation \mathbb{N} for the set of *natural numbers*. There seems to be no consensus on whether to include 0 as a natural number. Following the advice of Paul Halmos (1916–2006) in his seminal book *Naive Set Theory*, first published in 1960, we shall include 0 as a natural number. Thus \mathbb{N} stands for the set $\{0,1,2,3,\ldots\}$. We will on occasion use the somewhat less common notation \mathbb{P} for the set $\{1,2,3,\ldots\}$ of *positive integers*. Often we shall prefer to say things like "let x be a positive integer," instead of "let $x \in \mathbb{P}$."

1.1 Euclidean Division and the Greatest Common Divisor

In the seventh book of Euclid we are told that any integer b, upon division by a positive integer a, yields a unique quotient q and a unique remainder r such that

$$b = aq + r$$
 and $0 \le r < a$.

Indeed, q is the largest integer that is less than or equal to the fraction b/a, and r is simply b-aq. The procurement of q and r will be called **Euclidean division**.

If the remainder r = 0, then b = aq. In this case we say that a divides b, and write $a \mid b$. For instance, $17 \mid 51$ since $51 = 17 \cdot 3$. Conversely, a bit of reflection shows that if $r \neq 0$, then $a \nmid b$. For example, the Euclidean division

$$-91 = 17 \cdot (-6) + 11$$
 where $0 < 11 < 17$,

reveals that 17 does not divide -91. For visual clarity we may, on occasion just as above, use a dot to signify multiplication.

When $a \mid b$, we say that b is a **multiple** of a, and also that a is a **divisor** or **factor** of b. Here are a few simple things to note.

- If $a \mid b$ and $b \mid c$, then $a \mid c$.
- If $a \mid b$ and $a \mid c$ and x, y are any integers, then $a \mid bx + cy$.
- If $b \neq 0$ and $a \mid b$, then $|a| \leq |b|$.
- If $a \mid b$ and $b \mid a$, then $a = \pm b$.



2 A Refresher on the Integers

The Greatest Common Divisor of Two Integers

If a, b are integers, an *integer combination* of a and b is any integer c built up as

$$c = ax + by$$
 for some integers x, y .

For example, the equation $6 \cdot (-3) + 4 \cdot 7 = 10$ reveals that 10 is an integer combination of 6 and 4. But 11 is not an integer combination of 6 and 4, because integer combinations of 6 and 4 have to be even integers. We might note that the same integer combination c could be formed as an integer combination of a and b in more than one way. For example 10 also equals $6 \cdot 1 + 4 \cdot 1$.

If a, b are integers which are not both 0, then there do exist integer combinations of a and b which are positive. For instance $a \cdot a + b \cdot b = a^2 + b^2 > 0$. Hence there exist integers s, t which cause as + bt to be minimal among the positive, integer combinations of a and b.

Now comes a noteworthy result.

Proposition 1.1 (Greatest common divisor). *Let* a, b *be integers, and suppose at least one of them is not* 0. *If an integer* d *satisfies any one of the following properties, then* d *satisfies them all.*

- 1. *d is the least among the positive, integer combinations of a and b.*
- 2. *d is a positive, integer combination of a and b, and d divides every integer combination of a and b.*
- 3. *d is a positive, common divisor of a and b, and every common divisor of a and b is a divisor of d.*
- 4. *d is the largest among the numbers that divide both a and b.*

Proof. We will prove that $1 \implies 2 \implies 3 \implies 4 \implies 1$.

 $1 \implies 2$. Let d = as + bt be that positive integer combination of a and b which is minimal, and let c = ax + by represent any integer combination of a and b. By Euclidean division there exist integers q, r such that

$$c = dq + r$$
 where $0 < r < d$.

And so

$$r = c - dq = ax + by - (as + bt)q = a(x - sq) + b(y - tq),$$

which is evidently another integer combination of a and b. By the minimality of d, the nonnegative remainder r cannot be positive. Thus r = 0, which gives $d \mid c$.

