Chapter 5

Divisibility and the Greatest Common Divisor

As we have already seen in our study of Pythagorean triples, the notions of divisibility and factorizations are important tools in number theory. In this chapter we will look at these ideas more closely.

Suppose that m and n are integers with $m \neq 0$. We say that m *divides* n if n is a multiple of m, that is, if there is an integer k such that $n = mk$. If m divides n, we write $m|n$. Similarly, if m does not divide n, then we write $m \nmid n$. For example,

$$3|6 \quad \text{and} \quad 12|132, \quad \text{since} \quad 6 = 3 \cdot 2 \quad \text{and} \quad 132 = 12 \cdot 11.$$

The divisors of 6 are 1, 2, 3, and 6. On the other hand, $5 \nmid 7$, since no integer multiple of 5 is equal to 7. A number that divides n is called a *divisor of* n.

If we are given two numbers, we can look for common divisors, that is, numbers that divide both of them. For example, 4 is a common divisor of 12 and 20, since $4|12$ and $4|20$. Notice that 4 is the largest common divisor of 12 and 20. Similarly, 3 is a common divisor of 18 and 30, but it is not the largest, since 6 is also a common divisor. The largest common divisor of two numbers is an extremely important quantity that will frequently appear during our number theoretic excursions.

> The *greatest common divisor* of two numbers a and b (not both zero) is the largest number that divides both of them. It is denoted by $\gcd(a, b)$. If $\gcd(a, b) = 1$, we say that a and b are *relatively prime*.

Two examples that we mentioned above are

$$\gcd(12, 20) = 4 \qquad \text{and} \qquad \gcd(18, 30) = 6.$$

Another example is

$$\gcd(225, 120) = 15.$$

We can check that this answer is correct by factoring $225 = 3^2 \cdot 5^2$ and $120 = 2^3 \cdot 3 \cdot 5$, but, in general, factoring a and b is not an efficient way to compute their greatest common divisor.[1]

The most efficient method known for finding the greatest common divisors of two numbers is called the *Euclidean algorithm*. It consists of doing a sequence of divisions with remainder until the remainder is zero. We will illustrate with two examples before describing the general method.

As our first example, we will compute $\gcd(36, 132)$. The first step is to divide 132 by 36, which gives a quotient of 3 and a remainder of 24. We write this as

$$132 = 3 \times 36 + 24.$$

The next step is to take 36 and divide it by the remainder 24 from the previous step. This gives

$$36 = 1 \times 24 + 12.$$

Next we divide 24 by 12, and we find a remainder of 0,

$$24 = 2 \times 12 + 0.$$

The Euclidean algorithm says that as soon as you get a remainder of 0, the remainder from the previous step is the greatest common divisor of the original two numbers. So in this case we find that $\gcd(132, 36) = 12$.

Let's do a larger example. We will compute

$$\gcd(1160718174, 316258250).$$

Our reason for doing a large example like this is to help convince you that the Euclidean algorithm gives a far more efficient way to compute gcd's than factorization. We begin by dividing 1160718174 by 316258250, which gives 3 with a remainder of 211943424. Next we take 316258250 and divide it by 211943424. This process continues until we get a remainder of 0. The calculations are given in

[1] An even less efficient way to compute the greatest common divisor of a and b is the method taught to my daughter by her fourth grade teacher, who recommended that the students make complete lists of all the divisors of a and b and then pick out the largest number that appears on both lists!

the following table:

$$\begin{array}{rcrcr}
1160718174 &=& 3 \times 316258250 &+& 211943424 \\
316258250 &=& 1 \times 211943424 &+& 104314826 \\
211943424 &=& 2 \times 104314826 &+& 3313772 \\
104314826 &=& 31 \times 3313772 &+& 1587894 \\
3313772 &=& 2 \times 1587894 &+& 137984 \\
1587894 &=& 11 \times 137984 &+& 70070 \\
137984 &=& 1 \times 70070 &+& 67914 \\
70070 &=& 1 \times 67914 &+& 2156 \\
67914 &=& 31 \times 2156 &+& \boxed{1078} \leftarrow \text{gcd} \\
2156 &=& 2 \times 1078 &+& 0
\end{array}$$

