MA4N1 : Theorem Proving with Lean

Special Cases of Dirichlet's Theorem on Arithmetic Progression

Additional Notes

Group Name : Dirichlet

Module Leader
Dr. Damiano Tester

Institute of Mathematics

February 4, 2024

Foreword

This document acts as a supplement to our GitHub repository containing the project for MA4N1: Theorem Proving with Lean. Here we aim to add some explanation to the Lean4 code we have written and justify the assumptions we have made, with reference to the proofs we have formalised.

The repository can be found following *this* link or alternatively by visiting `https://github.com/J-Atfield/ma4n1-dirichlet-project`. Contribution to the project was even from all members of the group and thus the commit history in the repository is not reflective of actual contributions made.

Contents

1 Statement of Theorem

Our project surrounds the special cases of Dirichlet's theorem on arithmetic progressions, which we will call Dirichlet's theorem for short. The theorem is stated is as follows.

Theorem 1.1 (Dirichlet)**.** *Let $a, m \in \mathbb{Z}$, with a and m coprime. Then there are infinitely many prime numbers of the form $a + km$, where $k \in \mathbb{N}$.*

The theorem was proved by Dirichlet and for our project we will be tackling the formalisation in Lean4 of some of the specific cases, namely the cases $4k + 1$, $6k + 1$ and $8k + 1$.

2 Foundations

In this section we will state the fundamental properties required to prove our results as well as discussing how they are within Lean4. We will firstly introduce the two most fundamental definitions to our project.

Definition 2.1. An integer a is a quadratic residue modulo n if

$$x^2 \equiv a \mod n$$

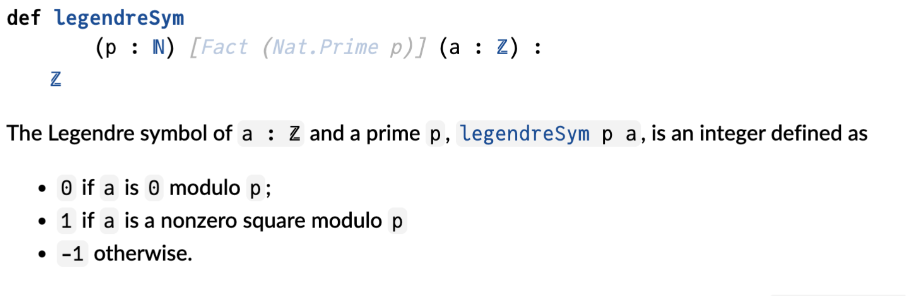
for some integer x.

In Lean4 we have that IsSquare (a : ZMod p) is equivalent to there being a solution to $x^2 = a \mod p$ i.e a is a quadratic residue modulo p, and we will use this throughout our proofs.

Definition 2.2. For p prime and $a \in \mathbb{Z}$ we define the Legendre symbol as

$$\left(\frac{a}{p}\right) = \begin{cases} 1 & \text{if } a \text{ is a quadratic residue modulo } p \text{ and } a \not\equiv \mod p \\ -1 & \text{if } a \text{ is not a quadratic residue midulo } p \\ 0 & \text{if } a \equiv 0 \mod p \end{cases}$$

The Legendre symbol is defined in Mathlib in the following way:



```
def legendreSym
      (p : ℕ) [Fact (Nat.Prime p)] (a : ℤ) :
  ℤ
```

The Legendre symbol of `a : ℤ` and a prime `p`, `legendreSym p a`, is an integer defined as

- `0` if `a` is `0` modulo `p`;
- `1` if `a` is a nonzero square modulo `p`
- `-1` otherwise.

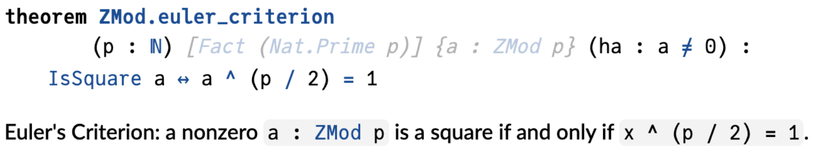
Figure 1: Legendre symbol in Mathlib4

Throughout our project we also regularly use Euler's Criterion.

