Math 5330
Spring 2018

The Greatest Common Divisor and Euclid's Algorithm

Much of this material is in our text book. I will do things in a slightly different order, but you can find this material in sections 1.3, 2.1, and 2.3 or our book. In this set of notes, we start with one last detail in our proof of the Fundamental Theorem: We needed the following lemma.

Lemma 1 *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 ready for the proof. We need a bit more theory.

Theorem 1 (The Division Algorithm) *Given an integer n and an integer $m \neq 0$, there are unique integers q and r for which*

$$(a) \qquad n = qm + r$$

and

$$(b) \qquad 0 \leqslant r < |m|.$$

Proof: The proof has two parts: first, establishing the existence of q and r, and second, showing that they are uniquely determined by m and n. We start by reducing to the case where m and n are both positive. Suppose we accept the theorem when m and n are both positive. If $m < 0$, then given q, r, for which $n = q(-m) + r$, we have $n = (-q)m + r$, meaning that we can use the same remainder, r, and the negative of the quotient, $-q$. If $m > 0$ but $n < 0$, then first divide $-n$ by m to get $-n = qm + r$. Then $n = (-q)m - r$. This will satisfy the theorem if $r = 0$. If $r > 0$, then we subtract and add m: $n = (-q)m - m + m - r = (-q-1)m + (m-r)$, and the new quotient is $-q - 1$ and the new remainder is $m - r$. That is, if we divide -75 by 32, first divide 75 by 32 to get $75 = 2 \times 32 + 11$. Now negate: $-75 = -2 \times 32 - 11$, and modify: $-75 = -3 \times 32 + 21$, giving a quotient of -3 and a remainder of 21.

We now prove that q and r exist in the case where $m > 0$, $n > 0$. If $n < m$ then we can write $n = 0m + n$, and $q = 0, r = n$ will satisfy both (a) and (b) above. Also, if q and r do not exist, the problem HAS to be with condition (b) since we can write $n = 0m + n$, meaning condition (a) can always be satisfied if we allow large remainders. So suppose that for some $n \geqslant m$, we can't find q and r so that (b) is satisfied. We use an infinite descent on n. To that end, let $n_1 = n - m$. If q and r exist for n_1, say $n_1 = q_1 m + r_1$ and $0 \leqslant r_1 < m$, then $n = n_1 + m = (q_1 + 1)m + r_1$. That is, if q and r do not exist for n, then they can't exist for $n - m$ either. Thus, we have our infinite descent, proving that q and r exist.

Finally, for uniqueness, suppose that $n = q_1 m + r_1$ and $n = q_2 m + r_2$ where $0 \leqslant r_1 < |m|$ and $0 \leqslant r_2 < |m|$. Then a subtraction gives $0 = (q_1 - q_2)m + r_1 - r_2$, or $r_2 - r_1 = (q_1 - q_2)m$. This means that $r_2 - r_1$ is divisible by m. However, $-(|m| - 1) \leqslant r_2 - r_1 \leqslant (|m| - 1)$, and the only number in this range divisible by m is 0, so $r_2 - r_1 = 0$. Since $r_1 = r_2$, we have $q_1 m = q_2 m$, so $q_1 = q_2$ as well. This completes 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

$\frac{6815}{27} = 252.407\ldots$. This tells us that $q = 252$. Subtracting off 252, $\frac{6815}{27} - 252 = .407\ldots$ so $r = 27 \times .407\ldots$. Multiplying by 27 gives $10.999\ldots$ (calculators are only so accurate) so we round to get $r = 11$, and we have $6815 = 252 \times 27 + 11$. Of course many calculators also have mod functions, which give you q and r without this fuss.

The Greatest Common Divisor

One last thing we need for our proof of the Fundamental Theorem is the greatest common divisor of two numbers. We define the greatest common divisor of m and n to be the largest integer, d, for which both m and n are divisible by d.

We have some notation for the greatest common divisor. We often write $\gcd(m, n)$ or $GCD(m, n)$ or even (m, n). So, for example, $(24,\ 18) = 6$. Since divisibility is such an important thing in this class, we introduce a symbol for it as well. We write $a \mid b$ to mean b is divisible by a, so $6 \mid 24$. We use $a \nmid b$ to say that b is not divisible by a (so $6 \nmid 27$.) The following is a result we will use repeatedly in this class.

