

GCD's and the Euclidean Algorithm

THE **greatest common divisor** OF a AND b , NOT BOTH 0, is the largest integer which divides both a and b . For example, the greatest common divisor of 21 and 35 is 7. We write $\gcd(a, b)$, as shorthand for the greatest common divisor of a and b . So $\gcd(35, 21) = 7$.

There are several ways to find the gcd of two integers, a and b (not both 0).

First, we could simply list all the positive divisors of a and b and pick the largest number that appears in both lists. Notice that 1 will appear in both lists. For the example above the positive divisors of 35 are 1, 5, 7, and 35. For 21 the positive divisors are 1, 3, 7, and 21. The largest number appearing in both lists is 7, so $\gcd(35, 21) = 7$.

Another way to say the same thing: If we let D_a denote the set of positive divisors of a , then $\gcd(a, b) =$ the largest number in $D_a \cap D_b$.

The reason $\gcd(0, 0)$ is not defined is that every positive integer divides 0, and so there is no largest integer that divides 0. From now on, when we use the symbol $\gcd(a, b)$, we will tacitly assume a and b are not both 0. The integers a and b can be negative. For example if $a = -34$ and $b = 14$, then the set of positive divisors of -34 is $\{1, 2, 17, 34\}$ and the set of positive divisors of 14 is $\{1, 2, 7, 14\}$. The set of positive common divisors of 14 and -34 is the set $\{1, 2, 17, 34\} \cap \{1, 2, 7, 14\} = \{1, 2\}$. The largest number in this set is $2 = \gcd(-34, 14)$.

Obviously then $\gcd(a, b) = \gcd(-a, b)$ since a and $-a$ have the same set of positive divisors. So when computing the $\gcd(a, b)$ we

may as well replace a and b by their absolute values if one or both happen to be negative.

Here are a few easy facts about gcd's:

- (1) If $a \neq 0$, then $\gcd(a, a) = a$.
- (2) $\gcd(a, 1) = 1$.
- (3) $\gcd(a, b) = \gcd(b, a)$. (The order a and b are given is not important,
but it is traditional to list them with $a \geq b$.)
- (4) If $a \neq 0$ and $a|b$, then $\gcd(a, b) = |a|$.
- (5) If $a \neq 0$, $\gcd(a, 0) = |a|$.

If $\gcd(a, b) = 1$, we say that a and b are **relatively prime**. When a and b are relatively prime, they have no common prime divisor. For example 12 and 35 are relatively prime.

22.1 Euclidean algorithm

It's pretty clear that computing $\gcd(a, b)$ by listing all the positive divisors of a and all the positive divisors of b , and selecting the largest integers that appears in both lists is not very efficient. There is a better way of computing $\gcd(a, b)$.

Theorem 22.1. *If a and b are integers (not both 0) and $a = sb + t$ for integers s and t , then $\gcd(a, b) = \gcd(b, t)$.*

Proof. *To prove the theorem, we will show that the list of positive integers that divide both a and b is identical to the list of positive integers that divide both b and $t = a - sb$. So, suppose $d|a$ and $d|b$. Then $d|(a - sb)$ so $d|t$. Hence d divides both b and t . On the other hand, suppose $d|b$ and $d|t$. Then $d|(sb + t)$, so that $d|a$. Hence d divides both a and b . It follows that $\gcd(a, b) = \gcd(b, t)$. ♣*

Euclid is given the credit for discovering this fact, and its use for computing gcd's is called the **Euclidean algorithm** in his honor.

The idea is to use the theorem repeatedly until a pair of numbers is reached for which the gcd is obvious. Here is an example of the Euclidean algorithm in action.

Example 22.2. Since $14 = 1 \cdot 10 + 4$, $\gcd(14, 10) = \gcd(10, 4)$. In turn $10 = 2 \cdot 4 + 2$ so $\gcd(10, 4) = \gcd(4, 2)$. Since $4 = 2 \cdot 2$, $\gcd(4, 2) = \gcd(2, 0) = 2$. So $\gcd(10, 14) = 2$.

