First Arts Lecture Notes 2004-2005

Chapter 2. Introduction to Number Theory

In this chapter, we intend to develop some simple properties of numbers that lead to the study of more complicated structures in abstract algebra. We are concerned throughout this chapter with integers, that is to say, whole numbers.

2.1 Definition An integer c is said to divide an integer b if there is some integer d with $b = cd$. We also say in this case that b is *divisible* by c.

Note that the only integer divisors of 1 and -1 are ± 1. This is an important fact. Furthermore, suppose that c divides both a and b. Then we claim that c divides $ra + sb$ for any integers r and s. For, we can write

$$a = cd, \quad b = ce$$

for certain integers d and e and then

$$ra + sb = rcd + sce = c(rd + se).$$

This shows that c divides exactly into $ra + sb$.

We now state a slight variation on the principle of mathematical induction, known as the *principle of complete induction.*

(2.2) Axiom (Principle of complete induction) Let $P(n)$ be a mathematical proposition made for each integer $n \geq 1$. Suppose that $P(1)$ is true. Suppose also that for any given value r of n, the assumption that all of $P(1)$, $P(2)$, $\ldots$, $P(r)$ are true implies that $P(r + 1)$ is true. Then $P(n)$ is true for all values of n.

We use this principle of complete induction to prove a well known fact used in arithmetic.

(2.3) Theorem (Division algorithm) Let a and b be positive integers. Then there exist unique integers $q \geq 0$ and r with

$$a = bq + r,$$

where $0 \leq r < b$. We call q the *quotient* and r the *remainder* in the division of a by b.

Proof Using the principle of complete induction, we assume that the result is true for all positive integers less than a, and then try to prove the result is true for a. Suppose first that $a < b$. Then we just take $q = 0$, and $r = a$. This is the only possible choice. If $a = b$, we take $q = 1$,

$r = 0$, and again this is the only possible choice. Finally, suppose that $a > b$. Then $a - b > 0$. By induction, since $0 < a - b < a$, we can write

$$a - b = sb + t,$$

where s and t are integers with $s \geq 0$ and $0 \leq t < b$. Then by addition,

$$a - b + b = a = (s + 1)b + t$$

and we just take $q = s + 1$, and $r = t$. This proves that we can find q and r. Their uniqueness follows from the uniqueness of s and t as they arise for $a - b$.

(2.4) Definition We say that a non-empty subset S of integers is a *subgroup* of **Z** if 0 is in S and whenever a and b are in S, the difference $a - b$ is also in S.

If the subgroup contains only 0, we say it is the zero subgroup.

(2.5) Proposition Let S be a non-zero subgroup of integers. Then if c is any number in S, and t is any integer, the integer multiple tc of c is also in S.

Proof We first show that $-c$ is also in S. We know that S contains 0, and as S is a subgroup, it also contains $0 - c = -c$. We prove by induction on n that S contains nc for any positive integer n. This is certainly true when $n = 1$. Suppose then that S contains kc, where k is an integer ≥ 1. As S therefore contains kc and $-c$, it contains

$$kc - (-c) = (k + 1)c,$$

as required. This proves that S contains nc for all positive integers n. But S contains the negative of each of its members and so $-(nc) = (-n)c$ is also in S. Since nc and $(-n)c$ are in S for any positive integer n, our proposition follows. ■

We would like now to give a description of all non-zero subgroups of **Z**. To do this, we must invoke another mathematical principle, known as the well ordering principle.

Axiom (Well ordering principle) Let S be a non-empty subset of non-negative integers. Then S contains a smallest member, m, say.

Note that if m is the smallest member of S, we have $m \leq x$ for all integers x in S.

(2.6) Theorem Let S be a non-zero subgroup of integers. Then S contains a smallest positive integer and consists of all possible integer multiples of this smallest positive integer.

Proof As S is not the zero subgroup, it contains a number, c, say, different from 0 and hence also $-c$, by Proposition 2.5. Now exactly one of c and $-c$ is positive, and as each of these numbers is in S, we see that S contains positive numbers. Let T be the set of positive numbers in S. We know now that T is non-empty and thus by the well-ordering principle, it contains a smallest positive integer, m, say. We will show that S consists of all integer multiples of m. Note that by Proposition 2.5, S certainly contains all integer multiples of m.

