How to deal with “infinitely many primes” proofs

Dan Fretwell

One particular stumbling block of the undergrad number theory student is in being asked to create proofs of theorems along the following lines:

“Prove that there are infinitely many primes of the form a mod n”.

These proofs have the appearance of being very lengthy and difficult to follow. The purpose of this document is to give a bit of an explanation as to what is going on, hopefully giving some kind of understanding of the general picture and allowing you to create your own such proofs.

1 Euclid’s proof

Let us start with Euclid’s proof that there are infinitely many primes. This is essentially the beginning of all the theorems that we will be interested in. Not only this but the clever tricks that Euclid used will become handy in future.

Theorem 1. *There are infinitely many primes.*

Proof. Suppose there are only finitely many primes. Call these $p_1, p_2, ..., p_n$.

Form the number:

$$N = (p_1p_2...p_n) + 1.$$

This number is certainly bigger than 1 so must have at least one prime factor. Let p be one of these.

Now p is prime. But by assumption there are only finitely many primes. So $p = p_i$ for some i.

However the fact that $p|N$ and $p|(p_1p_2...p_n)$ tells us that $p|1$. This is a contradiction! □

Let’s dissect this proof. There are three main steps:

1. First we made the assumption that there are only finitely many primes. We did this to try and get a contradiction.

2. We invented a specific number N that turned out to be very handy in getting our contradiction.

3. We showed that there exists a prime dividing N (which was very easy), then were able to use this and the original assumption to get the required contradiction.

In some sense this can be viewed as the proof that there are infinitely many primes congruent to 1 mod 2.

2 Working mod n

It took quite a while for mathematicians to start looking in more detail at the primes. There is so much more structure to them than meets the eye. In fact quite a lot of number theory is devoted to studying primes, we are far from finished.

One way to study the primes in more detail is to look at them "mod n". For example if we look at the sequence of primes mod 4 we get the following list:

$$2, 3, 1, 3, 3, 1, 1, 3, 3, 1, 3, ...$$

We can ask a few questions here:

- Why is there no 0 in this list?
- Will there ever be another 2 appear?
- Are there infinitely many 1's? Is this the same for 3's?

It is quite clear that any 0 in this list would correspond to a multiple of 4, so couldn't be prime. Similarly, with the exception of 2, every number of the form $4k + 2$ has 2 as a proper factor, so couldn't be prime.

We have answered the first two questions, now how about the last two? We will see that the answer to both is yes and we will prove them separately. In order to do this we will use tricks similar to Euclid's.

However we need to think a little about the steps we used last time:

1. We will still make the assumption that there will be only finitely many primes **of the given form**, trying to get a contradiction.
2. How will we invent N? This will have to be different in each case and we will have to rely on some piece of number theoretical knowledge to make the right choice.
3. We must then show that there is a prime **of the correct form** dividing N. Only then can we use our original assumption and get the contradiction (assuming a good N was chosen).

With this in mind, let's tackle the 3 mod 4 case.

Theorem 2. *There are infinitely many primes congruent to* 3 mod 4*.*

Proof. Suppose there are only finitely many primes congruent to 3 mod 4. Call these $p_1, p_2, ..., p_n$.

Form the number:

$$N = 4p_1p_2...p_n - 1.$$

Ok so why did we choose this to be N? Well in a minute we are going to see that this number MUST have a prime divisor that is 3 mod 4. This is what we want!

Now $N > 1$ is odd so N must have at least one odd prime divisor. Let p be one of these.

Remember, we now want to show that **there is** such a p that is 3 mod 4. We aren't done yet! We didn't have this problem with Euclid's theorem since we were just interested in "p being prime" not "p being prime and being 3 mod 4".

Suppose that **all** odd prime divisors of N are congruent to 1 mod 4 (i.e. that they are all not congruent to 3 mod 4). Then N would be factorised as a product of numbers that are all 1 mod 4, hence $N \equiv 1 \bmod 4$. However looking at our choice of N, it is clear that $N \equiv 3 \bmod 4$. So we get a contradiction, meaning that there must exist a prime p dividing N such that $p \equiv 3 \bmod 4$.

Now we have what we want! We have the right kind of p and so we just need to finish off the proof now.

But by assumption there are only finitely many primes congruent to 3 mod 4, hence $p = p_i$ for some i. So by the facts that $p|N$ and $p|(p_1p_2...p_n)$, it must be that $p|-1$. This is a contradiction. □

Chances are you will have to read this proof a few times before it becomes clear what is going on. Hopefully the narrative helps to give a view of things. You should compare this proof with Euclid's, observing similarities and what extra work needed to be done.

We now try to prove the same thing for primes that are 1 mod 4. We will see that things are not so easy.

Theorem 3. *There are infinitely many primes congruent to* 1 mod 4.

Proof. (Attempt!) Suppose there are only finitely many primes congruent to 1 mod 4. Call these $p_1, p_2, ..., p_n$.

You might be led to construct N as follows:

$$N = 4p_1p_2...p_n + 1.$$

This N is a bad choice but it is not obvious at the moment why. We shall continue until we struggle to proceed.

Now $N > 1$ is odd so N must have at least one odd prime divisor. Let p be one of these.

Like last time we are trying to show that **there is** such a p that is 1 mod 4.