The same example, presented a little more compactly, and without explicitly writing out the divisions, looks like

$$\gcd(14, 10) = \gcd(10, 4) = \gcd(4, 2) = \gcd(2, 0) = 2$$

At each step, the second number is replaced by the remainder when the first number is divided by the second, and the second moves into the first spot.

The process is repeated until the second number is a 0 (which must happen eventually since the second number never will be negative, and it goes down by at least 1 with each repetition of the process). The gcd is then the number in the first spot when the second spot is 0 in the last step of the algorithm.

Now, a more exciting example.

Example 22.3. Find the greatest common divisor of 540 and 252. We may present the computations compactly, without writing¹ out the divisions. We have

$$\gcd(540, 252) = \gcd(252, 36) = \gcd(36, 0) = 36.$$

¹ Do the divisions yourself to verify the results.

22.2 Efficiency of the Euclidean algorithm

Using the Euclidean algorithm to find gcd's is extremely efficient. Using a calculator with a ten digit display, you can find the gcd of two ten digit integers in a matter of a few minutes at most using the Euclidean algorithm. On the other hand, doing the same problem by first finding the positive divisors of the two ten digit integers would be a tedious project lasting several days. Some modern cryptographic systems rely on the computation of the gcd's of integers of hundreds of digits. Finding the positive divisors of such large integers, even with a computer, is, at present, a hopeless task. But a computer implementation of the Euclidean algorithm will produce the gcd of integers of hundreds of digits in the blink of an eye.

22.3 *The Euclidean algorithm in quotient/remainder form*

The Euclidean algorithm can also be written out as a sequence of divisions:

$$\begin{aligned}
 a &= q_1 \cdot b + r_1, & 0 < r_1 < b \\
 b &= q_2 \cdot r_1 + r_2, & 0 < r_2 < r_1 \\
 r_1 &= q_3 \cdot r_2 + r_3, & 0 < r_3 < r_2 \\
 &\vdots \\
 r_k &= q_{k+2} \cdot r_{k+1} + r_{k+2}, & 0 < r_{k+2} < r_{k+1} \\
 &\vdots \\
 r_{n-2} &= q_n \cdot r_{n-1} + r_n, & 0 < r_n < r_{n-1} \\
 r_{n-1} &= q_{n+1} \cdot r_n + 0
 \end{aligned}$$

The sequence of integer remainders $b > r_1 > \dots > r_k > \dots \geq 0$ must eventually reach 0. Let's say $r_n \neq 0$, but $r_{n+1} = 0$, so that $r_{n-1} = q_{n+1} \cdot r_n$. That is, in the sequence of remainders, r_n is the last non-zero term. Then, just as in the examples above we see that the gcd of a and b is the last nonzero remainder:

$$\begin{aligned}
 \gcd(a, b) &= \gcd(b, r_1) = \gcd(r_1, r_2) = \dots = \gcd(r_{n-1}, r_n) \\
 &= \gcd(r_n, r_{n+1}) = \gcd(r_n, 0) = r_n.
 \end{aligned}$$

Let's find $\gcd(317, 118)$ using this version of the Euclidean algorithm.

Here are the steps:

$$\begin{aligned}
 317 &= 2 \cdot 118 + 81 \\
 118 &= 1 \cdot 81 + 37 \\
 81 &= 2 \cdot 37 + 7 \\
 37 &= 5 \cdot 7 + 2 \\
 7 &= 3 \cdot 2 + 1 \\
 2 &= 2 \cdot 1 + 0
 \end{aligned}$$

Since the last non-zero remainder is 1, we conclude that $\gcd(317, 118) = 1$. So, in the terminology introduced above, we would say that 317 and 118 are relatively prime.

22.4 Exercises

Exercise 22.1. Use the Euclidean algorithm to compute $\gcd(a, b)$ in each case.

- a) $a = 233, b = 89$ b) $a = 1001, b = 13$ c) $a = 2457, b = 1458$
 d) $a = 567, b = 349$

Exercise 22.2. Compute $\gcd(987654321, 123456789)$.

Exercise 22.3. Write a step-by-step algorithm that implements the Euclidean algorithm for finding gcd's.

Exercise 22.4. If n is a positive integer, what is $\gcd(n, 2n)$?

22.5 Problems

Problem 22.1. Use the Euclidean algorithm to compute $\gcd(a, b)$ in each case.