Notice how at each step we divide a number A by a number B to get a quotient Q and a remainder R. In other words,

$$A = Q \times B + R.$$

Then at the next step we replace our old A and B with the numbers B and R and continue the process until we get a remainder of 0. At that point, the remainder R from the previous step is the greatest common divisor of our original two numbers. So the above calculation shows that

$$\gcd(1160718174, 316258250) = 1078.$$

We can partly check our calculation (always a good idea) by verifying that 1078 is indeed a common divisor. Thus

$$1160718174 = 1078 \times 1076733 \quad \text{and} \quad 316258250 = 1078 \times 293375.$$

There is one more practical matter to be mentioned before we undertake a theoretical analysis of the Euclidean algorithm. If we are given A and B, how can we find the quotient Q and the remainder R? Of course, you can always use long division, but that can be time consuming and subject to arithmetic errors if A and B are large. A pleasant alternative is to find a calculator or computer program that will automatically compute Q and R for you. However, even if you are only equipped with an inexpensive calculator, there is an easy three-step method to find Q and R.

Method to Compute Q and R on a Calculator So That $A = B \times Q + R$

1. Use the calculator to divide A by B. You get a number with decimals.
2. Discard all the digits to the right of the decimal point. This gives Q.
3. To find R, use the formula $R = A - B \times Q$.

For example, suppose that $A = 12345$ and $B = 417$. Then $A/B = 29.6043\ldots$, so $Q = 29$ and $R = 12345 - 417 \cdot 29 = 252$.

We're now ready to analyze the Euclidean algorithm. The general method looks like

$$\begin{array}{rccl}
a = & q_1 \times b & + & r_1 \\
b = & q_2 \times r_1 & + & r_2 \\
r_1 = & q_3 \times r_2 & + & r_3 \\
r_2 = & q_4 \times r_3 & + & r_4 \\
 & \vdots & & \\
r_{n-3} = & q_{n-1} \times r_{n-2} & + & r_{n-1} \\
r_{n-2} = & q_n \times r_{n-1} & + & \boxed{r_n} \leftarrow \text{gcd} \\
r_{n-1} = & q_{n+1} r_n & + & 0
\end{array}$$

If we let $r_0 = b$ and $r_{-1} = a$, then every line looks like

$$r_{i-1} = q_{i+1} \times r_i + r_{i+1}.$$

Why is the last nonzero remainder r_n a common divisor of a and b? We start from the bottom and work our way up. The last line $r_{n-1} = q_{n+1} r_n$ shows that r_n divides r_{n-1}. Then the previous line

$$r_{n-2} = q_n \times r_{n-1} + r_n$$

shows that r_n divides r_{n-2}, since it divides both r_{n-1} and r_n. Now looking at the line above that, we already know that r_n divides both r_{n-1} and r_{n-2}, so we find that r_n also divides r_{n-3}. Moving up line by line, when we reach the second line we will already know that r_n divides r_2 and r_1. Then the second line $b = q_2 \times r_1 + r_2$ tells us that r_n divides b. Finally, we move up to the top line and use the fact that r_n divides both r_1 and b to conclude that r_n also divides a. This completes our verification that the last nonzero remainder r_n is a common divisor of a and b.

But why is r_n the *greatest* common divisor of a and b? Suppose that d is any common divisor of a and b. We will work our way back down the list of equations. So from the first equation $a = q_1 \times b + r_1$ and the fact that d divides both a and b, we see that d also divides r_1. Then the second equation $b = q_2 r_1 + r_2$ shows us that d must divide r_2. Continuing down line by line, at each stage we will know that d divides the previous two remainders r_{i-1} and r_i, and then the current line $r_{i-1} = q_{i+1} \times r_i + r_{i+1}$ will tell us that d also divides the next remainder r_{i+1}. Eventually, we reach the penultimate line $r_{n-2} = q_n \times r_{n-1} + r_n$, at which point we conclude that d divides r_n. So we have shown that if d is any common divisor of a and b then d will divide r_n. Therefore, r_n must be the greatest common divisor of a and b.

This completes our verification that the Euclidean algorithm actually computes the greatest common divisor, a fact of sufficient importance to be officially recorded.