- $2 \implies 3$. We have that 0 < d = ax + by for some integers x, y and d divides every integer combination of a and b. Thereby d is a common divisor of a and b, because a and b are themselves integer combinations of a and b. Also if c is another common divisor of a and b, then $c \mid ax + by$ too. So $c \mid d$.
- $3 \implies 4$. By assumption every common divisor c of a and b is a divisor of d, and d > 0. Clearly then d > c.
- $4 \implies 1$. By the proven implications $1 \implies 2 \implies 3 \implies 4$ the unique least, positive, integer combination of a and b is the largest integer that divides both a and b. To repeat,



1.1 Euclidean Division and the Greatest Common Divisor

the largest integer that divides both a and b is the least, positive, integer combination of a and b.

Definition 1.2. The number d that satisfies any one, and thereby all, of the conditions of Proposition 1.1 is called the **greatest common divisor** of a and b. We write that positive integer as

$$gcd(a,b)$$
.

What should gcd(0,0) be? Well, since the greatest common divisor should be an integer combination of a and b, it makes sense to say that gcd(0,0) = 0, because that is the only possible integer combination of 0 and 0. The choice gcd(0,0) = 0 is further justified by the fact that 0 is the only divisor of 0, which is also divisible by all other divisors of 0.

The Euclidean Algorithm

For pairs of small integers, their greatest common divisor can be seen by inspection. For instance, gcd(42,30) = 6. When the integers become large, there is an efficient technique for finding their greatest common divisor, based on repeated use of Euclidean division. Clearly,

$$gcd(a,b) = gcd(b,a) = gcd(\pm a, \pm b), gcd(0,b) = |b| and gcd(b,b) = |b|.$$

Thus, for the purpose of computing greatest common divisors of a and b, we need only consider the situation where 0 < a < b.

If a, b are not both 0 and b = aq + r for some integers q, r, one can see that the set of common divisors of b and a is the same as the set of common divisors of a and b. Therefore

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

The preceding remark points the way for Euclidean division to find gcd(b, a).

Say 0 < a < b. Apply Euclidean division repeatedly as follows:

$$b = aq_1 + r_1 & 0 < r_1 < a$$

$$a = r_1q_2 + r_2 & 0 < r_2 < r_1$$

$$r_1 = r_2q_3 + r_3 & 0 < r_3 < r_2$$

$$r_2 = r_3q_4 + r_4 & 0 < r_4 < r_3,$$

to obtain strictly decreasing remainders $r_1 > r_2 > r_3 > \cdots \geq 0$.

Sooner or later an integer remainder becomes 0. In other words, there must be an index n such that

$$r_{n-3} = r_{n-2}q_{n-1} + r_{n-1}$$
 $0 < r_{n-1} < r_{n-2}$
 $r_{n-2} = r_{n-1}q_n + r_n$ $0 < r_n < r_{n-1}$
 $r_{n-1} = r_nq_{n+1} + 0$ $0 = r_{n+1}$.

3



4 A Refresher on the Integers

From the remark preceding the above algorithm:

$$r_n = \gcd(r_n, 0) = \gcd(r_{n-1}, r_n) = \gcd(r_{n-2}, r_{n-1}) = \cdots$$

= $\gcd(r_3, r_4) = \gcd(r_2, r_3) = \gcd(r_1, r_2) = \gcd(a, r_1) = \gcd(b, a).$

The last positive remainder r_n equals gcd(b,a). This famous process for obtaining greatest common divisors is called the *Euclidean algorithm*.

For example, here comes gcd(8316,4641):

$$8316 = 4641 \cdot 1 + 3675$$

$$4641 = 3675 \cdot 1 + 966$$

$$3675 = 966 \cdot 3 + 777$$

$$966 = 777 \cdot 1 + 189$$

$$777 = 189 \cdot 4 + 21$$

$$189 = 21 \cdot 9 + 0$$

Thus, gcd(8316, 4641) = 21.

Roughly, the number of steps in the Euclidean algorithm is no more than twice the base-two logarithm of b. Consequently, it comes as no surprise that machines can rapidly implement the Euclidean algorithm for enormous integers with hundreds of digits.