Theorem 2 *The greatest common divisor of two integers is a linear combination of those integers. That is, given integers m and n, not both zero, $\gcd(m, n) = mx + ny$ for some integers x and y.*

Proof: I will give the traditional proof of this result. Let S be the set of all integer linear combinations of m and n. That is, $S = \{mx + ny \,|\, x, y \in Z\}$. Let D be the smallest positive integer in S. Since D is in S, there are integers x_0 and y_0 for which $D = mx_0 + ny_0$. We claim that $\gcd(m, n) = D$. To verify the claim, let the true gcd by d. Since $d \mid m$ and $d \mid n$, it follows that d will divide everything in S. In particular, $d \mid D$. Since d and D are both positive, this means $D \geqslant d$.

Now we use the division algorithm to show that both m and n are multiples of D. We have $m = qD + r$, where $0 \leqslant r < D$. We plug in D as a combination of m and n: $m = q(mx_0 + ny_0) + r$, which can be written $r = (1 - qx_0)m + (-y_0)n$, so r is a combination of m and n, meaning $r \in S$. Since D is the smallest positive integer in S and $r < D$, r can't be positive. It also can't be negative, so $r = 0$, and $m = qD$. That is, D divides m. The same proof will show D divides n, so D is a common divisor of m and n. But d is the greatest common divisor of m and n so $D \leqslant d$. This means we have $D \leqslant d \leqslant D$, forcing $D = d$. ∎

Corolary 1 *If m and n are relatively prime, then there are integers x and y for which*

$$mx + ny = 1.$$

Finally, we may finish off the proof of the Fundamental Theorem. We prove Lemma 1 at the beginning of this set of notes. In class, I proved a slight generalization of the lemma: if

$m \mid ab$ and $\gcd(m, a) = 1$, then $m \mid b$. Here, I will settle just for the lemma itself.

Proof: Suppose p is prime and $p \mid ab$. If $p \mid a$, we are done, so suppose that $p \nmid a$. Since p is prime, p and a will be relatively prime, so for some integers x and y, $px + ay = 1$. Multiplying by b gives $pbx + aby = b$. Since both pb and ab are divisible by p, it follows that b is divisible by p as well. ■

When I was in (grade) school, I was taught to calculate $\gcd(a, b)$ as follows: Factor a and b, and for each prime divisor p common to each, use the minimum of the powers of the exponents for p as a divisor of a and b. For example, $11151 = 3^3 \times 7 \times 59$ and $3528 = 2^3 \times 3^2 \times 7^2$ so $(11151, 3528) = 3^2 \times 7 = 63$. This method works fine if a and b are easy to factor, as was always the case in school. However, calculating things like (123456, 123456789) would prove more difficult if we had to factor each of these two numbers. In fact, we do NOT have to factor a and b to determine $\gcd(a, b)$. To circumvent this, we use the following result.

Theorem 3 *For any integer k,* $\gcd(m, n) = \gcd(m - kn, n)$.

Proof: Let $\gcd(m, n) = d$ and $\gcd(m - kn, n) = D$. Since d is a divisor of both m and n, it is also a divisor of $m - kn$ so d is a common divisor of $m - kn$ and n, giving us $d \leqslant D$. Similarly, since D is a common divisor of $m - kn$ and n, it is also a divisor of the combination $(m - kn) + kn = m$. That is, D is a common divisor of m and n, so $D \leqslant d$. Thus, $D = d$. ■

Also, $\gcd(m, n) = \gcd(n, m)$ and by symmetry, $\gcd(m, n) = \gcd(m, n - km)$. So, for example, we have $\gcd(123456, 123456789) = \gcd(123456, 123456789 - 1000 \times 123456) = \gcd(123456, 789)$. This still looks hard, but we can iterate. We use a specific value for k, of course, the quotient of the two numbers. For example, since $123456 = 156 \times 789 + 372$ we have $\gcd(123456, 789) = \gcd(123456 - 156 \times 789, 789) = \gcd(372, 789)$. Next, since $789 = 2 \times 372 + 45$, $\gcd(372, 789) = \gcd(372, 45)$. We do one more iteration: $372 = 8 \times 45 + 12$ so $\gcd(372, 45) = gcd(12, 45) = 3$, because the prime factorizations of 12 and 45 are obvious. Summing up,

Corolary 2 *If $n = qm + r$, then* $\gcd(m, n) = \gcd(n, r)$.

