

Math 5330
Spring 2009

**The Greatest Common Divisor
and Euclid's Algorithm**
(Notes, part 3)

This set of notes is a supplement to what is in the book. The order in which I do things is slightly different than the book's order but much of what is in this set of notes can be found in sections 1.3, 2.1 and 2.3 of the book.

We have one last detail to complete on the proof of the Fundamental Theorem and that is to prove the lemma:

Lemma. If p is a prime and ab is divisible by p then a is divisible by p or b is divisible by p .

We are not quite in a position to do this yet; we require a bit more theory. First, we introduce some notation and basic arithmetic results.

Theorem (The Division Algorithm). If n is any integer and k is any positive integer, then there are **unique** integers q and r such that

$$1) \quad n = qk + r,$$

and

$$2) \quad 0 \leq r < k.$$

This says that we can divide n by k to get a quotient q and a remainder r , and we can always get a remainder which is non-negative and smaller than k .

Proof. The proof has two parts: the existence of q, r , and the fact that q, r have to be unique. Let's give a proof by infinite descent that q, r always exist provided $n \geq 0$. The proof goes like this: suppose this is not true. Suppose, for example, that we can't find q, r with $n = qk + r$, and $0 \leq r < k$. The problem HAS to be with r , since we can say $n = 0 \times k + n$, so lots of q 's exist. It must be that there is no q for which $r = n - qk$ satisfies $0 \leq r < k$. In this case, it must be that $n > k$, since if $0 < n \leq k$, we could use $q = 0$, and if $n = k$, we could use $q = 1, r = n - k$. Now consider the number $n_1 = n - k$. If $n_1 = q_1k + r_1$, where $0 \leq r_1 < k$, then $n = (q_1 + 1)k + r_1$, so we could use $q = q_1 + 1, r = r_1$ for n . This means that the nonexistence of r for n forces the nonexistence of r for $n - k$. Since $n - k > 0$, this gives the infinite descent.

To see that the values of q and r are unique (as long as $0 \leq r < k$), suppose that $n = q_1k + r_1$ and $n = q_2k + r_2$. Then subtracting gives $0 = (q_1 - q_2)k + r_1 - r_2$ or $r_2 - r_1 = (q_1 - q_2)k$. So $r_2 - r_1$ is divisible by k . If we assume that $0 \leq r_1 < k$ and $0 \leq r_2 < k$, then $-(k - 1) \leq r_2 - r_1 \leq (k - 1)$. But the only number between $-(k - 1)$ and $(k - 1)$ which is divisible by k is 0. This means that $r_2 - r_1 = 0$, so $r_1 = r_2$. Also, $r_2 - r_1 = (q_1 - q_2)k$ so $0 = (q_1 - q_2)k$. Thus, $q_1 = q_2$, completing the proof.

Here is how I often calculate quotients and remainders on a calculator: Suppose we want the quotient and remainder when 6815 is divided by 27.

Dividing on a calculator gives $\frac{6815}{27} = 252.407\cdots$. This tells us that $q = 252$.

If we subtract 252 from this number, we have $\frac{6815}{27} - 252 = 0.407\cdots$.

Multiplying this by 27 gives $10.999\cdots$. The calculator experienced a rounding error (calculators are not infinitely accurate) so we round this to $r = 11$, and get $6815 = 252 \cdot 27 + 11$. Of course many calculators also have mod functions which also do these things more naturally.

The Greatest Common Divisor

One last thing we need is the greatest common divisor of two numbers.

Definition. The greatest common divisor of two integers m and n is the **largest integer** d with the property that both m and n are divisible by d .

Some notation: The greatest common divisor of m and n is written **GCD(m, n)**, or often just **(m, n)**. So we have, for example, that $(24, 18) = 6$, etc. Since divisibility is such an important thing in this class, we introduce a symbol for it, too. We write $a \mid b$ to say that b is divisible by a . The expression " $a \mid b$ " is read " a **divides** b ," but means that b is divisible by a . So $6 \mid 24$ and $6 \mid 18$.

Theorem. The GCD of m and n is an integer linear combination of m and n . That is, there are integers x and y such that $\text{GCD}(m, n) = mx + ny$.

Proof. Let S be the set of all integer linear combinations of m and n . That is, $S = \{mx + ny \mid x, y \in \mathbb{Z}\}$ and let D be the smallest positive integer in S , say $D = mx_0 + ny_0$. We claim that $\text{GCD}(m, n) = D$. To see this, we let $d = \text{GCD}(m, n)$ and try to show that $D = d$. Now d is a divisor of m and n , so d will be a divisor of $mx + ny$ for every integer x and y . In particular, $d \mid D$. This means that $D \geq d$. Now we use the division algorithm to try to show that D is a common divisor of m and n : When we divide m by D we get $m = Dq + r$, where $0 \leq r < D$. Thus, $r = m - Dq = m - q(mx_0 + ny_0) = m(1 - qx_0) + n(-y_0)$. This means that r is in S . But $r < D$, and D is the smallest positive element of S . Thus, r cannot be positive. Since $r \geq 0$, this does not leave many options! It must be that $r = 0$. So we have that $m = Dq$, meaning that $D \mid m$. By a similar argument, $D \mid n$. Thus, D is a common divisor of m and n . Since d is the **greatest** common divisor of m and n , we have that $d \geq D$. We have shown that $d \geq D$ and $D \geq d$, which can only happen if $D = d$.

