GROUP THEORY: AN INTRODUCTION AND AN APPLICATION

NATHAN HATCH

Abstract. The ultimate goal of this paper is to prove that a prime p can be expressed as a sum of two squares if and only if $p = 2$ or $p \equiv 1 \pmod 4$. More time is spent on developing the tools to answer this question than on the question itself. To this end, the first section is entirely an introduction to basic group theory, and the second section extends this theory to prove that the group $\mathbb{Z}_p^\times$ is cyclic. The next section presents the Legendre symbol and proves one of the supplementary theorems of quadratic reciprocity. Finally, this last theorem provides the crucial first step in establishing the desired congruence conditions for primes of the form $x^2 + y^2$.

Contents

1. Introduction

A prime number p can be expressed in the form $x^2 + y^2$ for integral x and y if and only if $p = 2$ or $p \equiv 1 \pmod 4$. This observation was first made by Albert Girard in 1632, and the first person to claim to have a proof was Pierre de Fermat in 1640. For this reason, the statement is often called "Fermat's theorem on sums of two squares," or "Fermat's two-squares theorem" for short. The earliest written proof is (unsurprisingly) due to Leonhard Euler, who used the method of infinite descent to prove it in 1749, after a great deal of hard work [6].

Fermat's two-squares theorem provided the impetus for study of the general question of when a prime has the form $x^2 + ny^2$, where n is also an integer. This problem is by now a solved one; however, it took a lot of time—and many developments in number theory and algebra—to get there. The solution, which is discussed in David A. Cox's 1997 book *Primes of the Form* $x^2 + ny^2$, involves genus theory, reciprocity laws, and class field theory [5].

That said, this paper concerns itself merely with proving the original theorem about primes with the form x^2+y^2. Our proof will not follow that of Euler; instead,

Date: August 5, 2011.

our aim is to use more recent developments in number theory (which occurred partially *because* of studies of the aforementioned general problem) to offer a simpler explanation than the one Euler gave. The main topic is group theory, which is the subject of the first two sections. No background knowledge of groups is required; the first section begins with the basics and ends by proving Lagrange's theorem. In the second section we prove that $\mathbb{Z}_p^\times$ is cylic. (Note that some familiarity with fields and polynomials will make several of the lemmas in this section more understandable.) In section three, we introduce and discuss the Legendre symbol, using the cyclicity of $\mathbb{Z}_p^\times$ to prove the first supplementary law of quadratic reciprocity. With this knowledge, we finally have the necessary background for our proof of Fermat's two-squares theorem in section four.

2. Basic Group Theory

We begin, of course, with the definition of a group. Many of the proofs in this section are adapted from those presented in Laszlo Babai's 2011 University of Chicago VIGRE REU apprentice lecture series [1].

Definition 2.1. A **group** $(G, *)$ is a set G with a binary operation $* : G \times G \to G$ that satisfies the following three axioms:

(i) The operation $*$ is associative: for all $g, h, i \in G$, $(g * h) * i = g * (h * i)$.
(ii) There is an identity element $e \in G$ such that for any $g \in G$, $g * e = e * g = g$.
(iii) Every element $g \in G$ has an inverse $g^{-1} \in G$ satisfying $g * g^{-1} = g^{-1} * g = e$.

A group $(G, *)$ is **abelian** if the operation $*$ is commutative: for all $g, h \in G$, $g * h = h * g$.

The cardinality of the set G, denoted $|G|$, is called the **order** of $(G, *)$.

Example 2.2. The set of integers $\mathbb{Z}$ is an abelian group under the usual notion of addition. It has infinite order, and the identity element e is the number 0. However, $\mathbb{Z}$ is *not* a group under multiplication. The identity in this case would be 1, but most integers do not have an inverse in $\mathbb{Z}$. For example, $\frac{1}{2}$, the multiplicative inverse of 2, is not an integer.

Note that, when the group operation is clear from the context, often a group will be represented simply by the set G, rather than $(G, *)$. Also, it is a very common shorthand notation to omit the group operation when writing $g * h$, leaving simply gh. Next, we introduce the notion of subgroups, and we explain the concept of subgroups generated by subsets of a group.

Definition 2.3. A **subgroup** $(H, *)$ of a group $(G, *)$, denoted $H \leq G$, is a nonempty subset $H \subseteq G$, with the same group operation $*$, that satisfies closure under multiplication and inverses:

(i) $\forall g, h \in H$, $g * h \in H$, and
(ii) $\forall g \in H$, $g^{-1} \in H$.

