A Few Notes on the Division Algorithm (WIP)

David Meyer

dmm@{1-4-5.net,uoregon.edu}

Last update: October 1, 2019

1 Introduction

TBD

2 Fields and the Division Algorithm

Definition 2.1. Let f and g be polynomials in $F[x]$. Then we say that f divides g, or g is divisible by f if there is a polynomial h with $g = fh$.

For integers, the greatest common divisor of two integers a and b is the largest integer dividing both a and b. This definition doesn't quite work for polynomials. In particular, while we cannot talk about "largest" polynomial in the same manner as we do for integers, we can talk about the degree of a polynomial.

Recall that the degree of a nonzero polynomial f is the largest integer m for which the coefficient a_m of x^m is nonzero. For example, if $f(x) = a_n x^n + \ldots + a_1 x + a_0$ with $a_n \neq 0$, then the degree of $f(x)$, written $\deg f$, is n. The degree function allows us to measure size of polynomials.

There is one extra complication with this definition of $\deg f$. Consider for example that any polynomial of the form ax^2 with $a \neq 0$ divides x^2 and x^3. Thus, there isn't a unique polynomial of highest degree that divides a pair of polynomials. To get around this uniqueness problem the convention is to consider the *monic* polynomials, that is, those polynomials whose leading coefficient is 1. For example, x^2 is the monic polynomial of degree 2 that divides both x^2 and x^3, while $5x^2$ is not monic.

Finally, I'll use ab to represent $a \cdot b$ where the meaning is unambiguous.

Definition 2.2. Let f and g be polynomials over F, where f and g are not both zero. Then a greatest common divisor of f and g, $\gcd(f, g)$, is a monic polynomial of largest degree that divides both f and g.

Unfortunately, there is a problem with Definition 2.2 which has to do with uniqueness. In particular, could there be more than one greatest common divisor of a pair of polynomials? As we will se below, the answer turns out to be no.

Note that the main reason for assuming that the coefficients of our polynomials lie in a field is to ensure that the division algorithm is valid. Why? Because division is well behaved in integral domains (where there are no zero divisors), and because every field is an integral domain and every finite integral domain is a field. But why is that?

Theorem 2.1. *Every field F is an integral domain.*

Proof: Recall that if F is a field then each non-zero $r \in F$ has an inverse r^{-1}. So suppose $r, s \in F$ and $r \neq 0$ such that $r \cdot s = 0$. Then the claim is that $s = 0$. Why? Consider

$$\begin{array}{rcll} r \cdot s & = & 0 & \# \text{ assumption with } r \neq 0 \\ & \Rightarrow & r^{-1} \cdot (r \cdot s) = r^{-1} \cdot 0 & \# \text{ multiply both sides by } r^{-1} \\ & \Rightarrow & r^{-1} \cdot (r \cdot s) = 0 & \# \ x \cdot 0 = 0 \\ & \Rightarrow & (r^{-1} \cdot r) \cdot s = 0 & \# \text{ multiplication is associative} \\ & \Rightarrow & 1 \cdot s = 0 & \# \ r^{-1} \cdot r = 1 \\ & \Rightarrow & s = 0 & \# \ 1 \cdot x = x \end{array} \tag{1}$$

So r is not a zero divisor. But every non-zero element r of the field F has an inverse (r is a "unit") so F has no zero divisors and is by definition an integral domain. $\square$

Theorem 2.2. *Every finite integral domain is a field.*

Proof: The proof is based on the fact that since R is an integral domain it has a cancellation law (or equivalently, R has no zero divisors). Having a cancellation law means that

$$ab = ac \implies b = c \tag{2}$$

To see why any finite integral domain R is a field, consider $R = \{r, r^2, r^3, \ldots, r^n\}$ where $r^k \neq 0$ for $1 \leq k \leq n$. Since R is finite we will have $r^k = r^l$ for some k and l such that $k > l$. Then