- a) $a = 216, b = 111$ b) $a = 1001, b = 11$ c) $a = 663, b = 5168$
 d) $a = 1357, b = 2468$

Problem 22.2. Compute $\gcd(733103, 91637)$.

Problem 22.3. If p is a prime, and n is any integer, what are the possible values of $\gcd(p, n)$?

Problem 22.4. Prove or give a counterexample: If p and q are two different primes, then $\gcd(2p, 2q) = 2$.

Problem 22.5. If p is a prime, and m is a positive integer, determine $\gcd(p, p^m)$.

Problem 22.6. If p is a prime, and $m \leq n$ are positive integers, determine $\gcd(p^m, p^n)$.

Problem 22.7. Show that if n is a positive integer, then $\gcd(n, n + 1) = 1$.

GCD's Reprised

THE gcd OF a AND b IS DEFINED to be the largest integer that divides them both. But there is another way to describe that gcd. First, a little vocabulary: by a **linear combination** of a and b we mean any expression of the form $as + bt$ where s, t are integers. For example, $4 \cdot 5 + 10 \cdot 2 = 40$ is a linear combination of 4 and 10. Here are some more linear combinations of 4 and 10:

$$4 \cdot 1 + 10 \cdot 1 = 14, \quad 4 \cdot 0 + 10 \cdot 0 = 0, \quad \text{and,} \quad 4 \cdot (-11) + 10 \cdot 1 = -34.$$

23.1 The $\gcd(a, b)$ as a linear combination of a and b

If we make a list of all possible linear combinations of 4 and 10, an unexpected pattern appears: $\dots, -6, -4, -2, 0, 2, 4, 6, \dots$. Since 4 and 10 are both even, we are sure to see only even integers in the list of linear combinations, but the surprise is that *every* even number is in the list. Now here's the connection with gcd's: The gcd of 4 and 10 is 2, and the list of all linear combinations is exactly all multiples of 2. Let's prove that was no accident.

Theorem 23.1. *Let a, b be two integers (not both zero). Then the smallest positive number in the list of the linear combinations of a and b is $\gcd(a, b)$. In other words, the $\gcd(a, b)$ is the smallest positive integer that can be written as a linear combination of a and b .*

Proof. Let $L = \{as + bt \mid s, t \text{ are integers and } as + bt > 0\}$. Since a, b are not both 0, we see this set is nonempty. As a nonempty set of positive inte-

gers, it must have a least element, say m . Since $m \in L$, m is a linear combination of a and b . Say $m = as_0 + bt_0$. We need to show $m = \gcd(a, b) = d$. As noted above, since $d|a$ and $d|b$, it must be that $d|(as_0 + bt_0)$, so $d|m$. That implies $d \leq m$. We complete the proof by showing m is a common divisor of a and b . The plan is to divide a by m and show the remainder must be 0. So write $a = qm + r$ with $0 \leq r < m$. Solving for r we get $0 \leq r = a - qm = a - q(as_0 + bt_0) = a(1 - qs_0) + b(-qt_0) < m$. That shows r is a linear combination of a and b that is less than m . Since m is the smallest positive linear combination of a and b , the only option for r is $r = 0$. Thus $a = qm$, and so $m|a$. In the same way, $m|b$. Since m is a common divisor of a and b , it follows that $m \leq d$. Since the reverse inequality is also true, we conclude $m = d$. ♣

And now we are ready for the punch-line.

Theorem 23.2. *Let a, b be two integers (not both zero). Then the list of all the linear combinations of a and b consists of all the multiples of $\gcd(a, b)$.*

Proof. Since $\gcd(a, b) = d$ certainly divides any linear combination of a and b , only multiples of d stand a chance to be in the list. Now we need to show that if n is a multiple of the d then n will appear in the list for sure. According to the last theorem, we can find integers s_0, t_0 so that $d = as_0 + bt_0$. Now since n is a multiple of d , we can write $n = de$. Multiplying both sides of $d = as_0 + bt_0$ by e gives $a(s_0e) + b(t_0e) = de = n$, and that shows n does appear in the list of linear combinations of a and b . ♣

So, without doing any computations, we can be sure that the set of all linear combinations of 15 and 6 will be all multiples of 3.