Theorem 5.1 (Euclidean Algorithm)**.** *To compute the greatest common divisor of two numbers a and b, let $r_{-1} = a$, let $r_0 = b$, and compute successive quotients and remainders*

$$r_{i-1} = q_{i+1} \times r_i + r_{i+1}$$

for $i = 0, 1, 2, \ldots$ until some remainder r_{n+1} is 0. The last nonzero remainder r_n is then the greatest common divisor of a and b.

There remains the question of why the Euclidean algorithm always finishes. In other words, we know that the last nonzero remainder will be the desired $\gcd$, but how do we know that we ever get a remainder that does equal 0? This is not a silly question, since it is easy to give algorithms that do not terminate; and there are even very simple algorithms for which it is not known whether or not they always terminate. Fortunately, it is easy to see that the Euclidean algorithm always terminates. The reason is simple. Each time we compute a quotient with remainder,

$$A = Q \times B + R,$$

the remainder will be between 0 and $B - 1$. This is clear, since if $R \geq B$, then we can add one more onto the quotient Q and subtract B from R. So the successive remainders in the Euclidean algorithm continually decrease:

$$b = r_0 > r_1 > r_2 > r_3 > \cdots .$$

But all the remainders are greater than or equal to 0, so we have a strictly decreasing sequence of nonnegative integers. Eventually, we must reach a remainder that equals 0; in fact, it is clear that we will reach a remainder of 0 in at most b steps. Fortunately, the Euclidean algorithm is far more efficient than this. You will show in the exercises that the number of steps in the Euclidean algorithm is at most seven times the *number of digits* in b. So, on a computer, it is quite feasible to compute $\gcd(a, b)$ when a and b have hundreds or even thousands of digits!

Exercises

5.1. Use the Euclidean algorithm to compute each of the following gcd's.
(a) $\gcd(12345, 67890)$ **(b)** $\gcd(54321, 9876)$

5.2. 💻 Write a program to compute the greatest common divisor $\gcd(a, b)$ of two integers a and b. Your program should work even if one of a or b is zero. Make sure that you don't go into an infinite loop if a and b are both zero!

Chapter 6

Linear Equations and the Greatest Common Divisor

Given two whole numbers a and b, we are going to look at all the possible numbers we can get by adding a multiple of a to a multiple of b. In other words, we will consider all numbers obtained from the formula

$$ax + by$$

when we substitute all possible integers for x and y. Note that we are going to allow both positive and negative values for x and y. For example, we could take $a = 42$ and $b = 30$. Some of the values of $ax + by$ for this a and b are given in the following table:

	$x=-3$	$x=-2$	$x=-1$	$x=0$	$x=1$	$x=2$	$x=3$
$y=-3$	-216	-174	-132	-90	-48	-6	36
$y=-2$	-186	-144	-102	-60	-18	24	66
$y=-1$	-156	-114	-72	-30	12	54	96
$y=\ \ 0$	-126	-84	-42	0	42	84	126
$y=\ \ 1$	-96	-54	-12	30	72	114	156
$y=\ \ 2$	-66	-24	18	60	102	144	186
$y=\ \ 3$	-36	6	48	90	132	174	216

Table of Values of $42x + 30y$

Our first observation is that every entry in the table is divisible by 6. This is not surprising, since both 42 and 30 are divisible by 6, so every number of the form $42x + 30y = 6(7x + 5y)$ is a multiple of 6. More generally, it is clear that every number of the form $ax + by$ is divisible by $\gcd(a, b)$, since both a and b are divisible by $\gcd(a, b)$.

A second observation, which is somewhat more surprising, is that the greatest common divisor of 42 and 30, which is 6, actually appears in our table. Thus from the table we see that

$$42 \cdot (-2) + 30 \cdot 3 = 6 = \gcd(42, 30).$$

Further examples suggest the following conclusion:

$$\boxed{\begin{array}{c}\text{The smallest positive value of}\\ ax + by\\ \text{is equal to } \gcd(a, b).\end{array}}$$

There are many ways to prove that this is true. We will take a constructive approach, via the Euclidean algorithm, which has the advantage of giving a procedure for finding the appropriate values of x and y. In other words, we are going to describe a method of finding integers x and y that are solutions to the equation

$$ax + by = \gcd(a, b).$$

Since, as we have already observed, every number $ax+by$ is divisible by $\gcd(a, b)$, it will follow that the smallest positive value of $ax + by$ is precisely $\gcd(a, b)$.