$$
\begin{array}{rcll}
r^k & = & r^l & \# \ R \text{ is a finite integral domain} \\
& \Rightarrow & r \cdot r^{k-1} = r \cdot r^{l-1} & \# \text{ factor out } r \\
& \Rightarrow & r^{k-1} = r^{l-1} & \# \text{ use cancellation law (cancel } r\text{, Equation 2)} \\
& \Rightarrow & r \cdot r^{k-2} = r \cdot r^{l-2} & \# \text{ factor out } r \\
& \Rightarrow & r^{k-2} = r^{l-2} & \# \text{ use cancellation law (cancel } r\text{, Equation 2)} \\
& \vdots & & \# \text{ iterate } l-1 \text{ times} \\
& \Rightarrow & r^{k-l+1} = r^1 & \# \ldots \\
& \Rightarrow & r \cdot r^{k-l} = r \cdot r^0 & \# \text{ factor out } r \\
& \Rightarrow & r^{k-l} = r^0 & \# \text{ use cancellation law (cancel } r\text{, Equation 2)} \\
& \Rightarrow & r^{k-l} = 1 & \# \ r^0 = 1
\end{array}
$$

So $r^{k-l} = 1$. If $k - l = 1$ then r a unit since $r^{k-l} = r^1 = 1$ so r^{-1} is $\frac{1}{r}$. Otherwise $k - l > 1$ and $r^{k-l} = 1 \Rightarrow r^{k-l-1} = \frac{1}{r}$. So $r^{-1} = r^{k-l-1}$ and every $r \neq 0 \in R$ has an inverse. Thus every non-zero $r \in R$ is a unit and so R is a field. $\square$

So every field is an integral domain and every finite integral domain is a field and so division behaves reasonably in these cases.

A few conventions: set $\deg 0 = -\infty$ and $-\infty + -\infty = -\infty$ and $-\infty + n = -\infty$ for $n \in \mathbb{Z}$.

Lemma 2.1. Let F be a field and let f and g be polynomials over F. Then $\deg fg = \deg f + \deg g$.

Proof: If either $f = 0$ or $g = 0$, then the equality $\deg fg = \deg f + \deg g$ is true by our convention above. So, suppose that $f \neq 0$ and $g \neq 0$. Write $f = a_n x^n + \ldots + a_0$ and $g = b_m x^m + \ldots + b_0$ with $a_n \neq 0$ and $b_m \neq 0$. Therefore, $\deg f = n$ and $\deg g = m$. The definition of polynomial multiplication yields

$$fg = (a_n b_m)x^{n+m} + (a_n b_{m-1} + a_{n-1} b_m)x^{n+m-1} + \ldots + a_0 b_0$$

Since every field is an integral domain and since the coefficients come from a field we know that there are no zero divisors and so we can conclude that $a_n b_m \neq 0$, and so $\deg fg = n + m = \deg f + \deg g$, as desired. $\square$

Definition 2.3. The characteristic of a field F, $\text{char}(F)$, is the smallest positive integer n such that

$$n \cdot 1 = \underbrace{1 + 1 + \cdots + 1}_{n \text{ times}} = 0$$

The characteristic has all kinds of interesting properties. For example

Lemma 2.2. Let F be a field. Then

1. If $\text{char}(F) > 0$ then $\text{char}(F)$ is prime
2. If F is a finite field then $\text{char}(F) > 0$
3. If F is a finite field then $\text{char}(F)$ is prime

Proof:

1. Assume that n is not prime so that there exists a nontrivial factorization of $\text{char}(F) = n = p \cdot q$. Then

$$\begin{array}{rcll} 0 & = & n \cdot 1 & \# \text{ definition of } \text{char}(F) \\ & = & (p \cdot q) \cdot 1 & \# \ n = p \cdot q \\ & = & p \cdot (q \cdot 1) & \# \text{ multiplication is associative} \\ & = & (p \cdot 1) \cdot (q \cdot 1) & \# \ p = p \cdot 1 \end{array}$$

   So $(p \cdot 1) \cdot (q \cdot 1) = 0$. But we know that F is a field and fields have no zero divisors so either $(p \cdot 1) = 0$ or $(q \cdot 1) = 0$. But this contradicts the assumption that $\text{char}(F)$ was the smallest positive integer n such that $n \cdot 1 = 0$ since $p, q \neq 0$ and $n = p \cdot q$. So $n = \text{char}(F)$ must have been prime.