23.2 Back-solving to express $\gcd(a, b)$ as a linear combination

In practice, finding integers s and t so that $as + bt = d = \gcd(a, b)$ is carried out by using the Euclidean algorithm applied to a and b and then *back-solving*.

Example 23.3. Let $a = 35$ and $b = 55$. Then the Euclidean algorithm gives

$$55 = 35 \cdot 1 + 20$$

$$35 = 20 \cdot 1 + 15$$

$$20 = 15 \cdot 1 + 5$$

$$15 = 5 \cdot 3 + 0$$

The penultimate equation allows us to write $5 = 1 \cdot 20 + (-1) \cdot 15$ as a linear combination of 20 and 15. We then use the equation $35 = 20 \cdot 1 + 15$, to write $15 = 1 \cdot 35 + (-1) \cdot 20$. We can substitute this into the previous expression for 5 as a linear combination of 20 and 15 to get $5 = 1 \cdot 20 + (-1) \cdot 15 = 1 \cdot 20 + (-1) \cdot (1 \cdot 35 + (-1) \cdot 20)$. Which can be simplified by collecting 35's and 20's to write $5 = 2 \cdot 20 + (-1) \cdot 35$. Now we can use the top equation to write $20 = 1 \cdot 55 + (-1) \cdot 35$ and substitute this into the expression giving 5 as a linear combination of 35 and 20. We get $5 = 2 \cdot (1 \cdot 55 + (-1) \cdot 35) + (-1) \cdot 35$. This simplifies to $5 = 2 \cdot 55 + (-3) \cdot 35$.

23.3 Extended Euclidean Algorithm

The sort of computation in section 23.2 gets a little tedious, keeping track of equations and coefficients. Moreover, the back-substitution method isn't very pleasant from a programming perspective since all the equations in the Euclidean algorithm need to be saved before solving for the coefficients in a linear combination for the $\gcd(a, b)$. The Extended Euclidean Algorithm is a forward-substitution method that allows us to compute linear combinations as we calculate remainders on the way to finding $\gcd(a, b)$.

There are two new ideas that we need to add to our Euclidean Algorithm for computing $\gcd(a, b)$: every remainder can be written as a linear combination of a and b , and we can use the quotient-remainder computation to generate new linear combinations. An example is the best way to see how this works.

Example 23.4. Let $a = 6567$ and $b = 987$. We would like to find $\gcd(a, b) = \gcd(6567, 987)$ and coefficients s and t in a linear combi-

nation: $a(s) + b(t) = \gcd(a, b)$. Suppose that we had found the remainders $r_3 = 303$ and $r_4 = 39$, and had found corresponding linear combinations:

$$\text{eqn 3: } 6567(2) + 987(-13) = 303, \text{ and}$$

$$\text{eqn 4: } 6567(-3) + 987(20) = 39.$$

To find the next remainder, r_5 , the Euclidean Algorithm has us calculate the quotient, $q_4 = 7$, and then the remainder as $r_5 = r_3 - r_4 \times q_4 = 30$. The Extended Euclidean Algorithm finds the next equation, (eqn 5), by performing the same operation on (eqn 3) and (eqn 4):

$$(6567(2) + 987(-13)) - (6567(-3) + 987(20)) \times (7) = (303) - (39) \times (7),$$

$$6567(23) + 987(-153) = 30.$$

In order to get the process started we need two initial equations based on the values of a and b . That is, we will pretend that they are “remainders”, say $r_{-1} = 6567$ and $r_0 = 987$, that come before the first true remainder, r_1 . The complete Extended Euclidean Algorithm process is shown in the table. The last equation with a non-zero remainder is a linear combination of 6567 and 987 equal to the $\gcd(6567, 987) = 3$. That is, we have

$$6567(101) + 987(-672) = 3.$$

Notice that every equation in Example 23.4 has the same form:

$$6567(s_i) + 987(t_i) = r_i. \quad (23.1)$$

The only values that change are the linear combination coefficients, s_i and t_i , the remainders, r_i , and the quotient values q_i . If we remember the equation form 23.1, we could just write the changing values in a tabular form (see Table 23.1).