This definition immediately implies that all subgroups must contain the identity e. To see this, we can simply take any $h \in H$ (which we are guaranteed to have, since H is nonempty), and multiply it by its inverse, which by definition must also be in H: $hh^{-1} = e$.

Definition 2.4. Let G be a group and $T \subseteq G$. The **subgroup generated by T**, denoted $\langle T \rangle$, is the unique subgroup $K \leq G$ satisfying

(i) $T \subseteq K$, and
(ii) for all subgroups $H \leq G$, $T \subseteq H$ implies $K \subseteq H$.

If T consists of a single element g, the notation $\langle\{g\}\rangle$ is simplified to $\langle g\rangle$.

Definition 2.4 requires some justification. We need to show that such a subgroup always exists, and that it is unique. To prove this, we can use the fact that an arbitrary intersection of subgroups is itself a subgroup.

Lemma 2.5. *Let G be a group and $\{H_i\}_{i\in I}$ be an arbitrary set of subgroups. Let $J = \bigcap_{i\in I} H$. Then J is also a subgroup.*

Proof. First note that J is nonempty, because all subgroups contain the identity e. Now, pick any $g, h \in J$ (not necessarily distinct); we wish to show that $gh \in J$ and $g^{-1} \in J$. For any $i \in I$, we have $g \in H_i$ and $h \in H_i$. Then since H_i is closed under multiplication and inverses, both gh and g^{-1} must be elements of H_i as well. But since this is true for all $i \in I$, they must also be elements of the intersection J. □

With this fact in mind, we can now show that Definition 2.4 is valid.

Proposition 2.6. *If G is a group and $T \subseteq G$, then $\langle T\rangle$ exists and is unique.*

Proof. Define the set $\Gamma = \{H \leq G : T \subseteq H\}$. Note that Γ is not empty, because G itself is necessarily a subgroup containing T. Let

$$K = \bigcap_{H\in\Gamma} H.$$

We claim $K = \langle T\rangle$. Once proven, this automatically shows that $\langle T\rangle$ is unique, since there can be only one set satisfying the definition of K.

Now, since every subgroup $H \in \Gamma$ contains T, T must be a subset of their intersection. Therefore, by Lemma 2.5, K is itself a subgroup of G such that $T \subseteq K$. It is also clear that K must be the smallest such subgroup, since it is by definition a subset of any other subgroup containing T. Hence, $K = \langle T\rangle$. □

An example should help to clarify what is meant by the generation of a subgroup.

Example 2.7. As before, consider the group $(\mathbb{Z}, +)$. It is easy to see that the even integers $\mathbb{E}$ are a subgroup of $\mathbb{Z}$: the sum of two integers is even, and if an integer n is even, so is its inverse $-n$. This subgroup can be generated by a number of different sets; for example $\langle 2\rangle$ and $\langle\{2, 4, 28\}\rangle$ both equal $\mathbb{E}$. However, the set $\{4, 28\}$ generates the multiples of four. In general, we observe that $\langle T\rangle$ is the set of all compositions and inverses of elements in T.

We have already defined the notion of the order of a group. There is also a notion of order for individual elements of a group.

Definition 2.8. Let G be a group and $g \in G$. The **order** of g, denoted $ord(g)$, is the least natural number n satisfying $g^n := \underbrace{g * g * \cdots * g}_{n \text{ times}} = e$. If no such n exists, we say g has (countably) **infinite order**.

It is no coincidence that the same word is used in two different contexts here. As the following proposition shows, the two concepts are closely related.

Proposition 2.9. *If G is a group and $g \in G$, then $ord(g) = |\langle g\rangle|$.*

Proof. First of all, let $S = \{g^k : k \in \mathbb{Z}\}$. We claim $S = \langle g \rangle$. It is easy to see that S is closed under multiplication and inverses; therefore, S is a subgroup of G which contains g. By definition, this means $\langle g \rangle \subseteq S$. To see that $S \subseteq \langle g \rangle$, consider any $g^k \in S$. If $k = 0$, it is clear that $g^k \in \langle g \rangle$, since $g^0 = e$ and all subgroups contain the identity. If $k > 0$, again we must have $g^k \in \langle g \rangle$, since $\langle g \rangle$ is closed under multiplication and $g^k = \underbrace{g * g * \ldots * g}_{k \text{ times}}$. Finally, $k < 0$ means $g^k \in \langle g \rangle$, since by the previous case $g^{-k} \in \langle g \rangle$ and $\langle g \rangle$ is closed under inverses. This covers all cases, so the desired equality has been established.

