Second proof of the irreducibility of the cyclotomic polynomials

December 4, 2013

Let me start by recalling some highlights of what we did last time. Fix an algebraic closure $\overline{\mathbb{Q}}$ of $\mathbb{Q}$.

- The group of n^{th} roots of unity in $\overline{\mathbb{Q}}$ is cyclic of order n.
- A generator for this group is called a primitive n^{th} root of unity, and is denoted ζ_n. There are exactly $\phi(n)$ primitive n^{th} roots of unity.
- Theorem:
  1. The degree $[\mathbb{Q}(\zeta_n) : \mathbb{Q}]$ equals $\phi(n)$.
  2. The natural injective homomorphism $Gal(\mathbb{Q}(\zeta_n)/\mathbb{Q}) \to (\mathbb{Z}/n\mathbb{Z})^\times$ is an isomorphism.
  3. The n^{th} cyclotomic polynomial $\Phi_n(X)$ is irreducible.

Recall that the statements (1,2,3) in the last bullet point are, furthermore, essentially equivalent. We gave a proof of 1 (and hence all of them) in the special case $n = p$ prime last time. Today we're going to give a proof of 2 (and hence all of them) for arbitrary n.

Even though these statements (1,2,3) are equivalent, there's a sense in which 2 is better than 1. I mean, 1 just says that the cardinalities of the groups $Gal(\mathbb{Q}(\zeta_n)/\mathbb{Q})$ and $(\mathbb{Z}/n\mathbb{Z})^\times$ are the same, whereas 2 provides a canonical isomorphism between these groups. So we should learn something by directly proving 2, as we're about to.

To begin with, let's recall the definition of the homomorphism $Gal(\mathbb{Q}(\zeta_n)/\mathbb{Q}) \to (\mathbb{Z}/n\mathbb{Z})^\times$. Essentially, it records the effect of a Galois automorphism σ on the n^{th} roots of unity. More precisely, $\sigma \in Gal(\mathbb{Q}(\zeta_n)/\mathbb{Q})$ gets sent to $k \in (\mathbb{Z}/n\mathbb{Z})^\times$ if and only if $\sigma(\zeta) = \zeta^k$ for all n^{th} roots of unity ζ. It's also equivalent to say that this happens just for $\zeta = \zeta_n$.

So, what we need to prove is that given any $k \in (\mathbb{Z}/n\mathbb{Z})^\times$, we can find a Galois automorphism of $\mathbb{Q}(\zeta_n)/\mathbb{Q}$ which raises ζ_n to the k^{th} power.

Let me start with a remark, which will not actually be a part of the formal proof, but will serve as motivation. The remark is that we can do this pretty easily when $k = -1$. (We can also do it when $k = 1$ by taking $\sigma = id$, but that's not interesting).

How is this? Well, from an algebraic perspective it's no easier to handle $k = -1$ than any other $k \neq 1$. But look: there's this analytic field $\mathbb{C}$ which contains ζ_n and also carries a certain nontrivial automorphism called "complex conjugation". Complex conjugation sends ζ to ζ^{-1} whenever ζ lies on the unit circle, so in particular whenever ζ is a root of unity. It follows that complex conjugation restricts to an automorphism of $\mathbb{Q}(\zeta_n)$ which hits $k = -1$ under the above homomorphism. So that handles $k = -1$.

That's actually a strange argument. Somehow this complex conjugation falls out of the sky and does what we want. To move towards a better understanding, let's extract the following principle:

Lemma 0.1. *Let F be any field of characteristic zero which contains a primitive n^{th} root of unity. If there is an automorphism of F which sends ζ_n to ζ_n^k, then k is in the image of our homomorphism $Gal(\mathbb{Q}(\zeta_n)/\mathbb{Q}) \to (\mathbb{Z}/n\mathbb{Z})^\times$.*

In other words, we don't have to actually think about producing an automorphism of our field $\mathbb{Q}(\zeta_n)$. We can just produce, by whatever means, an automorphism of some bigger field which happens to contains a primitive n^{th} root of unity.

In some sense I think this Lemma is "obvious" from a field theory perspective. But I'm not sure about that, so here's a proof. I recommend skipping it unless you're curious.

Proof. Certainly this is true if we redefine $\mathbb{Q}(\zeta_n)$ to be the subfield of F generated by ζ_n. For, any homomorphism $\sigma : F \to F$ must send ζ_n to another n^{th} root of unity, which is therefore a power of ζ_n; so σ sends $\mathbb{Q}(\zeta_n)$ inside itself. Applying the same argument to the inverse, we see that an automorphism of F restricts to an automorphism of $\mathbb{Q}(\zeta_n)$, which proves the claim.