According to Proposition 1.1, gcd(a,b) = ax+by for some integers x, y. Inside the Euclidean algorithm lies the method of obtaining such x, y. It is a matter of backtracking up along the algorithm. For instance, in the preceding worked example:

$$\gcd(8316, 4641) = 21$$

$$= 777 \cdot 1 - 189 \cdot 4$$

$$= 777 \cdot 1 - (966 - 777 \cdot 1) \cdot 4$$

$$= 777 \cdot 5 - 966 \cdot 4$$

$$= (3675 - 966 \cdot 3) \cdot 5 - 966 \cdot 4$$

$$= 3675 \cdot 5 - 966 \cdot 19$$

$$= 3675 \cdot 5 - (4641 - 3675 \cdot 1) \cdot 19$$

$$= 3675 \cdot 24 - 4641 \cdot 19$$

$$= (8316 - 4641 \cdot 1) \cdot 24 - 4641 \cdot 19$$

$$= 8316 \cdot 24 - 4641 \cdot 43.$$

Thus, gcd(8316,4641) = 8316x + 4641y where x = 24 and y = -43.

Calculations such as these can of course be done by machine when the integers in question are big.



1.1 Euclidean Division and the Greatest Common Divisor

5

Coprime Integers

Definition 1.3. If the greatest common divisor of integers a and b equals 1, we say that a, b are *coprime*.

Evidently, two integers are coprime if and only if their only common divisors are ± 1 . In accordance with Proposition 1.1, a, b are coprime precisely when

$$ax + by = 1$$

for some integers x, y. A couple of properties stand out for pairs of coprime integers.

Proposition 1.4. If a,b,c are integers and a,b are coprime and $a \mid bc$, then $a \mid c$.

Proof. We know that ax + by = 1 for some integers x, y. Then acx + bcy = c. Since $a \mid bc$ and evidently $a \mid ac$, it follows that $a \mid acx + bcy = c$.

Proposition 1.5. If a, b, c are integers and $a \mid c$ and $b \mid c$, then $ab \mid c$.

Proof. We know that ax + by = 1 for some integers x, y. Then acx + cby = c. Since $b \mid c$, it is clear that $ab \mid ac$, and likewise $ab \mid cb$ because $a \mid c$. So, $ab \mid acx + cby = c$.

EXERCISES

- 1. Use the Euclidean algorithm to find gcd(3150, 3003) and express this greatest common divisor as an integer combination of 3150 and 3003.
- 2. Using your favorite software, write a program to calculate the greatest common divisor of two integers by means of the Euclidean algorithm, and to express the greatest common divisor as an integer combination of the integers. By using your program, or otherwise, find gcd(2452548,2943234) and express this greatest common divisor as an integer combination of the given integers.
- **3.** These exercises can be done by using Proposition 1.1.
 - (a) If a, b are non-zero integers and k is any integer, show that gcd(ka, kb) = k gcd(a, b).
 - (b) If a, b are non-zero integers and k is a positive integer, show that gcd(a + kb, b) = gcd(a, b).
 - (c) If $a \mid bc$, show that $a \mid b \gcd(a, c)$.
 - (d) If a, b are coprime integers and $c \mid at$ and $c \mid bt$, show that $c \mid t$.
 - (e) If a, b, c are integers with a, c coprime, prove that gcd(ab, c) = gcd(b, c).
 - (f) If a, b are each coprime with c, show that ab is coprime with c.
- **4.** If a, b, c are integers with a, b non-zero, show that the equation ax + by = c has integer solutions x, y if and only if $gcd(a, b) \mid c$.

Find an integer solution x, y to the equation 91x + 55y = 12.

- 5. Show that 1 is the only complex number which satisfies both equations $x^{245} = 1$ and $x^{297} = 1$.
- **6.** Let a, b be positive integers. Let $g = \gcd(a, b)$ and $\ell = ab/g$.
 - (a) Explain very briefly why $a \mid \ell$ and $b \mid \ell$. Thus ℓ is a common multiple of a and b.



6 A Refresher on the Integers

(b) It turns out that ℓ is the *least common multiple* of a and b. Prove this claim by showing that if a positive integer m is a common multiple of a and b, then $\ell \mid m$.

Hint. To get from $a \mid m, b \mid m$ to $\ell \mid m$, it suffices to show that $ab \mid mg$. Use Proposition 1.1.

