Math 75 notes, Lecture 23

P. Pollack and C. Pomerance

A little review

Last time we introduced the *formal derivative* of a polynomial: If F is a field and $f(x) = \sum_{i=0}^{d} a_i x^i \in F[x]$, we put

$$D(f(x)) = \sum_{i=1}^{d} i a_i x^{i-1}.$$

We saw that with this definition, we have number of results familiar from calculus. For example, the formal derivative is an F-linear map from $F[x]$ to $F[x]$ and satisfies the product rule,

$$D(fg) = fD(g) + gD(f).$$

Moreover, we completely characterized those polynomials f for which $D(f) = 0$: Of course constant polynomials always have derivative zero, but we saw that if F has characteristic p then there are infinitely many more examples; in that case, the polynomials with derivative zero are exactly the polynomials of the form $u(x^p)$, where $u(x) \in F[x]$. Our last theorem from last time gave us our first application of the derivative: if F is a finite field, then f and $D(f)$ are relatively prime exactly when f is squarefree.

Here we present a different application of the formal derivative, one which has a very different flavor. We use it to attack a polynomial version of the (in)famous problem known as Fermat's Last Theorem.

Fermat's last theorem for polynomials

Let's recall the statement of Fermat's last theorem. That theorem concerns solutions to the equation $x^n + y^n = z^n$. When $n = 2$, the positive integer solutions to this equation are known as Pythagorean triples, and as you probably learned long before you got to Dartmouth, these solutions exactly describe the possible sides of right triangles. What about when $n > 2$? In this case there are still solutions, but none very satisfying; e.g., we can take $x = 0$ and set $y = z$, but that is hardly interesting. Fermat's last theorem says that if we want non-boring solutions we are out of luck: if $n \geq 3$, then $x^n + y^n = z^n$ has no solutions in nonzero integers x, y, z.

What if we look at the same equation $x^n + y^n = z^n$, but instead of asking for integer solutions, we ask for polynomial solutions. In that case we are led to the following theorem:

Theorem 1. *Let F be a field of characteristic 0 and let $n \geq 3$ be an integer. Suppose that f, g, and h are three nonzero polynomials in $F[x]$ satisfying $f^n + g^n = h^n$, and that f, g, and h are relatively prime, i.e., there is no irreducible polynomial $p(x)$ dividing all three of f, g, and h. Then f, g, and h are constant.*

The restriction to fields of characteristic 0 leaves out finite fields, which have been the main object of study in this course. But as you'll see in your homework, one can prove a (slightly more complicated) version of the same theorem over these fields. Also, it is not a significant restriction to insist that f, g, and h are coprime, since we can always reduce to this case; indeed, if $f^n + g^n = h^n$, and d is the monic polynomial of largest degree dividing f, g, and h, then $(f/d)^n + (g/d)^n = (h/d)^n$, and now $f/d, g/d$, and h/d are coprime.

You should think of Theorem 1 as, morally, making the same sort of claim as the usual Fermat's Last Theorem; both say that there are no 'interesting' solutions to $x^n + y^n = z^n$ when $n \geq 3$. In the polynomial version, the class of 'uninteresting' solutions is a bit larger though, and includes not only the cases when one of the three summands is zero, but also the case when all three are constant. For example, when $n = 3$ and $F = \mathbb{C}$, there is nothing at all exciting or surprising about the solution

$$1^3 + 1^3 = (\sqrt[3]{2})^3.$$

Finally, notice that just as in the integer case, it is essential that we require $n \geq 3$ here: There are many 'interesting' solutions to $f^2 + g^2 = h^2$. In fact, if $a(x)$ is any polynomial in $F[x]$, then it is easily checked that

$$(a(x)^2 - 1)^2 + (2a(x))^2 = (a(x) + 1)^2.$$

Despite the surface similarities between Fermat's last theorem and our polynomial analogue, there is a profound disparity in the difficulty of their proofs. Fermat's last theorem resisted attack for hundreds of years before finally being dispensed with (in the mid 90s) by Andrew Wiles. The proof spans hundreds of pages and uses some of the most sophisticated tools of modern number theory.

By contrast, the proof of the polynomial Fermat's last theorem was known already to 19th century mathematicians. Here we describe a simple, modern proof.