How might we solve the equation $ax + by = \gcd(a, b)$? If a and b are small, we might be able to guess a solution. For example, the equation

$$10x + 35y = 5$$

has the solution $x = -3$ and $y = 1$, and the equation

$$7x + 11y = 1$$

has the solution $x = -3$ and $y = 2$. We also notice that there can be more than one solution, since $x = 8$ and $y = -5$ is also a solution to $7x + 11y = 1$.

However, if a and b are large, neither guesswork nor trial and error is going to be helpful. We are going to start by illustrating the Euclidean algorithm method for solving $ax + by = \gcd(a, b)$ with a particular example. So we are going to try to solve

$$22x + 60y = \gcd(22, 60).$$

The first step is to perform the Euclidean algorithm to compute the gcd. We find

$$\begin{aligned}
60 &= 2 \times 22 + 16\\
22 &= 1 \times 16 + 6\\
16 &= 2 \times 6 + 4\\
6 &= 1 \times 4 + 2\\
4 &= 2 \times 2 + 0
\end{aligned}$$

This shows that $\gcd(22, 60) = 2$, a fact that is clear without recourse to the Euclidean algorithm. However, the Euclidean algorithm computation is important because we're going to use the intermediate quotients and remainders to solve the equation $22x + 60y = 2$. The first step is to rewrite the first equation as

$$16 = a - 2b, \qquad \text{where we let } a = 60 \text{ and } b = 22.$$

We next substitute this value into the 16 appearing in the second equation. This gives (remember that $b = 22$)

$$b = 1 \times 16 + 6 = 1 \times (a - 2b) + 6.$$

Rearranging this equation to isolate the remainder 6 yields

$$6 = b - (a - 2b) = -a + 3b.$$

Now substitute the values 16 and 6 into the next equation, $16 = 2 \times 6 + 4$:

$$a - 2b = 16 = 2 \times 6 + 4 = 2(-a + 3b) + 4.$$

Again we isolate the remainder 4, yielding

$$4 = (a - 2b) - 2(-a + 3b) = 3a - 8b.$$

Finally, we use the equation $6 = 1 \times 4 + 2$ to get

$$-a + 3b = 6 = 1 \times 4 + 2 = 1 \times (3a - 8b) + 2.$$

Rearranging this equation gives the desired solution

$$-4a + 11b = 2.$$

(We should check our solution: $-4 \times 60 + 11 \times 22 = -240 + 242 = 2$.)

We can summarize the above computation in the following efficient tabular form. Note that the left-hand equations are the Euclidean algorithm, and the right-hand equations compute the solution to $ax + by = \gcd(a, b)$.

$$\begin{array}{r@{\;=\;}l|r@{\;=\;}l}
a & 2 \times b + 16 & 16 & a - 2b \\
b & 1 \times 16 + 6 & 6 & b - 1 \times 16 \\
\multicolumn{2}{c|}{} & & b - 1 \times (a - 2b) \\
\multicolumn{2}{c|}{} & & -a + 3b \\
16 & 2 \times 6 + 4 & 4 & 16 - 2 \times 6 \\
\multicolumn{2}{c|}{} & & (a - 2b) - 2 \times (-a + 3b) \\
\multicolumn{2}{c|}{} & & 3a - 8b \\
6 & 1 \times 4 + 2 & 2 & 6 - 1 \times 4 \\
\multicolumn{2}{c|}{} & & (-a + 3b) - 1 \times (3a - 8b) \\
\multicolumn{2}{c|}{} & & -4a + 11b \\
4 & 2 \times 2 + 0 & \multicolumn{2}{l}{}
\end{array}$$

Why does this method work? As the following table makes clear, we start with the first two lines of the Euclidean algorithm, which involve the quantities a and b, and work our way down.