But we already had a fixed $\mathbb{Q}(\zeta_n)$; how do we know it's isomorphic to this new one, which I'll now denote $\mathbb{Q}(\zeta_n)_F \subset F$? This sort of thing was implicit in the last lecture, but now I'll give an argument. First, we can enlarge F to its algebraic closure $\overline{F}$; then we can cut it down to the algebraic closure of $\mathbb{Q}$ in $\overline{F}$. This doesn't change the subfield $\mathbb{Q}(\zeta_n)_F$, because ζ_n is algebraic over $\mathbb{Q}$. Thus we can assume F is an algebraic closure of $\mathbb{Q}$. But then any isomorphism $F \simeq \overline{\mathbb{Q}}$ restricts to an isomorphism $\mathbb{Q}(\zeta_n)_F \simeq \mathbb{Q}(\zeta_n)$, by the same argument as in the first paragraph. □

So here's the picture. There are some special huge fields $F \supset \mathbb{Q}$ which are potentially analytic in nature. One of them is $F = \mathbb{C}$. They all shine their special light on the cyclotomic fields. If we just turn on one light we won't see much; but if we turn them all on that we get a full picture. The field $F = \mathbb{C}$ shone a light on the subgroup $\{\pm 1\}$ of $(\mathbb{Z}/n\mathbb{Z})^\times$. What we'll do to prove the theorem is find, for every prime p not dividing n, a field $F = \mathbb{Q}_p^{unr}$ which shines a light on the subgroup of $(\mathbb{Z}/n\mathbb{Z})^\times$ generated by p. This will be enough by our second lemma:

Lemma 0.2. *Suppose we know that for every prime p not dividing n, the homomorphism $Gal(\mathbb{Q}(\zeta_n)/\mathbb{Q}) \to (\mathbb{Z}/n\mathbb{Z})^\times$ hits p. Then $Gal(\mathbb{Q}(\zeta_n)/\mathbb{Q}) \to (\mathbb{Z}/n\mathbb{Z})^\times$ in fact hits every k, so that we're done.*

There's a good proof of this lemma, and a bad proof. The bad proof invokes Dirichlet's theorem on primes in arithmetic progressions to say that for any k relatively prime to n there exists a prime p congruent to k modulo n. This clearly implies the lemma, but Dirichlet's result involves some difficult (and amazing) mathematics which is strange to use in this context. The good proof is to say that every k can at least be written as a product of primes $p_1 \dots p_n$; then if $\sigma_1 \in Gal(\mathbb{Q}(\zeta_n)/\mathbb{Q})$ hits p_1, etc., it follows that $\sigma_1 \dots \sigma_n$ hits $k = p_1 \dots p_n$. In other words, the image of any group homomorphism is closed under products.

By combining the two lemmas we see that it suffices to prove that for every prime p not dividing n, there exists a field of characteristic zero which contains a ζ_n and admits an automorphism sending $\zeta_n \mapsto \zeta_n^p$. So that's what we'll do.

I'll let you meditate on that for a second.

OK. Here's a first observation: if we drop the requirement of "characteristic 0", then it's easy to produce such a field.

More precisely, it's easy as long as you know about the Frobenius:

Proposition 0.3. *Let F be a field of characteristic p. Define a map $Frob : F \to F$ by $Frob(x) = x^p$. Then $Frob$ is a field homomorphism.*

Proof. Certainly, $Frob(xy) = Frob(x)Frob(y)$. The surprising fact is that $Frob(x+y) = Frob(x) + Frob(y)$, i.e.

$$(x+y)^p = x^p + y^p$$

for $x, y \in F$, a field of characteristic p. This follows from the fact that the two-variable integer polynomial $(X+Y)^p - X^p - Y^p \in \mathbb{Z}[X,Y]$ has all its coefficients divisible by p. Indeed, if you look at these coefficients, they're binomial coefficients where p divides the numerator but not the denominator. □

Oftentimes the Frobenius $Frob : F \to F$ is actually an *automorphism*, i.e. it is surjective, and hence invertible (any field homomorphism is automatically injective). When this happens, we say F is *perfect*. For example, a finite field is perfect, since an injective map from a finite set to itself is always bijective. An algebraically closed field is also perfect, because finding a $Frob$-preimage of a amounts to solving the equation $X^p - a = 0$. The rational function field $\mathbb{F}_p(T)$ is not perfect: you can't hit T.