Typically ℓ is written as lcm(a,b), and we have the identity

$$ab = \gcd(a, b) \operatorname{lcm}(a, b).$$

- 7. One might wonder about the efficiency of the Euclidean algorithm. Suppose a, b are integers such that 0 < a < b, and that the Euclidean algorithm is applied to obtain gcd(a,b). By inspecting the algorithm one can see that each line of the algorithm consists of one Euclidean division, that there is a total of n+1 lines, and that gcd(a,b) appears on the nth line as r_n . We ask ourselves: how big could n get in terms of b?
 - (a) If 0 < a < b and b = aq + r for a quotient q and remainder r with $0 \le r < a$, show that $q \ge 1$, and then r < b/2.
 - (b) Suppose that $r_1, r_2, r_3, \ldots, r_n$ are the positive strictly decreasing remainders which appear in the Euclidean algorithm used to obtain r_n as gcd(a, b).

If n is even, explain how the following inequalities emerge:

$$r_2 < \frac{b}{2}, \quad r_4 < \frac{b}{2^2}, \quad r_6 < \frac{b}{2^3}, \dots, \quad r_n < \frac{b}{2^{n/2}}.$$

If n is odd, explain how the following inequalities emerge:

$$r_1 < \frac{b}{2}, \quad r_3 < \frac{b}{2^2}, \quad r_5 < \frac{b}{2^3}, \ldots, \quad r_n < \frac{b}{2^{(n+1)/2}} < \frac{b}{2^{n/2}}.$$

(c) If gcd(a, b) appears as r_n on the nth line of the Euclidean algorithm, use the fact $1 \le r_n$ to deduce that $2^{n/2} < b$, and then $n < 2\log_2(b)$.

Since line n + 1 with a zero remainder is needed to terminate the Euclidean algorithm, we learn that the number of Euclidean divisions used in the Euclidean algorithm to compute gcd(a,b) is at most

$$1 + 2\log_2(b)$$
.

(d) If a, b are positive integers of size at most 987654321234567, show that the Euclidean algorithm will compute gcd(a, b) in no more than 100 lines.

If a, b are positive integers of size at most 2^{500} , show that the Euclidean algorithm will compute gcd(a, b) in at most 1000 lines.

Despite such enormous possibilities, that is not a lot of lines for a computer program to execute.

8. Suppose that a, b, c are integers and that the only factors common to all three are ± 1 . Show that there exist integers x, y, z such that ax + by + cz = 1.

Generalize this to a statement about any number of integers a_1, a_2, \ldots, a_n .



1.2 Primes and Unique Factorization

7

1.2 Primes and Unique Factorization

Problems in \mathbb{Z} invariably boil down to problems about primes.

Definition 1.6. An integer p is called *prime* when $p \neq 0, 1, -1$ and the only divisors of p are ± 1 and $\pm p$. An integer n other than 0, 1, -1 and such that n has a divisor in addition to $\pm 1, \pm n$ is called **composite**.

Since p is prime whenever its negative is prime, attention is normally restricted to the positive primes, which, at a negligible loss of precision, tend to be called "prime" without the specification of positivity. Here are the primes up to 101:

With bigger integers it is no longer that easy to pick out primes. For example, a naive guess might be that 91 is prime, but 91 = 13.7.

Although we might feel confident that there are infinitely many primes, this is not obvious. But first comes a result which points to the importance of primes.

Proposition 1.7. If n is an integer and $n \ge 2$, then n can be factored into primes.

Proof. Use induction on *n*.

If n = 2, then n is factored as itself into primes. Suppose $2, 3, 4, \ldots, n-1$ can each be factored into primes. Now look at n. If n is prime, then n is a product of primes, namely itself. If n is composite, write $n = k\ell$, where 1 < k < n and $1 < \ell < n$. Since k, ℓ are among the integers $2, 3, 4, \ldots, n-1$, each of them factors into primes. That is,

$$k = p_1 \cdot p_2 \cdots p_r$$
 and $\ell = q_1 \cdot q_2 \cdots q_s$, where the p_i , q_i are primes.

Then of course, the equations

$$n = k\ell = p_1 \cdots p_r \cdot q_1 \cdots q_s$$

give a factorization of n into primes.

From the above comes a famous theorem already in Euclid's books.