Mason's theorem

Our main tool is an inequality due to Mason. To state his theorem we need a preliminary definition. Let F be a field and let f be a nonzero polynomial over F. We define

$$R(f) = \prod_{p|f} p,$$

where the product runs over the monic irreducibles $p(x)$ which divide f. (This is called the *radical* of f.) Concretely, if the unique factorization of f has the form $up_1^{e_1} \cdots p_r^{e_r}$ (where u is a unit and the p_i are distinct monic irreducibles), then $R(f) = p_1 \cdots p_r$. In words, $R(f)$ is the monic, squarefree divisor of f of largest degree.

Theorem 2 (Mason's theorem)**.** *Let F be any field. Suppose $f, g, h \in F[x]$ are nonzero and that f, g, and h are coprime, i.e., that there is no irreducible dividing all of f, g, and h. If f, g, and h satisfy the equation $f + g = h$, then either $D(f) = D(g) = D(h) = 0$, or*

$$\max\{\deg f, \deg g, \deg h\} \leq \deg R(fgh) - 1.$$

Remark. The hypothesis that no prime divides all three of f, g, and h is equivalent to the hypothesis that $\gcd(f, g) = \gcd(f, h) = \gcd(g, h) = 1$; this is because $f + g = h$, so that any prime dividing two of f and g necessarily divides the third. We'll use this a priori stronger-seeming hypothesis in our argument below.

The proof of Theorem 2 we give here is due to N. Snyder and was discovered while he was a high school student! We begin with an easy lemma:

Lemma 1. *Let f be a nonzero polynomial in $F[x]$. Then $f/\gcd(f, D(f))$ divides $R(f)$. As a consequence,*

$$\deg \gcd(f, D(f)) \geq \deg f - \deg R(f).$$

Proof. Write out the unique factorization of f as $f = up_1^{e_1} \cdots p_r^{e_r}$, so that $R(f) = p_1 \cdots p_r$. For each $1 \leq i \leq r$, we can write $f = p_i^{e_i} q_i$ (where q_i just collects all the other terms in the factorization) and apply the product rule to obtain that

$$\begin{aligned} D(f) &= p_i^{e_i} D(q_i) + q_i D(p_i^{e_i}) \\ &= p_i^{e_i} D(q_i) + q_i e_i p_i^{e_i - 1} D(p_i) \\ &= p_i^{e_i - 1}(p_i D(q_i) + e_i q_i D(p_i)). \end{aligned}$$

So $p_i^{e_i - 1}$ divides both $D(f)$ and f, and so must divide $\gcd(f, D(f))$. Since $p_i^{e_i}$ is the highest power of p_i dividing f, the p_i-part of the gcd of f and $D(f)$ is either $p_i^{e_i - 1}$ or $p_i^{e_i}$. This implies that

$$f/\gcd(f, D(f)) \mid up_1 \cdots p_r = uR(f).$$

Since u is a unit, we must also have $f/\gcd(f, D(f))$ dividing $R(f)$.

The final claim of the lemma follows immediately: We use the general rule that if $a \mid b$ and $b \neq 0$, then the degree of a is bounded by the degree of b. In our case, we get that

$$\deg\left(f/\gcd(f, D(f))\right) \leq \deg R(f).$$

But the degree of the left hand side is just the difference of the degrees of the numerator and denominator. Writing that out and rearranging gives the last claim. □

Proof of Theorem 2. We start with the equation $f + g = h$ and differentiate to obtain $D(f) + D(g) = D(h)$. Multiplying the first of these equations by $D(g)$ and the second by g, and subtracting the second from the first, we find

$$fD(g) - gD(f) = hD(g) - D(h)g.$$

This equality will be useful momentarily.

For now, observe that since $\gcd(f, D(f))$ divides f and $D(f)$, it must be that $\gcd(f, D(f))$ divides $fD(g) - gD(f)$. Similarly, $\gcd(g, D(g))$ also divides $fD(g) - gD(f)$. Finally, $\gcd(h, D(h))$ also divides $fD(g) - gD(f)$, because $fD(g) - gD(f) = hD(g) - D(h)g$. Since any two of f, g, and h are relatively prime, any two of $\gcd(f, D(f))$, $\gcd(g, D(g))$, and $\gcd(h, D(h))$ are also relatively prime, and so (by unique factorization) we have

$$\gcd(f, D(f)) \gcd(g, D(g)) \gcd(h, D(h)) \mid fD(g) - gD(f). \tag{1}$$

We will soon use this divisibility relation to compare the degrees of the right and left-hand sides, but before we can do this, we need to know that the right-hand side is nonzero (so that it has a degree).