To generate a new, i th row in a table consider the two consecutive previous rows as equations:

$$a(s_{i-2}) + b(t_{i-2}) = r_{i-2}, \text{ and } a(s_{i-1}) + b(t_{i-1}) = r_{i-1}.$$

After finding the quotient q_i so that $r_{i-2} = q_i r_{i-1} + r_i$, we subtract q_i times the $i - 1$ equation from the $i - 2$ equation.

Check this computation!

$$\begin{aligned} \text{eqn -1: } & 6567(1) + 987(0) = 6567, \\ \text{eqn 0: } & 6567(0) + 987(1) = 987, \\ \text{eqn 1: } & 6567(1) + 987(-6) = 645, (q_1 = 6), \\ \text{eqn 2: } & 6567(-1) + 987(7) = 342, (q_2 = 1), \\ \text{eqn 3: } & 6567(2) + 987(-13) = 303, (q_3 = 1), \\ \text{eqn 4: } & 6567(-3) + 987(20) = 39, (q_4 = 1), \\ \text{eqn 5: } & 6567(23) + 987(-153) = 30, (q_5 = 7), \\ \text{eqn 6: } & 6567(-26) + 987(173) = 9, (q_6 = 1), \\ \text{eqn 7: } & 6567(101) + 987(-672) = 3, (q_7 = 3), \\ \text{eqn 8: } & 6567(-329) + 987(2189) = 0, (q_8 = 3). \end{aligned}$$

The s_i and t_i are called *Bézout Coefficients*.

i	s_i	t_i	r_i	q_i
-1	1	0	6567	
0	0	1	987	
1	1	-6	645	6
2	-1	7	342	1
3	2	-13	303	1
4	-3	20	39	1
5	23	-153	30	7
6	-26	173	9	1
7	101	-672	3	3
8	-329	2189	0	3

Table 23.1: $i: 6567(s_i) + 987(t_i) = r_i, q_i$

Simplifying the resulting expression, we obtain

$$a(s_{i-2} - q_i \cdot s_{i-1}) + b(t_{i-2} - q_i \cdot t_{i-1}) = r_{i-2} - q_i \cdot r_{i-1}. \quad (23.2)$$

This is our new, i th equation:

$$a(s_i) + b(t_i) = r_i. \quad (23.3)$$

Comparing equations 23.2 with equation 23.3, we see that the s_i and t_i are calculated in exactly the same way from the previous two values as the remainder r_i is.

Since the equation number i plays no role in the computations, we need not include that column in the tableaux.

Example 23.5. Find $d = \gcd(55, 35)$ and a linear combination $55(s) + 35(t) = d$.

Complete the tableaux in the table using the Extended Euclidean Algorithm. Compare these calculations to those of Example 23.3, where we used the back-substitution method.

s_i	t_i	r_i	q_i
1	0	55	
0	1	35	
1	-1	20	1
-1		15	1
			1
			3

Often, when performing the Algorithm by hand it is more convenient to write the tableaux horizontally, as in the following example.

Example 23.6. Find $\gcd(107653, 22869)$, and write it as a linear combination of those two numbers. The complete Extended Euclidean Algorithm table is displayed in table 23.2, where the rows correspond to r_i , q_i , t_i , s_i , respectively.

107653	22869	16177	6692	2793	1106	581	525	56	21	14	7	0
		4	1	2	2	2	1	1	9	2	1	2
0	1	-4	5	-14	33	-80	113	-193	1850	-3893	5743	-15379
1	0	1	-1	3	-7	17	-24	41	-393	827	-1220	3267

Table 23.2: $\gcd(107653, 22869)$

Thus, we may conclude that

$$\gcd(107653, 22869) = 7 = (107653)(-1220) + (22869)(5743).$$

23.4 General Linear Combinations for $\gcd(a, b)$

In section 23.3 we saw how the Extended Euclidean Algorithm may be used to find a linear combination of the form $a(s) + b(t) = \gcd(a, b)$. It is often necessary to find all such linear combinations, (see section 25.3). In particular, such general linear combinations are found in the study of cryptography. The Algorithm stops when we reach a remainder of zero. It turns out that the corresponding values of s_i and t_i , which we haven't used yet, allow us to find the form for all linear combinations that equal $\gcd(a, b)$.