Theorem 2.3 (Euler's Criterion)**.** *Let p be an odd prime, and let $a \in \mathbb{Z}$ where p does not divide a . Then,*

$$\left(\frac{a}{p}\right) = a^{\frac{p-1}{2}} \mod p$$

Which is stated in lean in the following way.



```
theorem ZMod.euler_criterion
        (p : ℕ) [Fact (Nat.Prime p)] {a : ZMod p} (ha : a ≠ 0) :
    IsSquare a ↔ a ^ (p / 2) = 1
```

Euler's Criterion: a nonzero `a : ZMod p` is a square if and only if `x ^ (p / 2) = 1`.

Figure 2: Euler's criterion in Mathlib4.

3 The Fundamental Lemma

To prove the special cases in a consistent way we are going to use the following Fundamental Lemma. The lemma allows us to take any non constant polynomial over the integer ring and establish an infinite set of dividing primes. Mathematically the lemma is stated and proved in the following way.

Lemma 3.1. *Let $f(x) \in \mathbb{Z}[x]$ be a non constant polynomial. The set of primes p such that $p|f(n)$ for some $n \in \mathbb{Z}$ is infinite.*

Proof. For this proof we use Euclid's classic proof-by-contradiction idea which he used to establish the cardinality of the primes. If $f(x)$ has constant term zero the lemma becomes trivial. This follows from that $p|f(p)$ in such a case. Therefore let $f(x) = a_n x^n + \cdots + a_0$ then we have $a_0 \neq 0$. Hence it follows that $f(x)$ has structure,

$$f(x) = xg(x) + a_0$$

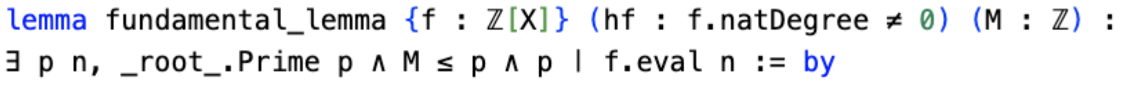
Where $g(x) \in \mathbb{Z}[x]$. For some $n \in \mathbb{Z}$ we will have $f(n) = ng(n) + a$, and as n gets very large it follows that $g(n)$ will become very large positive or very large negative. Without loss of generality we can assume very large positive. In particular, for any given $a \in \mathbb{Z}$, we can find an n large enough such that $f(n) > |a|$. Now let $k \in \mathbb{Z}$ be large, and consider,

$$\begin{aligned} f(k!\,a^2) &= a^2k!\,g(a^2k!\,) + a_0 \\ &= a(ak!\,g(a^2k!\,) + 1) \end{aligned}$$

Then we have that $ak!\,g(a^2k!\,) + 1$ is divisible by some prime (as $f(n) > |a|$) and that prime must be greater than k due to the $k!$ removing divisibility by all integers less than or equal to k. Hench proving we must have infinite primes satisfying our given relation.

□

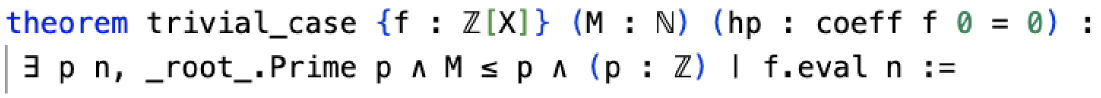
In Lean4 we stated this lemma in the following way, we found this was the most fitting statement for our special cases.



```
lemma fundamental_lemma {f : ℤ[X]} (hf : f.natDegree ≠ 0) (M : ℤ) :
∃ p n, _root_.Prime p ∧ M ≤ p ∧ p ∣ f.eval n := by
```

Figure 3: Statement of Fundamental Lemma in Lean4.

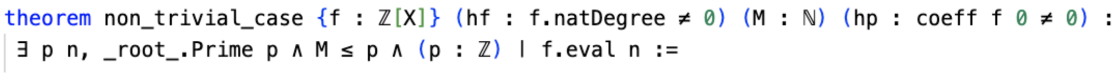
Following with the proof of Lemma 3.1 we began by proving the trivial case in Lean4.



```
theorem trivial_case {f : ℤ[X]} (M : ℕ) (hp : coeff f 0 = 0) :
  ∃ p n, _root_.Prime p ∧ M ≤ p ∧ (p : ℤ) ∣ f.eval n :=
```

Figure 4: Trivial case of Fundamental Lemma

After successfully formalising the trivial case, we then began to prove the general case, following the mathematical proof.



```
theorem non_trivial_case {f : ℤ[X]} (hf : f.natDegree ≠ 0) (M : ℕ) (hp : coeff f 0 ≠ 0) :
  ∃ p n, _root_.Prime p ∧ M ≤ p ∧ (p : ℤ) ∣ f.eval n :=
```

Figure 5: General case of Fundamental Lemma

which is where we have ran into a small amount of issues. Namely we had to sorry out three some what trivial looking statements in order to progress with the proof.