Iterating this gives a quick algorithm, called the **Euclidean Algorithm**, for finding the greatest common divisor.

The Extended Euclidean Algorithm

One of the most important computational problems in Number Theory is to find the x and y so that $\gcd(m, n) = mx + ny$. Often, finding x and y is as important as finding $\gcd(m, n)$, especially in the case where $\gcd(m, n) = 1$. I would guess this type of thing is needed millions of times each day, in fact. One very good algorithm, which also dates back

to Euclid, is to make repeated use of the division algorithm. Euclid's approach was to make the following computations:

$$\begin{aligned} m &= nq_0 + r_0 \\ n &= r_0q_1 + r_1 \\ r_0 &= r_1q_2 + r_2 \\ r_1 &= r_2q_3 + r_3 \\ &\vdots \\ r_{k-2} &= r_{k-1}q_k + r_k \\ r_{k-1} &= r_kq_{k+1} + 0. \end{aligned}$$

That is, we calculate quotients and remainders of successive earlier remainders until 0 appears as one of the remainders. In this case, the last nonzero remainder (the term r_k in the list above) will be the greatest common divisor. For example, with $\gcd(123456789, 123456)$

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

so, as before, the greatest common divisor is 3. The point of these calculations is that we can backtrack to write the greatest common divisor as a combination of the original numbers. Starting with the second to last line, we systematically eliminate remainders, other than the last one. I will underline each remainder to be eliminated. The equation in which that remainder appeared is used to eliminate it. The calculation goes as follows:

$$\begin{aligned} 3 &= 12 - \underline{9} = 12 - \underline{(45 - 12 \times 3)} \quad \text{(solving for 9 on the previous line)} \\ &= 4 \times \underline{12} - 45 = 4\underline{(372 - 8 \times 45)} - 45 \\ &= -33 \times \underline{45} + 4 \times 372 = -33\underline{(789 - 2 \times 372)} + 4 \times 372 \\ &= 70 \times \underline{372} - 33 \times 789 = 70\underline{(123456 - 156 \times 789)} - 33 \times 789 \\ &= -10953 \times \underline{789} + 70 \times 123456 \\ &= -10953\underline{(123456789 - 1000 \times 123456)} + 70 \times 123456 \\ &= 10953070 \times 123456 - 10953 \times 123456789. \end{aligned}$$

That is, $3 = 123456x + 123456789y$ where $x = 10953070$ and $y = -10953$. As a second example, we calculated that $\gcd(11151, 3528) = 63$, let's get 63 as a combination of 11151

and 3528. We apply Euclid's algorithm, and solve for remainders:

$$\begin{aligned} 11151 &= 3 \times 3528 + 567, & 567 &= 11151 - 3 \times 3528 \\ 3528 &= 6 \times 567 + 126, & 126 &= 3528 - 6 \times 567 \\ 567 &= 4 \times 126 + 63, & 63 &= 567 - 4 \times 126 \\ 126 &= 2 \times 63 + 0. \end{aligned}$$

Next, we backtrack, starting with the last line on the right, plugging remainders from the line above to replace the underlined number:

$$\begin{aligned} 63 &= 567 - 4 \times \underline{126} \\ &= 567 - 4\underline{(3528 - 6 \times 567)} \\ &= -4 \times 3528 + 25 \times \underline{567} \\ &= -4 \times 3528 + 25\underline{(11151 - 3 \times 3528)} \\ &= 25 \times 11151 - 79 \times 3528. \end{aligned}$$

One more small example: Write $\gcd(50, 70)$ as a combination of 50 and 70. You can probably do this in your head, but I will use the Euclidean algorithm anyway. We have $70 = 1 \times 50 + 20$, $50 = 2 \times 20 + 10$, and $20 = 2 \times 10 + 0$. Backtracking, $10 = 50 - 2 \times 20 = 50 - 2(70 - 50) = 3 \times 50 - 2 \times 70$. That is, the gcd is 10 and $10 = 3 \times 50 - 2 \times 70$.