Suppose that **all** odd prime divisors of N are congruent to 3 mod 4. Then N would be factorised as a product of numbers that are all 3 mod 4. However, does this necessarily mean that $N \equiv 3 \bmod 4$? It doesn't since products of numbers that are 3 mod 4 could be either 1 mod 4 or 3 mod 4 (e.g. $7*11 = 77 \equiv 1 \bmod 4$, $3*7*11 = 231 \equiv 3 \bmod 4$).

So our proof has failed! We can no longer guarantee that there is a p dividing N that is $1 \bmod 4$ and so we cannot get the contradiction we wanted.

□

In light of the problem mentioned above it appears we need a better way of "separating" primes congruent to 1 mod 4 from primes that are congruent to 3 mod 4.

How might we do this? Well we turn to the theory of quadratic residues. We know that for odd primes p, the Legendre symbol $\left(\frac{-1}{p}\right)$ is entirely dependent on p mod 4 so maybe we can use this somewhere? But in order to involve the Legendre symbol we need squares!

Proof. (Corrected) This time we construct N as follows:

$$N = 4(p_1 p_2 ... p_n)^2 + 1.$$

Why is this going to help? Well as usual N must have at least one odd prime divisor p.

But then:

$$4(p_1 p_2 ... p_n)^2 \equiv -1 \text{ mod } p.$$

In other words -1 is a square mod p, meaning that $\left(\frac{-1}{p}\right) = 1$. However this only happens when $p \equiv 1$ mod 4. We are now able to guarantee that there is a prime p dividing N that is congruent to 1 mod 4. (In fact we have proved that **all** of them must be of this form but we don't really need this).

From here we finish the proof in the usual way. Since $p \equiv 1$ mod 4 we must have that $p = p_i$ for some i (by the original assumption). But then $p|N$ and $p|4(p_1 p_2 ... p_n)^2$ tells us that $p|1$. This is a contradiction. □

Sometimes you have to use a mixture of techniques to prove these theorems. For example working mod 8 we know that (apart from 2) all primes must be either $1, 3, 5, 7$ mod 8. If we continue to prove our theorems in the same way as before we now have to eliminate three of these posibilities when studying primes that divide our N (earlier we had either 1 or 3 mod 4, and only had to eliminate one of these possibilities to get the other).

Maybe we can mix the approaches from Theorems 2 and 3. Let's try to use Legendre symbols first to eliminate some possibilities and then use proof by contradiction.

Theorem 4. *There are infinitely many primes congruent to* 7 mod 8.

Proof. Start as usual by supposing that there are only finitely many such primes. Call them $p_1, ..., p_n$.

Now we need a way to "separate" those 7 mod 8 primes from the others. Recall that:

$$\left(\frac{2}{p}\right) = \begin{cases} 1 & \text{if } p \equiv 1, 7 \text{ mod } 8 \\ -1 & \text{if } p \equiv 3, 5 \text{ mod } 8 \end{cases}$$

We have **almost** been able to isolate the 7 mod 8 primes.

Construct N as follows:

$$N = (p_1 p_2 ... p_n)^2 - 2.$$

As usual, $N > 1$ is odd and so has an odd prime divisor p.

Thus:

$$(p_1 p_2 ... p_n)^2 \equiv 2 \bmod p,$$

and so $\left(\frac{2}{p}\right) = 1$. But we have just seen that this implies that $p \equiv 1, 7 \bmod 8$.

We have done a good job of eliminating the posibilities $3, 5 \bmod 8$ as prime divisors but we are not yet done. Remember we want to guarantee that there is an 7 mod 8 prime divisor to get the contradiction!

Now suppose that **all** prime divisors of N are congruent to 1 mod 8. Then N would be factorised as a product of numbers that are 1 mod 8. So $N \equiv 1 \bmod 8$. But by looking at our definition of N we see that $N \equiv -1 \bmod 8$ (for each i we know that $p_i \equiv 7 \bmod 8$ so that $p_i^2 \equiv 1 \bmod 8$). This is a contradiction, so there must be at least one prime p dividing N that is 7 mod 8.

Finally we have guaranteed that the right kind of p divides N. We can finish the proof now in the usual way.

But there are only finitely many primes congruent to 7 mod 8 by assumption. So $p = p_i$ for some i. However $p|N$ and $p|(p_1 p_2 ... p_n)^2$ so $p|2$. This is a contradiction. □

Hopefully you will now be able to construct your own proofs of similar results. However you should notice that as the modulus increases it is getting harder to prove such statements. The tricks required are more complicated. In fact there comes a point where even the Legendre symbol and the other tricks do not help enough!

Fortunately Dirichlet was able to prove the general result in 1837:

Theorem 5. *Let n be a natural number. Then for any a coprime to n there are infinitely many primes congruent to a* mod *n.*

His proof was not along the lines of what we have done here but uses analytical techniques (specifically convergence of certain infinite series called Dirichlet L-series). The use of the Legendre symbol in our proofs above is connected with the use of one of these Dirichlet L-series in the general theory. You will get to study his proof of the general theorem if you go on to study the level 4 module “Analytic Number Theory”.

3 Summary

So to summarise, here are the key steps:

1. Assume that there are only finitely many primes **of the given form**.
2. Create N using number theoretical knowledge with the idea of trying to isolate the primes that you are interested in. Usually $N = a(p_1 p_2 ... p_n) + b$ or $a(p_1 p_2 ... p_n)^2 + b$ for some nicely chosen numbers a and b.
3. Prove that N has a prime divisor **of the correct form** then use the assumption to get a contradiction.