Let b be any integer in S. By Theorem 2.3 (the division algorithm), there exist integers q and r with $0 \leq r < m$ and

$$b = qm + r.$$

Now we know by Proposition 2.5 that qm is in S, as it is an integer multiple of m. Thus, as S is a subgroup, containing b and qm,

$$b - qm = r$$

is in S. Thus r is in S. As $0 \leq r < m$, the definition of m as the smallest positive integer in S implies that r must be 0. Hence $b = qm$ is an integer multiple of m, as required. ■

Given a positive integer m, let $m\mathbf{Z}$ denote the set of all integer multiples of m. So,

$$m\mathbf{Z} = \{\, 0, \pm m, \pm 2m, \pm 3m, \dots \}.$$

It is easy to check that $m\mathbf{Z}$ is a subgroup of $\mathbf{Z}$, whose smallest positive member is m. Conversely, Theorem 2.6 shows that all subgroups of $\mathbf{Z}$ have this form.

(2.7) Definition Let b and c be any two integers. A positive integer d is called the *greatest common divisor* (gcd) of b and c if d divides both b and c and is the largest integer with this property. We often write $\gcd(b, c)$ for the gcd of b and c.

We first look at the concept of gcd from a theoretical point of view and then proceed to a very efficient method for finding the gcd.

(2.8) Theorem Let b and c be any two non-zero integers and let d be their gcd. Then there exist integers s and t with

$$d = sb + tc.$$

Proof Consider the subset S of $\mathbf{Z}$ consisting of all integers of the form

$$xb + yc,$$

where x and y run independently through all the integers. We claim that S is a subgroup of $\mathbf{Z}$. To prove this, take two elements $xb + yc$ and $x'b + y'c$ in S. Then the difference

$$xb + yc - (x'b + y'c) = (x - x')b + (y - y')c = x''b + y''c,$$

where $x'' = x - x'$, $y'' = y - y'$, is certainly an element of S. Thus, S is a subgroup of $\mathbf{Z}$. It is certainly non-zero as it contains

$$b = 1b + 0c \text{ and } c = 0b + 1c.$$

By Theorem 2.6, $S = d\mathbf{Z}$ for some positive integer d. Thus, since d is in S,

$$d = sb + tc$$

for certain integers s and t. Moreover, as both b and c are in S,

$$b = du, \quad c = dv$$

for certain integers u and v. This shows that d is a common divisor of b and c. Furthermore, if e is a common divisor of b and c, e then divides sb and tc for any s and t. Hence e divides $sb + tc = d$. Thus e is a divisor of d and it follows that d must be the gcd of b and c. ■

Note that if s and t have been found as above, then for any integer value of r, the numbers

$$s' = s + rc, \quad t' = t - rb$$

have the property that $d = s'b + t'c$, so that there are infinitely many possibilities for s and t.

The following result follows from the proof above. It is not obvious from the definition of gcd that such a result should hold.

(2.9) Corollary Let b and c be any two non-zero integers. Then any common divisor of b and c divides their gcd.

Now Theorem 2.8 is an existence theorem and it gives no indication of how to find the gcd or of how to find the two integers s and t. Luckily a practical procedure has existed since the time of Euclid. It uses a repeated process of division and formation of remainder, as in Theorem 2.3.

We may suppose that our integers b and c are both positive. Dividing c into b by the division algorithm gives

$$b = cq_1 + r_1,$$

where $0 \leq r_1 < c$. Any common divisor of b and c divides r_1, as we see from the equation above. Thus $\gcd(b, c)$ is certainly a divisor of both r_1 and c, which implies that $\gcd(b, c) \leq \gcd(c, r_1)$. Likewise, any common divisor of c and r_1 divides b and thus $\gcd(c, r_1) \leq \gcd(b, c)$. These two inequalities imply that

$$\gcd(b, c) = \gcd(c, r_1).$$

This is important, for it shows that we can replace the problem of finding $\gcd(b, c)$ by the problem of finding $\gcd(c, r_1)$, which is easier, as $c \leq b$ and $r_1 < c$. We continue the division process, dividing r_1 into c:

$$c = r_1 q_2 + r_2,$$

where $0 \leq r_2 < r_1$. As above,

$$\gcd(c, r_1) = \gcd(r_1, r_2) = \gcd(b, c).$$

Continue, dividing the new remainder into the previous remainder.

$$\begin{array}{ccc} r_1 & = & r_2 q_3 + r_3 \\ \vdots & \vdots & \vdots \\ r_{n-2} & = & r_{n-1} q_n + r_n \\ r_{n-1} & = & r_n q_{n+1} \end{array}$$

Eventually, the remainder must become 0, say when r_{n-1} is divided by r_n, since the non-negative remainders get smaller at each stage and cannot continue decreasing indefinitely. The argument given above shows that

$$\gcd(b, c) = \gcd(c, r_1) = \gcd(r_1, r_2) = \ldots = \gcd(r_{n-1}, r_n).$$

However, the equation $r_{n-1} = r_n q_n$ shows that r_n divides r_{n-1} and so $\gcd(r_{n-1}, r_n) = r_n$. Thus the gcd of b and c is r_n, which is the last non-zero remainder in the repeated division process.