Proposition 1.8. *There are infinitely many primes.*

Proof. Given any *finite* list of primes $p_1, p_2, ..., p_n$, here is how to come up with one more prime not on the list. Let

$$n = p_1 \cdot p_2 \cdot \cdot \cdot p_n + 1.$$

According to Proposition 1.7, n has a prime factor q. This q cannot be equal to any p_1, \ldots, p_n . Indeed, if q equalled some p_j , then

q would divide $n - p_1 \cdot p_2 \cdots p_n$, which equals 1.



8 A Refresher on the Integers

Since no prime is a factor of 1, our q is a prime not equal to any of p_1, p_2, \ldots, p_n . This permits the build-up of new primes at will.

Unique Factorization

The special thing about primes is that there is *only one way* to factor an integer into primes. Ambiguous factorings such as

$$24 = 6 \cdot 4 = 8 \cdot 3 = 1 \cdot 1 \cdot 12 \cdot 2$$

do not occur when only primes are involved in the factors. Despite one's sense that such unique factorization must be true, this is not obvious.

It is worth noting that

an integer a is coprime with a prime p if and only if $p \nmid a$.

Indeed, if $p \mid a$, then $gcd(p, a) = p \neq 1$. Conversely, if d = gcd(p, a) and $d \neq 1$, then the fact $d \mid p$ forces d = p, and then $d \mid a$.

The next result forms the cornerstone for the proof of unique factorization.

Proposition 1.9. An integer p, other than $0, \pm 1$, is prime if and only if it has the property that whenever p divides a product ab, then p already divides a or b.

Proof. Say p is prime and $p \mid ab$, and suppose $p \nmid a$. As observed, p and a are coprime. By Proposition 1.4 it follows that $p \mid b$.

For the converse, suppose that p divides a factor whenever p divides the product of two integers. Now let a be a factor of p. Thus p = ab for some other integer b. Clearly $p \mid ab$, and thus $p \mid a$ or $p \mid b$. If $p \mid a$, then the fact $a \mid p$ yields $a = \pm p$. If $p \mid b$, write b = qp for some integer q. Then p = ab = aqp. Cancel p to get aq = 1, and from that $a = \pm 1$. So, the only possible factors of p are $\pm p$ and ± 1 .

A strong case can be made that the alternative property in Proposition 1.9 ought to be taken as the *definition* of a prime number. Indeed, this is what makes the proof of unique factorization work.

Proposition 1.9 readily extends to the product of several integers. Namely, if a prime p divides the product $a_1a_2a_3\cdots a_k$, then p already divides at least one of the a_j . The proof to follow, that factorization of integers into primes is unique, rests on the shoulders of Proposition 1.9.

Proposition 1.10. If $p_1, p_2, ..., p_n$ and $q_1, q_2, ..., q_m$ are two lists of primes (positive) with repetitions allowed, and if

$$p_1 \cdot p_2 \cdots p_n = q_1 \cdot q_2 \cdots q_m$$

then, after a rearrangement of the q_i , the lists are identical. That is

$$m = n$$
 and $p_1 = q_1, p_2 = q_2, ..., p_n = q_n$.



1.2 Primes and Unique Factorization

Proof. Clearly $p_1 | q_1 \cdot q_2 \cdots q_m$. By Proposition 1.9, p_1 divides some q_j . Rearrange the q_j , and say $p_1 | q_1$. Since q_1 is prime, $p_1 = 1$ or $p_1 = q_1$. The first option cannot hold because p_1 is

9

$$p_1 \cdot p_2 \cdot \cdot \cdot p_n = p_1 \cdot q_2 \cdot \cdot \cdot q_m$$
.

Cancel p_1 to get

prime. Thus $p_1 = q_1$, and then

$$p_2 \cdot p_3 \cdot \cdot \cdot p_n = q_2 \cdot q_3 \cdot \cdot \cdot q_m$$
.

Repeat the argument, with the necessary rearrangement of the q_i , to get $p_2 = q_2$, and then

$$p_3 \cdot p_4 \cdot \cdot \cdot p_n = q_3 \cdot q_4 \cdot \cdot \cdot q_m$$
.