Suppose, on the contrary, that the right-hand side vanishes. Then $fD(g) = gD(f)$, and so f divides $gD(f)$. Since f and g are coprime, it follows that f divides $D(f)$. But this is only possible if $D(f) = 0$. The same argument shows that in this case $D(g) = 0$. But we already remarked above that $D(f) + D(g) = D(h)$, so we have $D(f) = D(g) = D(h) = 0$, which is one of the possibilities allowed for in our statement of Mason's theorem.

So we can assume that $fD(g) - gD(f)$ is nonzero. Then we are justified in saying that the degree of the left-hand side of (1) is at most the degree of the right-hand side of (1). From the definition of the derivative, it's easy to get an upper bound on the degree of the right-hand side; we have

$$\deg\left(fD(g) - gD(f)\right) \leq \deg f + \deg g - 1.$$

Lemma 1 gives us a lower bound on the left-hand side. Breaking up the degree of the product as the sum of the degrees, we see the left-hand side of (1) ist at least

$$\deg f + \deg g + \deg h - \deg R(f) - \deg R(g) - \deg R(h).$$

Combining this with the upper bound found above and rearranging, we get

$$\begin{aligned}\deg h &\leq \deg R(f) + \deg R(g) + \deg R(h) - 1\\ &= \deg R(f)R(g)R(h) - 1 = \deg R(fgh) - 1.\end{aligned}$$

(We have that $R(fgh) = R(f)R(g)R(h)$, since no two of f, g, and h share an irreducible factor.)

So we've proved 1/3 of what we wanted to prove; we've shown that $\deg h$ is bounded by the right-hand expression in Mason's theorem, and we need to do the same for $\deg f$ and $\deg g$. But this follows immediately if we apply what we've already proved not to $f + g = h$, but to the equivalent rearranged equations $h + (-f) = g$ and $h + (-g) = f$. □

Back to Fermat's last theorem

Let's now prove Fermat's last theorem for polynomials. So suppose $n \geq 3$, and that f, g, and h are nonzero polynomials over F (now assumed to be of characteristic zero) with $f^n + g^n = h^n$, and with no irreducible dividing all three of f, g, and h. We have to show that all of f, g, and h are constant.

Mason's theorem gives us only two options, either $D(f^n) = D(g^n) = D(h^n) = 0$, or

$$\max\{\deg f^n, \deg g^n, \deg h^n\} \leq \deg R(f^n g^n h^n) - 1. \tag{2}$$

Since F has characteristic zero, in the first case f^n, g^n, and h^n are all constant polynomials. But then f, g, and h must also all be constant, as we wanted to show.

Let's show that the second option leads to a contradiction. First notice we can easily bound the right-hand side of (2) from above. Since the radical of a polynomial depends only on the irreducibles dividing the polynomial and not the exponents to which they occur,

$$R(f^n g^n h^n) = R(fgh).$$

But $R(fgh)$ divides fgh, and so an upper bound for the right hand side of (2) is

$$\deg f + \deg g + \deg h - 1. \tag{3}$$

Now we obtain a contradictory lower bound for the left-hand side of (2). Since the maximum of any three numbers is at least their average, the left hand side of (2) is at least

$$\frac{\deg f^n + \deg g^n + \deg h^n}{3} = \frac{n}{3}(\deg f + \deg g + \deg h) \geq \deg f + \deg g + \deg h,$$

since $n \geq 3$, contradicting (3).

What about the integers?

Ok, that wasn't trivial, but it wasn't hundreds of pages of work either. So it seems worthwhile to ask if there is an analogue of Mason's theorem for integers; this then might lead to a much shorter proof of Fermat's last theorem than the one currently known.

The answer seems to be 'yes and no.' Yes, there is an analogous statement for integers, known as the 'abc conjecture,' and this statement does indeed have many marvelous consequences (including a proof of Fermat's last theorem for all sufficiently large exponents). But no, it isn't a theorem! In fact, there does not currently appear to be a plan on the mathematical table that has much hope of settling it. Note that the argument we gave above for Mason's theorem fails to even get off the ground, as it relies heavily on the properties of the formal derivative, and no such tool is available for studying integer arithmetic.