4 Special cases

In this section we prove the existence of infinitely many primes in progressions of certain forms. We start by proving that there exist infinitely many primes of the forms $4k+1$ and then progress to, $6k+1$ and $8k+1$. In this section we will also go over how we have adapted the statement and proofs of the theorems to achieve formalisation in Lean4.

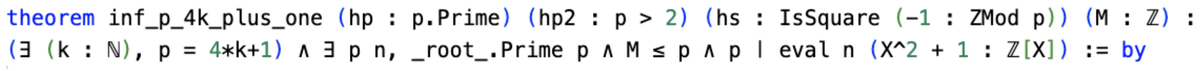
4.1 $4k+1$

The mathematical statement and proof of the special case where our primes take the form of $4k+1$ is as follows.

Theorem 4.1. *There are infinitely many primes of the form* $4k+1$.

Proof. Using Fermat's little theorem and Wilson's theorem we can quickly deduce that the quadratic congruence $x^2+1 \equiv 0 \mod p$ where p is an odd prime has a solution if and only if $p \equiv 1 \mod 4$. Hence it follows that for every p for which $x^2+1 \equiv 0 \mod p$ is solvable is of the form $4k+1$. By Lemma 3.1 there will be infinitely many such primes, and our statement is therefore established. □

In Lean4 we state this theorem in the following way:

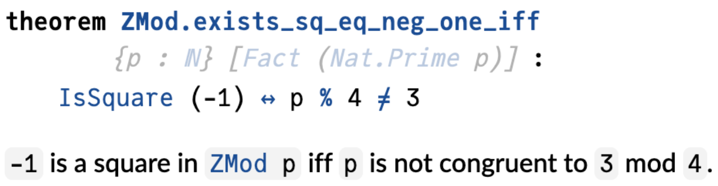


```
theorem inf_p_4k_plus_one (hp : p.Prime) (hp2 : p > 2) (hs : IsSquare (-1 : ZMod p)) (M : ℤ) :
(∃ (k : ℕ), p = 4*k+1) ∧ ∃ p n, _root_.Prime p ∧ M ≤ p ∧ p ∣ eval n (X^2 + 1 : ℤ[X]) := by
```

Figure 6: Statement of $4k+1$ special case in Lean4.

This statement shows there are infinitely many primes of the form $4k+1$ due to the fact that for any integer M in the statement we can find a prime of the form $4k+1$ greater than M.

In order to prove this Theorem in Lean4 we capitalised on some results within the Mathlib4 library, including the following very helpful theorem.

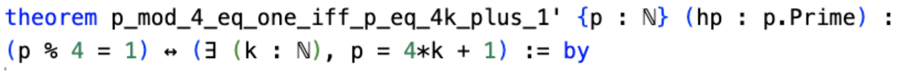


```
theorem ZMod.exists_sq_eq_neg_one_iff
      {p : ℕ} [Fact (Nat.Prime p)] :
    IsSquare (-1) ↔ p % 4 ≠ 3
```

-1 is a square in ZMod p iff p is not congruent to 3 mod 4.

Figure 7: $x^2 \equiv -1 \mod p$ if and only if $p \neq 4k + 3$.

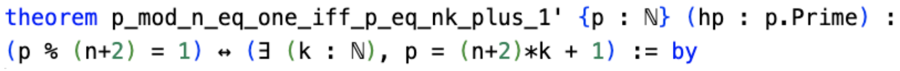
This result made the proof dramatically simpler as in order to show that $x^2 + 1 \equiv 0 \mod p$ where p is an odd prime has a solution if and only if $p \equiv 1 \mod 4$ we could just negate this statement. We consequently proved this if and only if relation and took advantage of it to obtain that $p \equiv 1 \mod 4$ from which we were able to complete the proof, with the help of the supplementary result below.



```
theorem p_mod_4_eq_one_iff_p_eq_4k_plus_1' {p : ℕ} (hp : p.Prime) :
(p % 4 = 1) ↔ (∃ (k : ℕ), p = 4*k + 1) := by
```

Figure 8: Equivalence relation for $4k + 1$.