Continuing in this fashion, after suitable rearrangement of the q_j , we end up with one of the following possibilities:

- n < m, and $p_1 = q_1, \ldots, p_n = q_n, 1 = q_{m-n} \cdots q_m$
- m < n, and $p_1 = q_1, \ldots, p_m = q_m, p_{n-m} \cdots p_n = 1$
- m = n, and $p_1 = q_1, \ldots, p_n = q_n$.

The first two situations cannot happen, since only ± 1 can be factors of 1 and the p_j , q_j are not ± 1 . So indeed, m = n and all $p_j = q_j$, after some rearranging of the q_j .

The Multiplicity of a Prime inside an Integer

It is customary to collect repeated primes in the unique factorization of a positive integer a as follows:

$$a=p_1^{e_1}\cdot p_2^{e_2}\cdots p_n^{e_n},$$

where p_i are now distinct primes and the exponents $e_i \ge 1$.

Definition 1.11. The unique number of times e_j in which a prime appears in the unique factorization of a non-zero integer a is called the *multiplicity* of p_j inside a. It is convenient to allow multiplicities to equal 0. A prime p has multiplicity 0 in a when $p \nmid a$.

For example, with

$$a = 2^3 \cdot 3^0 \cdot 5^1 \cdot 11^8 \cdot 29^0$$
.

the primes 3 and 29 have multiplicity 0 in a, while the multiplicity of 11 in a is 8.

Proposition 1.10 is saying that for every positive integer a and every prime p, there is a unique non-negative multiplicity of p in a. Denote this multiplicity by $m_a(p)$. For example, $m_{320}(2) = 6$ because $320 = 2^6 \cdot 5^1$, while $m_{320}(7) = 0$.

For each positive a, the function $p \mapsto m_a(p)$ from the set of primes to the set of non-negative integers counts the number of times that each prime p appears in the unique factorization of a.

© in this web service Cambridge University Press

www.cambridge.org



10 A Refresher on the Integers

Here are a few observations about the multiplicity function. If a, b are positive integers, then

- $m_{ab}(p) = m_a(p) + m_b(p)$,
- $m_a(p) > 0$ for at most finitely primes p,
- $m_a(p) = m_b(p)$ for all primes p if and only if a = b,
- $m_a(p) = 0$ for all primes p if and only if a = 1.
- If m is a function defined on the set of primes with values in the set of non-negative integers, and if m(p) > 0 for at most a finite number of primes p, then there is a unique positive integer a such that $m(p) = m_a(p)$ for all primes p. Indeed, if p_1, p_2, \ldots, p_n are the distinct primes at which $m(p_i) > 0$, the required a is given by

$$a = p_1^{m(p_1)} p_2^{m(p_2)} p_3^{m(p_3)} \cdots p_n^{m(p_n)}.$$

Divisibility and Unique Factorization

The notion of divisibility fits nicely into the language of multiplicities.

Proposition 1.12. A positive integer a divides another positive integer b if and only if $m_a(p) \le m_b(p)$ for every prime p.

Proof. Suppose $a \mid b$. That is, b = ac for some c. Then

$$m_b(p) = m_{ac}(p) = m_a(p) + m_c(p)$$
 for all primes p.

Clearly $m_a(p) \le m_b(p)$, because $m_c(p) \ge 0$.

Conversely, suppose $m_a(p) \le m_b(p)$ for all primes p. For each such prime, let $m(p) = m_b(p) - m_a(p)$. Since $m(p) \ge 0$ for all primes p and m(p) > 0 for at most a finite number of primes, let c be the unique positive integer that has m as its multiplicity function, i.e. $m(p) = m_c(p)$. Then

$$m_{ac}(p) = m_a(p) + m_c(p) = m_a(p) + m(p) = m_a(p) + m_b(p) - m_a(p) = m_b(p)$$

for all primes p. It follows that ac = b, in other words $a \mid b$.

Greatest Common Divisors in Terms of Multiplicity Functions

The greatest common divisor of two positive integers can be expressed in terms of the multiplicity function that arises from unique factorization.

Proposition 1.13. If a,b are positive integers, then gcd(a,b) is the positive integer c whose multiplicity function is given by

$$m_c(p) = \min(m_a(p), m_b(p))$$
 for all primes p.