Corollary. If m and n are relatively prime, then there are integers x and y such that

$$mx + ny = 1.$$

We are finally in a position to prove the Lemma at the end of the second set of notes! The proof goes like this: Suppose that p is a divisor of ab . Then there are two cases to consider: first, it might be that p is a divisor of a . If so, great. If not, we have to show that p is a divisor of b . Now if p is not a divisor of a , then it has no factors in common with a (being prime, p does not have many factors!) so a and p are relatively prime. By the corollary, there are integers x and y such that $ax + py = 1$. If we multiply this by b , $abx + pby = b$. Now pby is divisible by p and we are assuming that ab is divisible by p . This means that $abx + pby$ is divisible by p , so $p \mid b$. This completes the proof of the lemma, which finally completes the proof of the Fundamental Theorem of Arithmetic.

Here is how I was taught to calculate $\text{GCD}(m, n)$ in school: factor m and n into primes. Then (m, n) will be the product of the primes that divide both m and n , raised to the smaller of the two powers for which they divide m and n separately. For example, to calculate $(11151, 3528)$, we first factor these numbers: $11151 = 3^3 \cdot 7 \cdot 59$, $3528 = 2^3 \cdot 3^2 \cdot 7^2$. So $(11151, 3528) = 3^2 \cdot 7 = 63$. This method works fine if m and n are easy to factor. However,

calculating things like $(123456, 123456789)$ can be quite hard if we must first factor the two numbers. But, in fact, we do not have to factor m and n in order to calculate (m, n) .

Theorem. $(m, n) = (m - kn, n)$ for any integer k .

Proof: Let $(m, n) = d$ and $(m - kn, n) = D$. Since d is a divisor of both m and n , it is also a divisor of $m - kn$. This means that d is a common divisor of $m - kn$ and n , so $d \leq D$. Similarly, D is a common divisor of n and $m - kn$, so D will be a common divisor of n and m . Thus, $D \leq d$, forcing $d = D$, as desired.

Obviously, it is also true that $(m, n) = (m, n - km)$. So, for example, $(123456, 123456789) = (123456, 123456789 - 1000 \cdot 123456) = (123456, 789)$. This still looks hard, but we can do the following: use a specific value of k , namely the (integer) quotient when m is divided by n . That is, if $m = qn + r$, then with $k = q$, we have $(m, n) = (r, n)$. Since $123456 = 156 \cdot 789 + 372$, $(123456, 789) = (372, 789)$. Now we iterate this procedure: $789 = 2 \cdot 372 + 45$, so $(372, 789) = (372, 45)$; $372 = 8 \cdot 45 + 12$ so $(372, 45) = (12, 45)$; $45 = 3 \cdot 12 + 9$ so $(12, 45) = (12, 9)$; $12 = 9 \cdot 1 + 3$ so $(12, 9) = (3, 9)$, and $9 = 3 \cdot 3 + 0$ so $(3, 9) = (3, 0)$. At some point, the calculation becomes easy, and we see $(123456, 123456789) = 3$. These calculations can be summarized in the following result:

Corollary (to previous theorem) If $m = qn + r$, then $(m, n) = (n, r)$.

The Euclidean Algorithm

Given m and n , how do we find x and y so that $(m, n) = mx + ny$? The following algorithm dates back at least to Euclid: In the calculation of (m, n) using the corollary above, perform the following steps:

$$m = nq_0 + r_0$$

$$n = r_0q_1 + r_1$$

$$r_0 = r_1q_2 + r_2$$

$$r_1 = r_2q_3 + r_3$$

$$\vdots$$

$$\begin{aligned}r_{k-2} &= r_{k-1}q_k + r_k \\r_{k-1} &= r_kq_{k+1} + 0.\end{aligned}$$

That is, we calculate quotients and remainders of successive earlier remainders until 0 appears as a remainder. In this case, the last nonzero remainder (r_k) will be the greatest common divisor. For example, with (123456789, 123456), the calculations will be as follows:

$$\begin{aligned}123456789 &= 123456 \cdot 1000 + 789 \\123456 &= 789 \cdot 156 + 372 \\789 &= 372 \cdot 2 + 45 \\372 &= 45 \cdot 8 + 12 \\45 &= 12 \cdot 3 + 9 \\12 &= 9 \cdot 1 + 3 \\9 &= 3 \cdot 3 + 0.\end{aligned}$$