What we find, then, is that the field $\overline{\mathbb{F}_p}$ seems to satisfy all of our needs. Indeed, $\overline{\mathbb{F}_p}$ contains a primitive n^{th} root of unity: remember, the abstract proof of the existence of ζ_n we gave in the first lecture also worked for algebraically closed fields of characteristic not dividing n. It also admits an automorphism which sends ζ_n to ζ_n^p, because, in fact, it admits an automorphism (the Frobenius) which sends *everything* to its p^{th} power! There's just one catch, though, which is that $\overline{\mathbb{F}_p}$ has characteristic p, not 0. It doesn't have $\mathbb{Q}(\zeta_n)$ as a subfield.

So the game will be this. We will somehow try to "perturb" the algebraic structure of $\overline{\mathbb{F}_p}$ so that:

- We move from characteristic p to characteristic 0;
- The primitive n^{th} root of unity moves along with us;
- So does the Frobenius automorphism.

It turns out, by some sort of algebraic miracle, that such perturbations are possible in great generality.

Actually, it's not such a miracle. In fact you were taught something very similar in grade school. Remember, $\mathbb{Q}$ is a field of characteristic zero, but you were taught how to calculate $+$ and $\times$ inside it in terms of similar operations in $\mathbb{Z}/10\mathbb{Z}$, by using decimal expansions. And of course we might as well substitute our prime p for 10.

Now, unfortunately for us, there is something slightly non-algebraic about those algorithms you were taught in school. In fact, it turns out that there's no way to algebraically describe $\mathbb{Q}$ just in terms of $\mathbb{Z}/p\mathbb{Z}$. In particular you can't transfer the Frobenius from $\mathbb{Z}/p\mathbb{Z}$ to $\mathbb{Q}$. The problem is that pesky "carry" rule. It says that, when summing, the n^{th} digits contribute to the $(n+1)^{st}$ digit whenever

their sum in $\mathbb{Z}$ exceeds p. It's that "in $\mathbb{Z}$" part that's the problem: calculating in $\mathbb{Z}$, for our purposes, amounts to assuming what we're trying to prove.

But a mathematician named Witt saw a way around this. He figured out how to modify the rules of arithmetic for base-p-expansions in such a way that, now, everything is described purely in terms of the structure of $\mathbb{F}_p$ as a field. But by doing this you no longer end up with the rational numbers $\mathbb{Q}$ — instead you end up with a field $\mathbb{Q}_p$ of "p-adic numbers" (with a ring $\mathbb{Z}_p$ of "p-adic integers" as an intermediary). You can look up Witt's construction, if you like: it goes by the name of "Witt vectors". It involves some interesting integral polynomials which are generalizations of the $F(X,Y) = \frac{1}{p}[(X+Y)^p - X^p - Y^p]$ which came up in the proof of the existence of Frobenius.

The great thing for us is that Witt's construction works exactly the same way with any perfect field of characteristic p replacing $\mathbb{F}_p$. So we can apply it to our field $\overline{\mathbb{F}_p}$. The output of the Witt vector construction, in that case, is a ring $\mathbb{Z}_p^{unr}$ satisfying the following properties:

- There is a natural identification $\mathbb{Z}_p^{unr}/p = \overline{\mathbb{F}_p}$, and $\mathbb{Z} \to \mathbb{Z}_p^{unr}$ is injective. Furthermore, $\mathbb{Z}_p^{unr}$ is an *integral domain* (more on that later).
- Every automorphism of $\overline{\mathbb{F}_p}$ lifts to an automorphism of $\mathbb{Z}_p^{unr}$.
- If $f(X)$ is a monic polynomial in $\mathbb{Z}_p^{unr}[X]$, and if $f(X) \pmod p$ is separable in $\overline{\mathbb{F}_p}[X]$, then the map $\{\text{roots of } f(X) \text{ in } \mathbb{Z}_p^{unr}\} \to \{\text{roots of } f(X) \text{ in } \overline{\mathbb{F}_p}\}$ given by reducing $\pmod p$ is a bijection.