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

Linear Diophantine Equations

As I've said before, a **Diophantine** equation is an equation in which we look for solutions over the integers, rather than the real numbers. This is strange terminology because it is not the equation but the type of solution we want that makes it Diophantine, but we still call it a Diophantine equation. Examples we have already spent a good time talking about are $x^2 + y^2 = z^2$, $x^2 + ky^2 = z^2$, $x^2 + y^2 = kz^2$ and $x^4 + y^4 = z^2$. The most basic kind of Diophantine equation is the linear one, $ax + by = c$. Here, the problem is to find integer solutions (x, y) that satisfy the equation, where a, b, c are also integers. We do this as an application of the Euclidean algorithm.

In order for solutions to exist, we need $\gcd(a, b)$ to be a divisor of c. Letting $d = \gcd(a, b)$, and assuming $d \mid c$, let $c = kd$. We can find a solution as follows: Use the Euclidean algorithm to write d as a combination of a and b, say $au + bv = d$. Then multiplying by k, we get $a(ku) + b(kv) = kd = c$, so $x = ku$, $y = kv$ will be a solution to the equation.

But suppose we want ALL (integer) solutions to $ax + by = c$. We try to get the rest from the first solution. The analysis goes like this. Suppose we have two solutions, (x_1, y_1) and (x_2, y_2) to $ax + by = c$. Then a subtraction gives $a(x_1 - x_2) + b(y_1 - y_2) = 0$, and we can

write this $\frac{a}{b} = \frac{y_2 - y_1}{x_1 - x_2}$. Remember that for relatively prime numbers m, n and relatively prime numbers k, l, if $\frac{m}{n} = \frac{k}{l}$, then $m = k$ and $n = l$. If we know that m and n are relatively prime, but don't know about k and l then we can still say that for some integer, s, $k = ms$ and $l = ns$ where s is $\gcd(m, n)$. Since a/d and b/d are relatively prime, it follows that for some integer k, $y_2 - y_1 = k\frac{a}{d}$ and $x_1 - x_2 = k\frac{b}{d}$. We have proven the following result.

Theorem 4 *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 $\gcd(a, b) = d$ and $d \mid c$ then given one solution, (x_0, y_0) to the equation, the set of all solutions is*

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

For example, suppose we want all solutions to $50x + 70y = 90$. We know that $d = 10$, and $3 \times 50 - 2 \times 70 = 10$. Since $10 \mid 90$ our equation will have solutions. We multiply $3 \times 50 - 2 \times 70 = 10$ by 9 to get $27 \times 50 - 18 \times 70 = 90$, so one solution is $(x_0, y_0) = (27, -18)$. The set of all solutions is

$$\left\{ \left(27 - k\frac{70}{10}, -18 + k\frac{50}{10} \right) \mid k \in Z \right\} = \{(27 - 7k, -18 + 5k) \mid k \in Z\}.$$

If we don't like our initial solution, (27, -18), we can modify it by picking an appropriate k-value. Here, we could pick $k = 3$ to get a new solution (6, -3) or $k = 4$ to get (-1, 2). In particular, we could rewrite the set of all solutions to be $\{(-1 - 7k, 2 + 5k) \mid k \in Z\}$, or as $\{(6 + 7m, -3 - 5m) \mid m \in Z\}$. That is, we can let (x_0, y_0) be ANY solution, not just the first one we discover, and there is nothing magical about the name k or even its sign.

Euler's method for solving $ax + by = c$.

No matter how one solves $ax + by = c$, the result has to be consistant with Theorem 4. That is, assuming $d = \gcd(a, b)$ and $d \mid c$, the solution must be $\left\{ \left(x_0 - k\frac{b}{d},\ y_0 + k\frac{a}{d} \right) \mid k \in Z \right\}$ for some (x_0, y_0). The approach described above using Euclid's algorithm is perfectly fine, but it tends to produce rather large initial solutions (x_0, y_0). Euler introduced a method that, in general, gives smaller initial solutions. I will explain Euler's method by way of several examples. As a first example, let's redo $50x + 70y = 90$. Euclid's method gave (27, -18) as a solution. Euler's approach is to take $ax + by = c$ and solve for the variable with the smallest coefficient. In $50x + 70y = 90$, he would solve for x: $x = \frac{90 - 70y}{50}$. It is hard not to divide through by 10, but let's resist since it is not necessary. Next, we divide both 90 and 70 by 50 and rewrite this $x = 1 - y + \frac{40 - 20y}{50}$. Since x and $1 - y$ are both integers, the fraction must also be an integer, call it z and write $z = \frac{40 - 20y}{50}$ so $20y + 50z = 40$.