Which we than later generalised so we could use it for all cases.



```
theorem p_mod_n_eq_one_iff_p_eq_nk_plus_1' {p : ℕ} (hp : p.Prime) :
(p % (n+2) = 1) ↔ (∃ (k : ℕ), p = (n+2)*k + 1) := by
```

Figure 9: General equivalence relation.

4.2 $6k + 1$

Now moving onto our next special case.

Theorem 4.2. *There are infinitely many primes of the form* $6k + 1$.

Proof. Consider the congruence $x^2 + 3 \equiv 0 \mod p$. This will have solutions for those p for which -3 is a quadratic residue modulo p. Consider the following manipulation of Legendre symbols

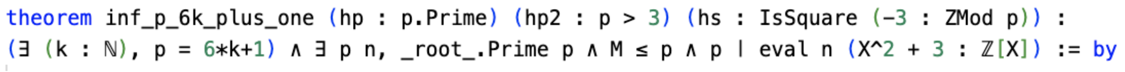
$$\left(\frac{-3}{p}\right) = \left(\frac{-1}{p}\right)\left(\frac{3}{p}\right) = \left(\frac{-1}{p}\right)\left(\frac{p}{3}\right)(-1)^{\frac{p-1}{2}} = \left(\frac{p}{3}\right)$$

Now using Euler's criterion we have that,

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

Clearly we have $p \equiv 1 \mod 2$ for all primes except 2. Further we get from the Chinese Remainder Theorem that every integer satisfying this system is of the form $6k + 1$, hence by the Fundamental lemma we are done. □

We state this in Lean4 in the same way as in the $4k+1$ case.



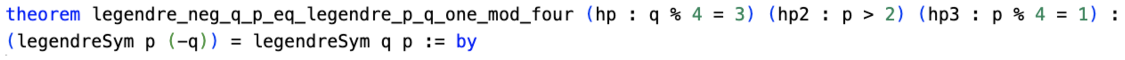
```
theorem inf_p_6k_plus_one (hp : p.Prime) (hp2 : p > 3) (hs : IsSquare (-3 : ZMod p)) :
(∃ (k : ℕ), p = 6*k+1) ∧ ∃ p n, _root_.Prime p ∧ M ≤ p ∧ p ∣ eval n (X^2 + 3 : ℤ[X]) := by
```

Figure 10: Statement of $6k+1$ special case in Lean4.

Since the mathematical proof works with the consideration of the relation $x^2+3 \equiv 0 \mod p$ we must make the assumption in our statement so that we can deduce the required implications. This meant the bulk of the formalisation of this theorem was hence centred on proving that manipulation of the Legendre symbol below held.

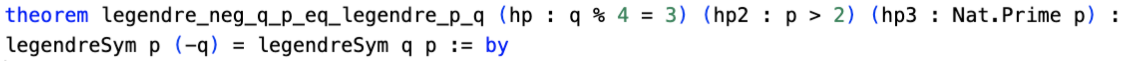
$$\left(\frac{-3}{p}\right) = \left(\frac{-1}{p}\right)\left(\frac{3}{p}\right) = \left(\frac{-1}{p}\right)\left(\frac{p}{3}\right)(-1)^{\frac{p-1}{2}} = \left(\frac{p}{3}\right)$$

This led to us proving some nice general results such as,



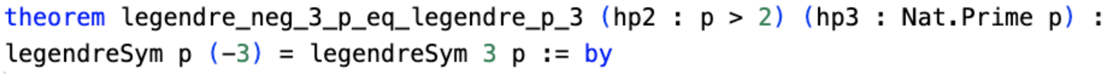
```
theorem legendre_neg_q_p_eq_legendre_p_q_one_mod_four (hp : q % 4 = 3) (hp2 : p > 2) (hp3 : p % 4 = 1) :
(legendreSym p (-q)) = legendreSym q p := by
```

and



```
theorem legendre_neg_q_p_eq_legendre_p_q (hp : q % 4 = 3) (hp2 : p > 2) (hp3 : Nat.Prime p) :
legendreSym p (-q) = legendreSym q p := by
```

which allowed us to obtain,



```
theorem legendre_neg_3_p_eq_legendre_p_3 (hp2 : p > 2) (hp3 : Nat.Prime p) :
legendreSym p (-3) = legendreSym 3 p := by
```

4.3 $8k+1$

The mathematical statement and proof of the final special case where our primes take the form of $8k+1$ is as follows.