The first point expresses that $\mathbb{Z}_p^{unr}$ is a perturbation of $\overline{\mathbb{F}_p}$ to characteristic zero. If you set $p = 0$ in $\mathbb{Z}_p^{unr}$, then you get back $\overline{\mathbb{F}_p}$; but p itself is not zero in $\mathbb{Z}_p^{unr}$. This is akin to the relationship between $\mathbb{Z}$ and $\mathbb{F}_p$, but its important to bear in mind that the connection between $\mathbb{Z}_p^{unr}$ and $\overline{\mathbb{F}_p}$ is tighter, though we haven't listed the precise property that expresses that ($\mathbb{Z}_p^{unr}$ is a *strict p-ring*). The second and third points tell us that many structures carried by $\overline{\mathbb{F}_p}$ "lift" to $\mathbb{Z}_p^{unr}$. The meaning of "lift" is as follows: in the first case, we mean that for every automorphism $\sigma : \overline{\mathbb{F}_p} \to \overline{\mathbb{F}_p}$, there exists an automorphism $\widetilde{\sigma} : \mathbb{Z}_p^{unr} \to \mathbb{Z}_p^{unr}$ such that $\widetilde{\sigma}(x) \equiv \sigma(x) \pmod p$. In the second case, we mean that for every root $a \in \overline{\mathbb{F}_p}$ of $f(X) \pmod p$, there is a unique root $\widetilde{a}$ of $f(X)$ in $\mathbb{Z}_p^{unr}$ with $\widetilde{a} \equiv a \pmod p$.

Witt's construction describes the elements of $\mathbb{Z}_p^{unr}$ as formal infinite sums $[c_0] + [c_1]\cdot p + [c_2]\cdot p^2 + \ldots$, with the "digits" c_n all lying in $\overline{\mathbb{F}_p}$. Then one has to say what $+$ and $\times$ are. They are given by complicated formulas which, however, are described purely in terms of the $+$ and $\times$ in $\overline{\mathbb{F}_p}$.

We remark that many other constructions the ring $\mathbb{Z}_p^{unr}$ are possible. A particularly nice (and elementary!) one was given by Cuntz and Deninger just earlier this month! See

http://arxiv.org/abs/1311.2774

Now, let's return to this *integral domain* business. That's a fancy name, but it just means that whenever $ab = 0$, it must be that either $a = 0$ or $b = 0$. This is exactly the property one needs to be able to embed $\mathbb{Z}_p^{unr}$ into a field. A minimal such field can be built by formally considering fractions a/b with $b \neq 0$, and using the obvious rules of $+$ and $\times$, just like how one calculates in $\mathbb{Q}$ in terms of calculations in $\mathbb{Z}$. The resulting field, called the *fraction field* of $\mathbb{Z}_p^{unr}$, is denoted $\mathbb{Q}_p^{unr}$.

Now we are done: this $\mathbb{Q}_p^{unr}$ is a field of characteristic 0 which contains a primitive n^{th} root of unity ζ_n and admits an automorphism which sends ζ_n to ζ_n^p. It's characteristic 0 because $\mathbb{Z}_p^{unr}$ was an integral

domain in which no prime number is zero. It contains a ζ_n and has such an automorphism by the "lifting" properties described above, since $\overline{\mathbb{F}_p}$ contains such structure. (Take $f(x) = X^n - 1$, and $\sigma = Frob$.)

To summarize: we're trying to construct Galois automorphisms. The way we do it is by starting with the Frobenius in characteristic p, then showing that this Frobenius somehow lifts to characteristic 0. It no longer raises everything to the p^{th} power, but it still deserves to be called the Frobenius! If we take all primes together, then we actually generate the whole Galois group this way. So in a sense the Frobenius controls everything.

As a general remark, note that to go from characteristic p to characteristic 0, we necessarily had to pass through a ring $\mathbb{Z}_p^{unr}$ which was not a field. We ran into this in the last proof as well, where the intermediary ring was the less exotic ring $\mathbb{Z}$, and we were going from $\mathbb{F}_p$ to $\mathbb{Q}$.

Maybe it's worth augmenting this proof with some philosophical reflection. (Probably not, but I'm going to do it anyway.)

This field $\mathbb{Q}_p^{unr}$ may seem very complicated and unfamiliar. And it is. But in a sense, from the perspective of mathematics, it's not much more complicated than, say, the real numbers $\mathbb{R}$. Imagine if we, as human beings, didn't have such a direct experience with the concept of continuum in our day-to-day lives. Imagine everything we knew was discrete. Then we would really be number theorists. Our favorite ring would have to be $\mathbb{Z}$, and a ring like $\mathbb{R}$, or even worse $\mathbb{C}$, wouldn't occur to us for a long time. And when it did, we'd think it was kind of strange and miraculous, but a bit hard to construct.

This is exactly the situation for fields like $\mathbb{Q}_p^{unr}$. Perhaps if we were different sorts of beings with different immediate sensory experience, they would be completely intuitive for us. But as it stands, we just have to remember that, philosophically, they're on equal footing with $\mathbb{C}$, so we should do our best to get comfortable with them, continuum-afflicted beings that we are.