Now we repeat: $y = \dfrac{40-50z}{20} = 2 - 2z - \dfrac{10z}{20}$. Again, the remaining fraction must be an integer, which I will call w. We have $10z = 20w$, an equation so simple we can write down a solution: $w = 0$, $z = 0$ being the simplest. We now backtrack, sticking these values in previous equations until we have values for x and y. We have $y = 2 - 2z - \frac{10z}{20} = 2$, and $x = 1 - y + \frac{40-20y}{50} = -1$. So Euler's method gave the much smaller solution $x = -1$, $y = 2$.

As a second example, $3528x + 11151y = 315$. We found $3528x + 11151y = 63$ has $x = -79$, $y = 25$ as a solution, and $315 = 5 \times 63$ so $3528x + 11151y = 315$ has $x = -395$, $y = 125$ as a solution from Euclid's method. Using Euler's method, $x = \dfrac{315 - 11151y}{3528} = -3y + \dfrac{315 - 567y}{3528}$. We set the fraction equal to z, giving $3528z + 567y = 315$, so $y = \dfrac{315 - 3528z}{567} = -6z + \dfrac{315 - 126z}{567}$. Set this fraction equal to w, and $567w + 126z = 315 \rightarrow z = \dfrac{315 - 567w}{126} = 2 - 4w + \dfrac{63 - 63w}{126} = 2 - 4w + \dfrac{1-w}{2}$. We could continue, setting $\frac{1-w}{2} = s$, say, but it is easy enough to find a solution to the z/w equation, $w = 1, z = -2$. Backtracking, $y = -6z + \frac{315-126z}{567} = 13$, $x = -3y + \frac{315-567y}{3528} = -41$. Euler's method gave the solution $x = -41$, $y = 13$. Using Theorem 4, the general solution is $\{(-41 + 177t, 13 - 56t) \mid t \in Z\}$.

As a final example, suppose we want all solutions to $11x + 26y = 500$. Since $\gcd(11, 26) = 1$, with Euclid's algorithm we would first find a solution to $11x + 26y = 1$. We have $26 = 2 \times 11 + 4$, $11 = 2 \times 4 + 3$, $4 = 1 \times 3 + 1$, $3 = 3 \times 1 + 0$. Backtracking, $1 = 4 - 3 = 4 - (11 - 2 \times 4) = 3 \times 4 - 11 = 3(26 - 2 \times 11) - 11 = -7 \times 11 + 3 \times 26$. Since $11x + 26y = 1$ has solution $x = -7$, $y = 3$, we multiply by 500 to get a solution of $x = -3500$, $y = 1500$ for $11x + 26y = 500$. The general solution will be $\{(-3500 + 26t, 1500 - 11t) \mid t \in Z\}$.

With Euler's approach, we don't go through the equation $11x + 26y = 1$, but start directly with $11x + 26y = 500$. We solve for x, $x = \dfrac{500 - 26y}{11} = 45 - 2y + \dfrac{5 - 4y}{11}$. We set $z = \dfrac{5 - 4y}{11}$, so $11z + 4y = 5$, and iterate this approach, in each case, solving for the variable with smallest coefficient. We have $y = \dfrac{5 - 11z}{4} = 1 - 2z + \dfrac{1 - 3z}{4}$, let $w = \dfrac{1 - 3z}{4} \rightarrow 4w + 3z = 1$, and I am tempted to stop here, since there is an obvious solution to this last equation, $w = 1$, $z = -1$. It would not have hurt (other than taking a little more time) to do one more step, say $z = \dfrac{1 - 4w}{3} = -w + \dfrac{1 - w}{3}$, $u = \dfrac{1 - w}{3} \rightarrow 3u + w = 1$. As soon as one of the variables has 1 as a coefficient, you can let the other variable (u here) be zero. This would lead back to $w = 1$, $z = -1$ again. In any case, $y = \frac{5-11z}{4} = 4$, $x = 45 - 2y + \frac{5-4y}{11} = 36$, so one solution is $x = 36$, $y = 4$, leading to the solution set $\{(36 + 26t, 4 - 11t) \mid t \in Z\}$. Note that $x = 10$, $y = 15$ is also a solution to $11x + 26y = 500$, and this might be considered "smaller" than $x = 36$, $y = 4$. Euler's method is not guaranteed to produce the "best" solution, but it usually beats out the Euclid approach.