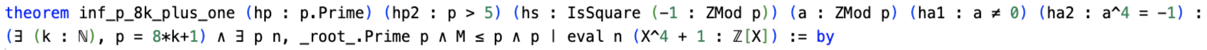
Theorem 4.3. *There are infinitely many primes of the form* $8k+1$.

Proof. We want to establish that the odd primes p for which $x^4+1 \equiv 0 \mod p$ admits a solution are of the form $8k+1$. By making the substitution $y = x^2$ we have that $y^2 = 0 \mod p$ and so from Theorem 4.1 we have that p must be of the form $4k+1$. Furthermore suppose a was a solution to $x^4 \equiv -1 \mod p$, we have from Fermat's little theorem,

$$1 \equiv a^{p-1} \equiv (a^4)^{\frac{p-1}{4}} \equiv (-1)^{\frac{p-1}{4}} \mod p.$$

Now the only way for this equation to hold is if $p = 8k+1$ and since $x^4+1 \equiv 0 \mod p$ for infinitely many primes, by Lemma 3.1 we are done. □

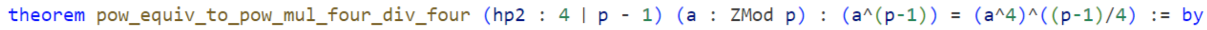
We state this in Lean4 in the following way:



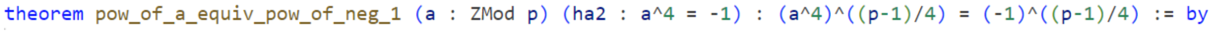
```
theorem inf_p_8k_plus_one (hp : p.Prime) (hp2 : p > 5) (hs : IsSquare (−1 : ZMod p)) (a : ZMod p) (ha1 : a ≠ 0) (ha2 : a^4 = −1) :
(∃ (k : ℕ), p = 8*k+1) ∧ ∃ p n, _root_.Prime p ∧ M ≤ p ∧ p ∣ eval n (X^4 + 1 : ℤ[X]) := by
```

Figure 11: Statement of $8k + 1$ special case in Lean4.

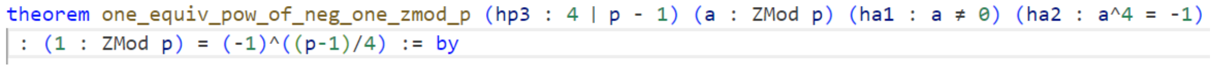
In comparison to the $4k + 1$ and $6k + 1$ cases, for this special case we require the further assumptions of $p > 5$ however this does not effect the infinite nature of p as well as $(ha1 : a \neq 0)$ and $(ha2 : a^4 = -1)$. The inclusion of $ha1$ and $ha2$ is to allow us to prove the congruence relation that is given as a result of Fermat's little theorem, with them being sensible as a is assumed to be a solution to $x^4 \equiv -1 \mod p$. Stating this relation in Lean4 required it's split into smaller individual equivalence relations, each following on from the last.



```
theorem pow_equiv_to_pow_mul_four_div_four (hp2 : 4 ∣ p - 1) (a : ZMod p) : (a^(p-1)) = (a^4)^((p-1)/4) := by
```

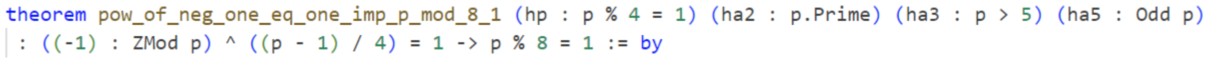


```
theorem pow_of_a_equiv_pow_of_neg_1 (a : ZMod p) (ha2 : a^4 = -1) : (a^4)^((p-1)/4) = (-1)^((p-1)/4) := by
```



```
theorem one_equiv_pow_of_neg_one_zmod_p (hp3 : 4 ∣ p - 1) (a : ZMod p) (ha1 : a ≠ 0) (ha2 : a^4 = -1)
  : (1 : ZMod p) = (-1)^((p-1)/4) := by
```

Besides from these the formalisation of this special case follows very closely to the structure used in the previous two cases, the only difference being that we needed to prove certain consequences of the aforementioned congruence relation, most importantly the result below which allows us to obtain the desired form for p.



```
theorem pow_of_neg_one_eq_one_imp_p_mod_8_1 (hp : p % 4 = 1) (ha2 : p.Prime) (ha3 : p > 5) (ha5 : Odd p)
  : ((-1) : ZMod p) ^ ((p - 1) / 4) = 1 -> p % 8 = 1 := by
```