At the end of class we had a vote on what to do next. I said the options were either for us to take a step back and develop some general theory to help explain the concepts we've come across so far (algebraic number theory), or for us to visit the historical roots of this subject in Gauss's study of regular n-gons. I was gunning for the second option, but the first option won at the polls.

I think that was wisely voted. We have run across a lot of concepts here so far already, and it should be good to consolidate them in a more general study. We can get back to n-gons letter when we're hankering for something more concrete.

So, to finish this lecture I'll give a brief introduction to some ideas in algebraic number theory.

We can take as motivation the factorization

$$p = (1 - \zeta_p)(1 - \zeta_p^2) \dots (1 - \zeta_p^{p-1})$$

we ran across in the first proof of the irreducibility of cyclotomic polynomials. Remember, this is an equality in this ring $\mathbb{Z}[\zeta_p]$ we found sitting inside our field $\mathbb{Q}(\zeta_p)$.

Algebraic number theory studies arbitrary *number fields*, which are just fields F which are finite extensions of $\mathbb{Q}$. The first important observation in algebraic number theory is that for any such F there is a canonical "ring of integers" $\mathcal{O}_F \subset F$, generalizing $\mathbb{Z}[\zeta_p] \subset \mathbb{Q}(\zeta_p)$ or, at a more basic level, $\mathbb{Z} \subset \mathbb{Q}$. In general it's not as simple as saying, "take $\mathbb{Z}$-linear combinations of powers of your generator

instead of $\mathbb{Q}$-linear combinations". That ring depends on the generator, and it's generally not even guaranteed to form a finitely generated $\mathbb{Z}$-module! We just sort of got lucky with $\mathbb{Q}(\zeta_p)$. But there are several different ways of describing this $\mathcal{O}_F$ in general.

Once we have this $\mathcal{O}_F$, there are two basic types of questions one can ask:

- This ring is supposed to be like $\mathbb{Z}$. How far does the analogy go? Which results about $\mathbb{Z}$ transfer to $\mathcal{O}_F$?
- We can ask about more direct relationships between $\mathbb{Z}$ and $\mathcal{O}_F$. There's a homomorphism $\mathbb{Z} \to \mathcal{O}_F$. For example, how do primes in $\mathbb{Z}$ break up when you put them inside $\mathcal{O}_F$?

One of the most beautiful stories in algebraic number theory relates to the first point. A fundamental result about $\mathbb{Z}$ is, well, the so-called "fundamental theorem of arithmetic", which says that any integer has a factorization into primes, and that such a factorization is unique up to units and reordering.

You may not appreciate this fact because it's so hard-wired into your intuitions about $\mathbb{Z}$. For example, I'll bet you never stop to think, when you're reducing a huge fraction to lowest terms, that the end result might depend on the choices you make along the way about which factors to cancel first.

Actually, the first person who realized that this fundamental theorem of arithmetic needed to be proved, and the first person to give a proof, was Gauss in his Disquisitiones Arithmeticae. That book also contains (some of) Gauss's study of cyclotomy. It's not totally unlikely that Gauss came to realize how fundamental this theorem of arithmetic was, because he saw that it failed in these more general contexts!

That is, $\mathcal{O}_F$ does generally not satisfy unique prime factorization.

However, it turns out that the failure is in some sense bounded. A remarkable fact is that one can attach to F an abelian group $Cl(F)$, called the class group of F, which sort of measures how far away $\mathcal{O}_F$ is from having unique prime factorization. Then the theorem is that $Cl(F)$ is finite. (Also, you prove this by drawing some pictures — the real numbers have their revenge on the p-adics...)

Now, the Big Theorem that we're (sort of) aiming for says something about these class groups in the case of $F = \mathbb{Q}(\zeta_p)$. Namely, a natural question turns out to be whether p divides $\#Cl(\mathbb{Q}(\zeta_p))$ or not. An answer is given by "Kummer's criterion":

Theorem 0.4. *Define rational numbers B_j for $j \geq 0$ as $j!$ times the coefficient of X^j in the power series expansion of*

$$\frac{X}{e^X - 1} = 1 - \frac{1}{2}X + \frac{1}{12}X^2 + \dots.$$

Also, let p be a prime.

Then p divides the order of $Cl(\zeta_p)$ if and only if p divides the numerator of some B_j for $0 \leq j \leq p-3$.

In more modern times tighter connections have been found between these $Cl(\mathbb{Q}(\zeta_p))$ and the Bernoulli numbers B_j. But we'll focus on the above result.