$$\begin{array}{r|l}
a = \ q_1 b + r_1 & r_1 = a - q_1 b \\
b = q_2 r_1 + r_2 & r_2 = b - q_2 r_1 \\
 & = b - q_2(a - q_1 b) \\
 & = -q_2 a + (1 + q_1 q_2) b \\
r_1 = q_3 r_2 + r_3 & r_3 = r_1 - q_3 r_2 \\
 & = (a - q_1 b) - q_3\bigl(-q_2 a + (1 + q_1 q_2) b\bigr) \\
 & = (1 + q_2 q_3) a - (q_1 + q_3 + q_1 q_2 q_3) b \\
\vdots & \qquad\qquad\vdots
\end{array}$$

As we move from line to line, we will continually be forming equations that look like

$$\text{latest remainder} = \text{some multiple of } a \text{ plus some multiple of } b.$$

Eventually, we get down to the last nonzero remainder, which we know is equal to $\gcd(a, b)$, and this gives the desired solution to the equation $\gcd(a, b) = ax + by$.

A larger example with $a = 12453$ and $b = 2347$ is given in tabular form on top of the next page. As before, the left-hand side is the Euclidean algorithm and the right-hand side solves $ax + by = \gcd(a, b)$. We see that $\gcd(12453, 2347) = 1$ and that the equation $12453x + 2347y = 1$ has the solution $(x, y) = (304, -1613)$.

We now know that the equation

$$ax + by = \gcd(a, b)$$

always has a solution in integers x and y. The final topic we discuss in this section is the question of how many solutions it has, and how to describe all the solutions. Let's start with the case that a and b are relatively prime, that is, $\gcd(a, b) = 1$, and suppose that (x_1, y_1) is a solution to the equation

$$ax + by = 1.$$

We can create additional solutions by subtracting a multiple of b from x_1 and adding the same multiple of a onto y_1. In other words, for any integer k we obtain a new solution $(x_1 + kb, y_1 - ka)$.[1] We can check that this is indeed a solution by computing

$$a(x_1 + kb) + b(y_1 - ka) = ax_1 + akb + by_1 - bka = ax_1 + by_1 = 1.$$

[1] Geometrically, we are starting from the known point (x_1, y_1) on the line $ax + by = 1$ and using the fact that the line has slope $-a/b$ to find new points $(x_1 + t, y_1 - (a/b)t)$. To get new points with integer coordinates, we need to let t be a multiple of b. Substituting $t = kb$ gives the new integer solution $(x_1 + kb, y_1 - ka)$.

$a = 5 \times b \quad + 718$	$718 = a - 5b$
$b = 3 \times 718 + 193$	$193 = b - 3 \times 718$
	$= b - 3 \times (a - 5b)$
	$= -3a + 16b$
$718 = 3 \times 193 + 139$	$139 = 718 - 3 \times 193$
	$= (a - 5b) - 3 \times (-3a + 16b)$
	$= 10a - 53b$
$193 = 1 \times 139 + 54$	$54 = 193 - 139$
	$= (-3a + 16b) - (10a - 53b)$
	$= -13a + 69b$
$139 = 2 \times 54 \quad + 31$	$31 = 139 - 2 \times 54$
	$= (10a - 53b) - 2 \times (-13a + 69b)$
	$= 36a - 191b$
$54 = 1 \times 31 \quad + 23$	$23 = 54 - 31$
	$= -13a + 69b - (36a - 191b)$
	$= -49a + 260b$
$31 = 1 \times 23 \quad + 8$	$8 = 31 - 23$
	$= 36a - 191b - (-49a + 260b)$
	$= 85a - 451b$
$23 = 2 \times 8 \quad + 7$	$7 = 23 - 2 \times 8$
	$= (-49a + 260b) - 2 \times (85a - 451b)$
	$= -219a + 1162b$
$8 = 1 \times 7 \quad + 1$	$1 = 8 - 7$
	$= 85a - 451b - (-219a + 1162b)$
	$= 304a - 1613b$
$7 = 7 \times 1 \quad + 0$	

So, for example, if we start with the solution $(-1, 2)$ to $5x + 3y = 1$, we obtain new solutions $(-1 + 3k, 2 - 5k)$. Note that the integer k is allowed to be positive, negative, or zero. Putting in particular values of k gives the solutions

$$\ldots(-13, 22),\ (-10, 17),\ (-7, 12),\ (-4, 7),\ (-1, 2),$$
$$(2, -3),\ (5, -8),\ (8, -13),\ (11, -18)\ldots.$$