To find the integers s and t so that $d = sb + tc$ is often rather laborious. It can be done working from the botttom or from the top. To explain working from the bottom, recall that r_n, the last non-zero remainder is the gcd. Now

$$r_n = r_{n-2} - r_{n-1} q_n$$

expresses r_n in terms of previous remainders. Next,

$$r_{n-1} = r_{n-3} - r_{n-2} q_{n-1}$$

and substituting this value into the previous equation, we get

$$r_n = r_{n-2} - (r_{n-3} - r_{n-2} q_{n-1}) q_n = r_{n-2}(1 + q_{n-1} q_n) - q_n r_{n-3}.$$

Then we substitute for r_{n-2} and obtain an expression for r_n in terms of r_{n-3} and r_{n-4}. Continuing in this way, we express the gcd r_n in terms of two successive remainders and eventually work up to b and c.

For computational purposes, it seems to be easier to work from the top. Our first three divisions give

$$\begin{aligned} r_1 &= b - cq_1 \\ r_2 &= c - r_1q_2 \\ r_3 &= r_1 - r_2q_3 \end{aligned}$$

and hence substituting

$$\begin{aligned} r_2 &= c - (b - cq_1)q_2 \\ r_2 &= -bq_2 + c(1 + q_1q_2) \\ r_3 &= b - cq_1 + bq_2q_3 - c(q_3 + q_1q_2q_3) \\ r_3 &= b(1 + q_2q_3) - c(q_1 + q_3 + q_1q_2q_3). \end{aligned}$$

This process may be continued until r_n is reached. When a calculator is used, only a small number of memory locations are needed to keep the various quotients and remainders.

Let b and c be non-zero integers and let d be their gcd. Let

$$b' = \frac{b}{d}, \quad c' = \frac{c}{d}.$$

It is easy to see that the gcd of b' and c' must be 1. For, if the gcd is e, then de is a common divisor of b and c and hence must be d. Numbers whose gcd is 1 are said to be *relatively prime.*

We want now to make use of Theorem 2.8 to investigate properties of relatively prime integers.

(2.10) Theorem Let b and c be relatively prime integers. Suppose that b divides a product ce, where e is some integer. Then b divides e.

Proof As b and c are relatively prime, their gcd is 1 and thus by Theorem 2.8, there exist integers s and t with

$$1 = sb + tc.$$

Multiplying this equation by e, we obtain $e = sbe + tce$. By hypothesis, b divides ce and hence tce. But b also divides sbe and thus divides $sbe + tce = e$, which gives us what we want. ■

(2.11) Corollary Let b and c be relatively prime integers. Suppose that b divides a product c^ne, where n is a positive integer and e is some integer. Then b divides e.

Proof We see that b divides $c(c^{n-1}e)$ and hence by Theorem 2.10, b divides $c^{n-1}e$. Then we deduce that b divides $c^{n-2}e$. Continuing in this way, we eventually arrive at the fact that b divides e, as required (this proof may be set up properly by induction). ■

It has been well known since the time of Pythagoras's Theorem that $\sqrt{2}$ is not a rational number. This means that we cannot find integers b and c so that

$$\left(\frac{b}{c}\right)^2 = 2.$$

This fact shows that there are numbers naturally occurring in mathematics that are not simple fractions (or equivalently, eventually repeating decimals). It is possible to generalize the proof of this result, by making use of Corollary 2.11.

(2.12) Theorem Let d be an integer and let n be a positive integer. Suppose that d is the n-th power of a rational number. Then d is the n-th power of an integer.

Proof Suppose that d is the n-th power of the rational number $\frac{b}{c}$, where b and c are integers. By removing common factors of b and c (so that $\frac{b}{c}$ is written in its lowest terms), we may assume that b and c are relatively prime and that $c > 0$. We claim now that in this case c must be 1. For we have

$$d = \left(\frac{b}{c}\right)^n \text{ implying that } b^n = c^n d.$$

This shows that c certainly divides $b^n = b^n 1$. Hence c divides 1, by Corollary 2.11. As $c > 0$, this forces c to equal 1, as required and thus $d = b^n$ is the n-th power of an integer. ■

This result implies that the n-th roots of integers are irrational unless the integer is itself the n-th power of an integer.