The last remainder to appear was 3, so $(123456789, 123456) = 3$. We now use a back substitution method to write 3 as a linear combination of 123456789 and 123456:

$$\begin{aligned}3 &= 12 - 9 \\&= 12 - (45 - 12 \cdot 3) = 12 \cdot 4 - 45 \\&= (372 - 45 \cdot 8)4 - 45 = 372 \cdot 4 - 45 \cdot 33 \\&= 372 \cdot 4 - (789 - 372 \cdot 2)33 = 372 \cdot 70 - 789 \cdot 33 \\&= (123456 - 789 \cdot 156)70 - 789 \cdot 33 = 123456 \cdot 70 - 789 \cdot 10953 \\&= 123456 \cdot 70 - (123456789 - 123456 \cdot 1000) \cdot 10953 \\&= 123456 \cdot 10953070 - 123456789 \cdot 10953\end{aligned}$$

so $3 = 123456789(-10953) + 123456 \cdot 10953070$.

A smaller example: Calculate $(1239, 168)$ and express it as a linear combination of 1239 and 168. We have

$$1239 = 168 \cdot 7 + 63$$

$$168 = 63 \cdot 2 + 42$$

$$63 = 42 \cdot 1 + 21$$

$$42 = 21 \cdot 2 + 0$$

So $(1239, 168) = 21$. Now comes the back substitution:

$$\begin{aligned} 21 &= 63 - 42 \\ &= 63 - (168 - 63 \cdot 2) = 63 \cdot 3 - 168 \end{aligned}$$

$$= (1239 - 168 \cdot 7)3 - 168$$

$$= 1239 \cdot 3 - 168 \cdot 22$$

That is, $21 = 1239 \cdot 3 - 168 \cdot 22$.

This is an algorithm to learn well--we will need it several times this course.

Linear Diophantine equations

A **Diophantine** equation is an equation in which we look for solutions over the integers (or occasionally over the rational numbers). This is strange terminology since it is not the equation but the type of solution to the equation that makes it Diophantine. We have already spent a good deal of time talking about equations of the form $ax^2 + by^2 = cz^2$. The most basic kind of Diophantine equation is a linear one: $ax + by = c$. Here, the problem is to find integer solutions (x, y) to $ax + by = c$, where a, b, c are given integers. We do this as an application of the Euclidean algorithm.

In order for a solutions to exist, we need (a, b) to be a divisor of c . Letting $d = (a, b)$, assuming that $d|c$, we can consider the related problem of solving $\frac{a}{d}x + \frac{b}{d}y = \frac{c}{d}$. Clearly, the solutions to this system will be the same as the solutions to the first. Thus, for simplicity, let's assume that $(a, b) = 1$, and solve $ax + by = c$ in this case. Here is how to find one solution: use the Euclidean algorithm to find integers u and v such that $au + bv = 1$. Multiplying by c gives

$a(uc) + b(vc) = c$. Thus, $x = uc$ and $y = vc$ gives a solution.

The next problem is to get all solutions, once we have one solution. Suppose that we have two solutions (x_1, y_1) and (x_2, y_2) to $ax + by = c$. Then $a(x_1 - x_2) + b(y_1 - y_2) = 0$. We can rewrite this: $\frac{a}{b} = \frac{y_2 - y_1}{x_1 - x_2}$. Now if a and b are relatively prime, and $\frac{a}{b} = \frac{m}{n}$, then it must be that for some k , $m = ka$ and $n = kb$. You should see if you can prove this fact. In any case, this gives $y_2 - y_1 = ka$ and $x_2 - x_1 = -kb$. That is, $x_2 = x_1 - kb$ and $y_2 = y_1 + ka$. Similarly, given any solution (x_1, y_1) and any integer k , (x_2, y_2) will also be a solution where $x_2 = x_1 - kb$ and $y_2 = y_1 + ka$. We have proven the following:

Theorem. An equation of the form $ax + by = c$ will have no solutions unless c is divisible by the greatest common divisor of a and b . If $(a, b) = d$ and $d|c$, then given one solution (x_0, y_0) to the equation, the set of all solutions is the set

$$\left\{ \left(x_0 - k \frac{b}{d}, y_0 + k \frac{a}{d} \right) \mid k \in \mathbb{Z} \right\}.$$

So, for example, suppose we wish to find all integer solutions to $1239x + 168y = 105$. We do the following: First, we calculate the greatest common divisor of 1239 and 168. Previously, we calculated that $(1239, 168) = 21$. Since $21 \mid 105$, there will be integer solutions. We now look at the problem of solving $59x + 8y = 5$ (the equation that results when we divide by 21). To solve this, we first write 1 as a linear combination of 59 and 8. This was also done previously, $1 = 59 \cdot 3 - 8 \cdot 22$. A solution to the equation will be $x_0 = 5 \cdot 3 = 15$, and $y_0 = 5 \cdot (-22) = -110$. The set of all solutions is

$$\{(15 - 8k, -110 + 59k) \mid k \in \mathbb{Z}\}.$$

If we don't like the solution $(15, -110)$, we can use the form for the set of all solutions. In particular, if $k = 2$, we get the solution $(-1, 8)$.