Still looking at the case that $\gcd(a, b) = 1$, we can show that this procedure gives all possible solutions. Suppose that we are given two solutions (x_1, y_1) and (x_2, y_2) to the equation $ax + by = 1$. In other words,

$$ax_1 + by_1 = 1 \qquad \text{and} \qquad ax_2 + by_2 = 1.$$

We are going to multiply the first equation by y_2, multiply the second equation by y_1, and subtract. This will eliminate b and, after a little bit of algebra, we are left with

$$ax_1y_2 - ax_2y_1 = y_2 - y_1.$$

Similarly, if we multiply the first equation by x_2, multiply the second equation by x_1, and subtract, we find that

$$bx_2y_1 - bx_1y_2 = x_2 - x_1.$$

So if we let $k = x_2y_1 - x_1y_2$, then we find that

$$x_2 = x_1 + kb \qquad \text{and} \qquad y_2 = y_1 - ka.$$

This means that the second solution (x_2, y_2) is obtained from the first solution (x_1, y_1) by adding a multiple of b onto x_1 and subtracting the same multiple of a from y_1. So every solution to $ax + by = 1$ can be obtained from the initial solution (x_1, y_1) by substituting different values of k into $(x_1 + kb, y_1 - ka)$.

What happens if $\gcd(a, b) > 1$? To make the formulas look a little bit simpler, we will let $g = \gcd(a, b)$. We know from the Euclidean algorithm method that there is at least one solution (x_1, y_1) to the equation

$$ax + by = g.$$

But g divides both a and b, so (x_1, y_1) is a solution to the simpler equation

$$\frac{a}{g}x + \frac{b}{g}y = 1.$$

Now our earlier work applies, so we know that every other solution can be obtained by substituting values for k in the formula

$$\left(x_1 + k \cdot \frac{b}{g},\ y_1 - k \cdot \frac{a}{g}\right).$$

This completes our description of the solutions to the equation $ax + by = g$, as summarized in the following theorem.

Theorem 6.1 (Linear Equation Theorem)**.** *Let a and b be nonzero integers, and let $g = \gcd(a, b)$. The equation*

$$ax + by = g$$

always has a solution (x_1, y_1) in integers, and this solution can be found by the Euclidean algorithm method described earlier. Then every solution to the equation can be obtained by substituting integers k into the formula

$$\left(x_1 + k \cdot \frac{b}{g},\ y_1 - k \cdot \frac{a}{g}\right).$$

For example, we saw that the equation

$$60x + 22y = \gcd(60, 22) = 2$$

has the solution $x = -4$, $y = 11$. Then our Linear Equation Theorem says that every solution is obtained from the formula

$$(-4 + 11k, 11 - 30k) \qquad \text{with } k \text{ any integer.}$$

In particular, if we want a solution with x positive, then we can take $k = 1$, which gives the smallest such solution $(x, y) = (7, -19)$.

In this chapter we have shown that the equation

$$ax + by = \gcd(a, b)$$

always has a solution. This fact is extremely important for both theoretical and practical reasons, and we will be using it repeatedly in our number theoretic investigations. For example, we will need to solve the equation $ax + by = 1$ when we study cryptography in Chapter 18. And in the next chapter we will use this equation for our theoretical study of factorization of numbers into primes.

Exercises

6.1. **(a)** Find a solution in integers to the equation

$$12345x + 67890y = \gcd(12345, 67890).$$

(b) Find a solution in integers to the equation

$$54321x + 9876y = \gcd(54321, 9876).$$

6.2. Describe all integer solutions to each of the following equations.

(a) $105x + 121y = 1$

(b) $12345x + 67890y = \gcd(12345, 67890)$

(c) $54321x + 9876y = \gcd(54321, 9876)$

6.3. 💻 The method for solving $ax + by = \gcd(a, b)$ described in this chapter involves a considerable amount of manipulation and back substitution. This exercise describes an alternative way to compute x and y that is especially easy to implement on a computer.

(a) Show that the algorithm described in Figure 6.1 computes the greatest common divisor g of the positive integers a and b, together with a solution (x, y) in integers to the equation $ax + by = \gcd(a, b)$.

(b) Implement the algorithm on a computer using the computer language of your choice.