Example The n-th root of 2 is irrational for $n \geq 2$. For if there is a rational number whose n-th power is 2, by what we have proved, there is also a positive integer with this property. This is impossible as the n-th power of a positive integer is either 1 or is at least 4.

Next, we move on to consider properties of primes and the factorization of integers into the product of primes.

(2.13) Definition An integer $p > 1$ is said to be a *prime* if its only integer divisors are $\pm p$ and ± 1.

Notice that 2 is the only even prime. The prime numbers are difficult to enumerate systematically, but the first few are 2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43.

Now let p be a prime integer. Then the only positive integers that divide p are 1 and p. Consequently, if b is any non-zero integer, the gcd of b and p is either 1, in which case b and p are relatively prime, or the gcd is p, in which case p divides b. We can now obtain a basic property of primes.

(2.14) Theorem Let p be a prime integer. Suppose that p divides a product bc, where b and c are integers. Then p divides b or p divides c.

Proof If p divides b, we are finished. Otherwise, if p does not divide b, b and p are relatively prime and then it follows from Theorem 2.10 that p divides c. ■

A straightforward extension of the argument above shows that if a prime p divides a product $b_1 \cdots b_n$ of n integers, it divides at least one of the integers b_i.

We are now in a position to prove what may be called the fundamental theorem of arithmetic. This is the statement that any integer may be factorized into a product of prime numbers and this factorization is essentially unique. While this fact is probably known intuitively to most educated people, it was not until the work of Gauss in 1801 that a convincing proof of the theorem was first given.

(2.15) Theorem Let $b > 1$ be an integer. Then b may be expressed as the product of prime integers. Furthermore, if we write $b = p_1 p_2 \cdots p_t$, where $p_1, \ldots, p_t$ are primes satisfying $p_1 \leq p_2 \leq \ldots \leq p_t$, this decomposition is unique.

Proof We first prove by induction on the size of b that b is a product of primes. If b is itself a prime, we are finished. If b is not a prime, it may be written as a product

$$b = cd,$$

where c and d are integers satisfying $1 < c < b$, $1 < d < b$. By induction, c and d are products of primes, say,

$$c = p_1 \cdots p_s, \quad d = q_1 \cdots q_t,$$

and then

$$b = p_1 \cdots p_s q_1 \cdots q_t$$

is expressed as a product of primes.

The harder part of the proof is to show the uniqueness of the factorization. Suppose that we have two prime factorizations

$$b = r_1 \cdots r_m = s_1 \cdots s_n,$$

where each r_i and s_j is a prime and we may assume that $r_1 \leq \ldots \leq r_m$ and $s_1 \leq \ldots \leq s_n$. Then the prime r_1 divides the product $s_1 \cdots s_n$ and by our remarks following the proof of Theorem 2.14, we see that r_1 divides at least one of $s_1, \ldots, s_n$, say, s_j. Since s_j is a prime and r_1 divides it, $r_1 = s_j$. This shows in particular that $r_1 \geq s_1$. On the other hand, repeating the argument with s_1 in place of r_1, we must obtain $s_1 \geq r_1$. These two inequalities imply that $r_1 = s_1$. Cancelling r_1 and s_1 from each side of the equation for b, we get

$$r_2 \cdots r_m = s_2 \cdots s_n.$$

We continue this process, showing that $r_2 = s_2$, and so on, as required. ■

We can use the prime factorization of integers to investigate the gcd of integers in a theoretical manner. Let b and c be positive integers. Write the prime factorizations of b and c in the form

$$b = p_1^{r_1} \ldots p_m^{r_m}, \quad c = p_1^{s_1} \ldots p_m^{s_m},$$

where $p_1, \ldots, p_m$ are different primes and at least one of the indices r_i, s_i is positive for $i = 1, \ldots, m$ (but we allow the possibility that some indices are 0). Then set t_i equal to the minimum of r_i and s_i for $i = 1, \ldots, m$. It is quite straightforward to see that

$$p_1^{t_1} \ldots p_m^{t_m}$$

is the greatest common divisor of b and c.

Example The gcd of $2^3 3^2 7^5 13$ and $2^2 3^3 7^2 11$ is easily seen to be $2^2 3^2 7^2$.

The Euclidean algorithm method for finding the gcd is much more efficient in practice, since there is in general no easy way to obtain the prime factorization of an integer.