It remains to show that $|S| = ord(g)$. We first show this for the case when g has infinite order. Assume that for all $n \in \mathbb{N}$, $g^n \neq e$. If this is true, then we claim that for all $i, j \in \mathbb{Z}$, $i \neq j$ implies $g^i \neq g^j$. Say, without loss of generality, that $i \geq j$. If $g^i = g^j$, then $g^{i-j} = g^i g^{-j} = g^j (g^j)^{-1} = e$, which is impossible if $i - j \in \mathbb{N}$. Therefore $i - j = 0$, and $i = j$. This is enough to show that the set S is countably infinite.

Now for the finite case. Suppose there is an $n \in \mathbb{N}$ such that $ord(g) = n$. Our proof will be complete if we can show that the set S is the same as another set $S' = \{e, g, g^2, \ldots, g^{n-1}\}$, where all elements of S' are distinct. Suppose there exist $k, l \in \{0, \ldots, n-1\}$ such that $g^k = g^l$. Then, as above, let $k \geq l$, and note that $0 \leq k - l < n$. As before, we see that $g^{k-l} = e$. But n is the least positive integer satisfying $g^n = e$, and therefore $k - l = 0$. This proves the elements of S' are distinct. Now, pick any $g^a \in S, a \in \mathbb{Z}$. By the Division Theorem, $a = qn + r$ for some $q \in \mathbb{Z}$, $r \in \{0, \ldots, n-1\}$. Then $g^a = g^{qn+r} = (g^n)^q g^r = g^r$, which is in S', as desired. Since clearly also $S' \subseteq S$, we conclude $S' = S$. □

We will return to the concept of subgroup generation in the discussion of cyclic groups in the next section. For now, we are interested in figuring out if there are any other special properties that a subgroup must satisfy. In particular, if a group G is finite, is the order of a subgroup related in some way to the order of the whole group? In fact, it is: the order of any subgroup must divide the order of the whole group. This important relationship is called Lagrange's Theorem. In order to prove it, we introduce a special type of subset, called a coset, that can be created from a subgroup.

Definition 2.10. Let $H \leq G$. For any element $g \in G$, the **left** and **right cosets** of H in G are the sets $gH := \{gh : h \in H\}$ and $Hg := \{hg : h \in H\}$, respectively.

One very important property of the left cosets (and also of the right) is that they create a partition of the group G. That is, any two cosets are either disjoint or equal, and the union of all cosets is the entire group. The latter statement is very easy to see, since for all $g \in G$, $g \in gH$. Now, to establish the former assertion, it is helpful to prove first the following lemma.

Lemma 2.11. *Let $H \leq G$ and $g_1, g_2 \in G$. Then $g_1 H = g_2 H$ if and only if there is an $h \in H$ such that $g_1 = g_2 h$.*

Proof. We first prove the forward direction. Since $e \in H$, we know $g_1 \in g_1 H$, which means $g_1 \in g_2 H$ since by assumption the two cosets are equal. Thus, by definition, there is an $h \in H$ such that $g_2 h = g_1$. To prove the other direction, we must show $g_1 H \subseteq g_2 H$. An identical argument will show $g_2 H \subseteq g_1 H$, proving equality. So, pick any $g_1' \in g_1 H$. Then there is an $h' \in H$ such that $g_1' = g_1 h'$. We must show

$g_1' \in g_2H$. But by assumption $g_1 = g_2h$, so

$$g_1' = g_1h' = (g_2h)h' = g_2(hh').$$

Then since $hh' \in H$, we conclude $g_1' \in g_2H$. □

The following proposition completes the proof that the cosets partition the group.

Proposition 2.12. *Let $H \leq G$ and $g_1, g_2 \in G$. If $g_1H \cap g_2H \neq \varnothing$, then $g_1H = g_2H$.*

Proof. By assumption, we can find an element g in $g_1H \cap g_2H$. That is to say, there exist $h_1, h_2 \in H$ such that $g_1h_1 = g = g_2h_2$. But then $g_1 = g_2h_2h_1^{-1}$, so by Lemma 2.11, $g_1H = g_2H$. □

Now, we have all the information we need to prove Lagrange's Theroem.

Theorem 2.13. (Lagrange's Theorem) *If G is a finite group and $H \leq G$, then $|H| \mid |G|$.*

Proof. Let $n = |G|$ and $m = |H|$. Let $S = \{gH : g \in G\}$ be the set of all cosets of H in G. Clearly $|S| \leq n$, because each coset is completely defined in terms of a single element of G. So set $k = |S|$.