Example 23.7. Consider example 23.4 wherein we wanted to express $\gcd(6567, 987)$ as a linear combination. We found at the end of the process that $\gcd(6567, 987) = 3$ and

$$\begin{aligned} 6567(101) + 987(-672) &= 3, \text{ and,} \\ 6567(-329) + 987(2189) &= 0. \end{aligned}$$

Now, for any integer, say n , we may multiply the last equation by n and retain zero on the right. Finally, if we add that new equation to the previous one, we obtain

$$\begin{aligned} 6567(101 - 329n) + 987(-672 + 2189n) &= 3, \text{ and,} \\ 6567(-329n) + 987(2189n) &= 0. \end{aligned}$$

The new penultimate equation gives the form for all the linear combinations equal to the $\gcd(6567, 987)$:

$$6567(101 - 329n) + 987(-672 + 2189n) = \gcd(6567, 987) = 3.$$

Each pair of Bézout Coefficients can be shown to be relatively prime¹. In particular, the last equation in the Extended Euclidean Algorithm,

$$a(s_{k+1}) + b(t_{k+1}) = 0,$$

has minimal coefficients s_{k+1} and t_{k+1} in that every other such pair is a common multiple of these two². This means that, given any integer n , ns_{k+1} and nt_{k+1} also satisfy the equation. In fact, it is easy to see

¹ See page 206 for the definition of relatively prime.

² See section 25.3.

that

$$a(ns_{k+1}) + b(nt_{k+1}) = 0.$$

Finally, to obtain the general solution, as we did in example 23.7, we add this to the Bézout equation for the $\gcd(a, b)$, obtaining

$$a(s_k + ns_{k+1}) + b(t_k + nt_{k+1}) = \gcd(a, b).$$

One of s_{k+1} and t_{k+1} will always be negative!

23.5 Exercises

Exercise 23.1. Determine $\gcd(13447, 7667)$ and write it as a linear combination of 13447 and 7667. Try both the method of back-substitution and the Extended Euclidean Algorithm to determine a suitable linear combination.

Exercise 23.2. What can you conclude about $\gcd(a, b)$ if there are integers s, t with $as + bt = 1$?

Exercise 23.3. What can you conclude about $\gcd(a, b)$ if there are integers s, t with $as + bt = 19$?

Exercise 23.4. What can you conclude about $\gcd(a, b)$ if there are integers s, t with $as + bt = 18$?

23.6 Problems

Problem 23.1. Determine $\gcd(41559, 39417)$ and write it as a linear combination of 41559 and 39417. Try both the method of back-substitution and the Extended Euclidean Algorithm to determine a suitable linear combination.

Problem 23.2. What can you conclude about $\gcd(a, b)$ if there are integers s, t with $as + bt = 12$?

Problem 23.3. What is the smallest positive integer that can be written as a linear combination of 2191 and 1351?

Problem 23.4. Definition: The least common multiple of the positive integers a and b is the smallest positive integer that is divisible by both a and b .

Example: the least common multiple of 24 and 18 is 72. Write that as $\text{lcm}(24, 18) = 72$. You might recall the notion of least common multiple from the time you learned how to add fractions. The idea was that to add two fractions, $\frac{a}{b}$ and $\frac{c}{d}$, first write the two as equivalent fractions with the same denominator. For example, to add $\frac{2}{3}$ and $\frac{5}{4}$, write them as $\frac{8}{12}$ and $\frac{15}{12}$, then add to get $\frac{23}{12}$. The least common denominator when adding $\frac{a}{b}$ and $\frac{c}{d}$ is the least common multiple of the denominators, $\text{lcm}(b, d)$

Determine $\text{lcm}(22, 33)$. (The answer is not 726.)

Problem 23.5. *The result when $\gcd(a, b)$ and $\text{lcm}(a, b)$ are multiplied is always a simple combination of a and b . Example: $\gcd(6, 4) \text{lcm}(6, 4) = 2 \cdot 12 = 24$. Try a few more examples, and see if you can guess the value of $\gcd(a, b) \text{lcm}(a, b)$.*