   Another way to think about this proof is to notice that if F is a field of characteristic n then the elements $\{0, 1, 2, \ldots, , n-1\}$ of F obey the rules for addition and multiplication modulo n. Therefore there is a copy of $\mathbb{Z}_n$ inside of F. Since $\mathbb{Z}_n$ has zero divisors when n is not prime (and a field is an integral domain and so has no zero divisors, see Theorem 2.1), it follows that the characteristic of a field must be prime.

   Thus every finite field F must have characteristic p for some prime p and the elements $\{0, 1, 2, \ldots, p-1\}$ form a copy of $\mathbb{Z}_p$ inside of F. This copy of $\mathbb{Z}_p$ is known as the *prime subfield of F*.

2. Since F is finite not every element can be unique. In particular, consider a finite field $F = \{f, f^2, f^3, \ldots, f^n\}$ where $f^k \neq 0$ for $1 \leq k \leq n$. Since F is finite we will have $(f^k \cdot 1) = (f^l \cdot 1)$ for some k and l where $k > l$. Then $(f^k - f^l) \cdot 1 = 0$ and so $\text{char}(F) > 0$.

3. Immediate by 1. and 2.

□

Proposition 2.1. (Division Algorithm) Let F be a field and let f and g be polynomials over F with f nonzero. Then there are unique polynomials q and r with $g = qf + r$ with $\deg r < \deg f$.

Proof: Let $\mathcal{S} = \{t \in F[x] \mid t = g - qf \text{ for some } q \in F[x]\}$. Now, we know $\mathcal{S} \neq \emptyset$ since $g \in \mathcal{S}$ (see Equation 4 below). We also know by the well ordering property of the integers [1] that there is a polynomial r of least degree in $\mathcal{S}$. Then by definition there is a $q \in F[x]$ with $r = g - qf$ and so $g = qf + r$.

We need to show that $\deg r < \deg f$. So suppose that $\deg r \geq \deg f$, say with $n = \deg f$ and $m = \deg r$. If $f = a_n x^n + \ldots + a_0$ and $r = r_m x^m + \ldots + r_0$ with $a_n \neq 0$ and $r_m \neq 0$, then by the method of long division of polynomials, we see that $r = (r_m a_n^{-1})x^{m-n} f + r'$ with $\deg r' < m = \deg r$. But then

$$\begin{array}{rcll} g & = & qf + r & \# \; g \in \mathcal{S} \text{ and well ordering principle} \\ & = & qf + (r_m a_n^{-1})x^{m-n} f + r' & \# \; r = (r_m a_n^{-1})x^{m-n} f + r' \\ & = & (q + r_m a_n^{-1} x^{m-n})f + r' & \# \text{ factor out } f \\ & \Longrightarrow & r' \in \mathcal{S} & \# \; r' = g - (q + r_m a_n^{-1} x^{m-n})f \in \mathcal{S} \end{array}$$

Since $\deg r' < \deg r$, $r' \in \mathcal{S}$ contradicts our choice of r (the assumption that r was the polynomial of least degree in $\mathcal{S}$). Therefore $\deg r \geq \deg f$ is false so $\deg r < \deg f$ is true and this proves existence of q and r.

We still have to show the uniqueness of q and r. To do this, suppose that $g = qf + r$ and $g = q'f + r'$ for some polynomials $q, q', r, r' \in F[x]$, and with both $\deg r < \deg f$ and $\deg r' < \deg f$. Then $qf + r = q'f + r'$, so $(q - q')f = r' - r$. Taking degrees and using Lemma 2.1, we have

$$\deg(q - q') + \deg f = \deg(r' - r) \tag{3}$$

Since $\deg r < \deg f$ and $\deg r' < \deg f$, we have $\deg(r' - r) < 0$. However, if $\deg(q' - q) \geq 0$, we get a contradiction of Equation 3. In particular, by our convention the only way for this to happen is for $\deg(q' - q) = \deg(r' - r) = -\infty$, and the only way for that to happen is if $q' - q = 0 = r' - r$. So $q' = q$ and $r' = r$, which shows uniqueness. $\square$

Proposition 2.2. Let F be a field and let f and g be polynomials over F, with not both f and g zero. Then $\gcd(f, g)$ exists and is unique. Furthermore, there are polynomials h and k with $\gcd(f, g) = hf + kg$.