Now, note that, by Proposition 2.12,

$$\sum_{gH \in S} |gH| = n.$$

We claim that each coset $gH \in S$ has the same number of elements as H itself. If true, this would imply

$$\sum_{gH \in S} |gH| = km,$$

and therefore $km = n$. Thus, once we prove this claim, we will have shown $m \mid n$, and our proof will be complete.

Pick any $g \in G$, and define the function $\mu_g : H \to gH$ where for all $h \in H$, $\mu_g(h) = gh$. Clearly, for any $gh' \in gH$, $\mu_g(h') = gh'$, so μ_g is surjective. Also,

$$\begin{aligned} \mu_g(h_1) = \mu_g(h_2) &\Rightarrow gh_1 = gh_2 \\ &\Rightarrow h_1 = h_2, \end{aligned}$$

since we can multiply by g^{-1} on the left. This means μ_g is also injective. Then since μ_g is bijective for any g, all cosets have the same cardinality as H, as desired. □

We can use an earlier result (Proposition 2.9) to prove an interesting corollary of Lagrange's Theorem, which will be the last result proven in this section.

Corollary 2.14. *If G is a finite group, then for all $g \in G$, $g^{|G|} = e$.*

Proof. From Proposition 2.9, $ord(g) = |\langle g \rangle| =: n$, and from Lagrange's Theorem, there exists $m \in \mathbb{N}$ such that $|G| = nm$. Then

$$g^{|G|} = (g^n)^m = e^m = e.$$

□

This introduction barely scratches the surface of group theory. However, the rest of this paper hopes to show that even this relatively small set of results has at least *one* interesting application already. We will need only one additional concept: cyclic groups. They are explained in the next section.

3. Cyclicity and $\mathbb{Z}_p^\times$

It is time to introduce a very specific group which will be of central importance in the next section, when we discuss the Legendre symbol. You may recall we mentioned that the integers under multiplication do not form a group. This is true; however, if we look at the integers modulo some prime p and take out the zero element, the remaining equivalence classes *do* form an abelian group under multiplication. We will now give a formal definition of this set and its binary operation, then prove that it satisfies the axioms presented in Definition 2.1. It is assumed that the reader has some familiarity with modular arithmetic.

Definition 3.1. Let p be a prime. Define $\mathbb{Z}_p^\times := \{[1], \ldots, [p-1]\}$ to be the set of equivalence classes of $\mathbb{Z}$ modulo p. Let $\cdot : \mathbb{Z}_p^\times \times \mathbb{Z}_p^\times \to \mathbb{Z}_p^\times$ be a binary operation where for all $([i], [j]) \in \mathbb{Z}_p^\times \times \mathbb{Z}_p^\times$, $[i] \cdot [j]$ maps to $[ij]$.

Proposition 3.2. *The operation $\cdot$ is well-defined. Furthermore, $(\mathbb{Z}_p^\times, \cdot)$ is an abelian group.*

Proof. The only concern about whether $\cdot$ is well-defined is the possibility that two nonzero equivalence classes might map to the zero equivalence class (which is not in $\mathbb{Z}_p^\times$). But this is precluded by the fact that p is prime: suppose $[i], [j] \in \mathbb{Z}_p^\times$. Then $p \nmid i$ and $p \nmid j$, so $p \nmid ij$. This means $[ij] \in \mathbb{Z}_p^\times$.

We now show $\mathbb{Z}_p^\times$ is an abelian group. The identity is the equivalence class [1], and the operation $\cdot$ is both associative and commutative by the associativity and commutativity of multiplication on $\mathbb{Z}$. It remains to show existence of inverses. Fix some $[i] \in \mathbb{Z}_p^\times$. Now, suppose there exist $[j], [k] \in \mathbb{Z}_p^\times$ such that $[ij] = [ik]$. Then, by definition of equality in $\mathbb{Z}_p^\times$, $p \mid (ij - ik)$. But

$$\begin{aligned} p \mid i(j-k) &\Rightarrow p \mid i \text{ or } p \mid (j-k) \\ &\Rightarrow p \mid (j-k) \qquad\qquad (\text{because } p \nmid i) \\ &\Rightarrow [j] = [k]. \end{aligned}$$

Therefore, all of the products $[i1], [i2], \ldots, [ii], \ldots, [i(p-1)]$ are distinct. There are $p-1$ such products. Since there are also $p-1$ elements in $\mathbb{Z}_p^\times$, exactly one of these products must be the identity. Hence, $[i]^{-1}$ exists. □

Henceforth, when discussing elements of $\mathbb{Z}_p^\times$, we will usually discard the brackets. For example, unless otherwise noted, we take 1 to mean [1].