Proof: Here we have to show that $\gcd(f, g)$ exists and that it is unique. To show that $\gcd(f, g)$ exists, let $\mathcal{S} = \{hf + kg : h, k \in F[x]\}$. Then $\mathcal{S}$ contains nonzero polynomials f and g (at least), since for some $f' \in \mathcal{S}$

$$\begin{array}{rcll} f' & = & hf + kg & \# \text{ definition of elements of } \mathcal{S} \\ & = & 1 \cdot f + 0 \cdot g & \# \text{ let } h = 1 \text{ and } k = 0 \\ & = & f & \# \text{ so } f \in \mathcal{S} \end{array} \tag{4}$$

Likewise for g (let $h = 0$ and $k = 1$). So $f, g \in \mathcal{S}$. But this means that there is a nonzero polynomial $d \in \mathcal{S}$ of smallest degree by the well ordering principle [1], so we can write $d = hf + kg$ for some $h, k \in F[x]$. By dividing by the leading coefficient of d, we may assume that d is monic without changing the condition that d is the polynomial of least degree in $\mathcal{S}$.

Now the claim is that $d = \gcd(f, g)$. To show that d is a common divisor of f and g, first consider f. By the division algorithm, we can write f as $f = qd + r$ for some polynomials q and r with $\deg r < \deg d$. Then

$$\begin{array}{rcll} f & = & qd + r & \# \ f \in \mathcal{S} \text{ by Equation 4}, \mathcal{S} \text{ has a division algorithm} \\ & \implies & r = f - qd & \# \text{ solve for } r \\ & \implies & r = f - q(hf + kg) & \# \ d = hf + kg \\ & \implies & r = f - qhf - qkg & \# \text{ arithmetic} \\ & \implies & r = (1 - qh)f + (-qk)g & \# \ F[x] \text{ is closed under } */+ \text{ so } r \text{ is of the form } hf + kg \\ & \implies & r \in \mathcal{S} & \# \ r \text{ is of the form } hf + kg \text{ so } r \in \mathcal{S} \end{array} \tag{5}$$

So $r \in \mathcal{S}$. Now, if $r \neq 0$ we have contradiction since we assumed that $\deg(r) < \deg(d)$ and that d was the polynomial of lowest degree in $\mathcal{S}$. Thus r must equal 0, which shows that $f = qd$, and so d divides f. Similarly, d divides g and so d is a common divisor of f and g.

If e is any other common divisor of f and g, then e divides any combination of f and g; in particular, e divides $d = hf + kg$. This forces $\deg(e) \leq \deg(d)$ by Lemma 2.1. Thus, d is the monic polynomial of largest degree that divides f and g, so d is a greatest common divisor of f and g. This shows that $\gcd(f, g)$ exists.

To show that $\gcd(f, g)$ is unique, suppose that d and d' are both monic common divisors of f and g of largest degree. By Equation 5, we can write both d and d' as combinations of f and g. In addition, the argument above shows that d divides d' and vice-versa. If $d' = ad$ and $d = bd'$, then $d = bd' = bad$. Taking degrees shows that $\deg(ba) = 0$, which means that a and b are both constants. However since d and d' are monic, for $d' = ad$ to be monic it must be that $a = 1$. Thus, $d' = ad = d$. This shows that the greatest common divisor is unique, completing the proof.

3 Acknowledgements

References

[1] MIT OCW. The Well Ordering Principle. https://ocw.mit.edu/courses/electrical-engineering-and-computer-science/6-042j-mathematics-for-computer-science-spring-2015/readings/MIT6_042JS15_Session3.pdf, 2015. [Online; accessed 23-Sep-2019].