As mentioned above, there are only $p-1$ elements in $\mathbb{Z}_p^\times$, since we took out the zero element. Then, by Lagrange's Theorem (more specifically, by Corollary 2.14), for any $a \in \mathbb{Z}_p^\times$, $a^{p-1} = 1$. This result is more commonly known as Fermat's Little Theorem.

Our ultimate goal in this section is to prove that $\mathbb{Z}_p^\times$ is cyclic. Of course, we must first explain what is meant by this.

Definition 3.3. An element g of a group G is called **primitive** if $\langle g \rangle = G$. A group G is called **cyclic** if it contains a primitive element (also called a **generator**).

Our proof of the cyclicity of $\mathbb{Z}_p^\times$ follows that of Keith Conrad, in his eponymous article on the subject [4]. We will need five lemmas. To begin, let us introduce a useful notation for the number of elements of $\mathbb{Z}_p^\times$ with a given order.

Definition 3.4. Let p be prime. Define the function $N_p : \{d \in \mathbb{N} : d \mid (p-1)\} \to \mathbb{N}$ by $N_p(d) = |\{g \in \mathbb{Z}_p^\times : ord(g) = d\}|$.

In other words, $N_p(d)$ returns the number of elements of $\mathbb{Z}_p^\times$ that have order d. Recall from our discussion of Lagrange's Theorem that the order of any element of $\mathbb{Z}_p^\times$ is a divisor of $p-1$, the order of $\mathbb{Z}_p^\times$. This means that the sum of $N_p(d)$ over all divisors of $p-1$ gives precisely the number of elements in $\mathbb{Z}_p^\times$:

$$\sum_{d|(p-1)} N_p(d) = p - 1. \tag{3.5}$$

This fact will be important later.

Before we begin proving our lemmas, there is one more concept that we must take into consideration. We originally defined the group $(\mathbb{Z}_p^\times, \cdot)$ in terms of multiplication on $\mathbb{Z}$, and the fact is that if we also include addition on $\mathbb{Z}$, then $(\mathbb{Z}_p, +, \cdot)$ becomes a *field*. This fact is crucial because it allows us to talk about polynomials over $\mathbb{Z}_p$ in a well-defined way. Unfortunately, it is beyond the scope of this paper to develop the theory behind fields and polynomials. Some prior experience with them will be helpful (though perhaps not necessary) to make sense of the following two lemmas, which deal with polynomials over a general field F. After proving these, we will bring the discussion back to the group $(\mathbb{Z}_p^\times, \cdot)$ by means of a short corollary.

Lemma 3.6. *Let F be a field and let $f \in F[x]$ have degree d such that $d \geq 1$. If there is an $\alpha \in F$ such that $f(\alpha) = 0$, then there is another function $g \in F[x]$ of degree $d-1$ such that $f(x) = (x-\alpha)g(x)$.*

Proof. Write f as

$$f(x) = a_d x^d + \ldots + a_1 x + a_0, \tag{3.7}$$

where for all i, $a_i \in F$, and $a_d \neq 0$. Now we plug α into this expression to obtain

$$0 = a_d \alpha^d + \ldots + a_1 \alpha + a_0. \tag{3.8}$$

Now we subtract Eq.(3.8) from Eq.(3.7). The constant term cancels, and we are left with

$$f(x) = a_d(x^d - \alpha^d) + \ldots + a_1(x - \alpha). \tag{3.9}$$

For each $i \in \{1, \ldots, d\}$, we note that we can factor $x^i - \alpha^i$ as

$$x^i - \alpha^i = (x-\alpha)(x^{i-1} + x^{i-2}\alpha + \ldots + x\alpha^{i-2} + \alpha^{i-1}).$$

Hence, we can factor out $(x-\alpha)$ from every term in Eq.(3.9), leaving us with an expression of the desired form $f(x) = (x-\alpha)g(x)$. Note that the coefficients of g are sums of coefficients of f, and the coefficient of x^{d-1} in g is a_d, which is nonzero. Thus, g is a polynomial of degree $d-1$ with coefficients in F, as required. □

With this knowledge, proving the inductive step of the following lemma becomes much easier.

Lemma 3.10. *Let F be a field and let $f \in F[x]$ have degree d such that $d \geq 1$. Then f has at most d roots in F.*

Proof. We induct on d.

If $d = 1$, then f has the form $f(x) = ax + b$, where $a \neq 0$. In this case, $f(\alpha) = 0$ implies $\alpha = -a^{-1}b$, so f has exactly one root. In particular, we can say it has at most d roots, and this establishes the base case.

For the inductive step, fix $d \in \mathbb{N}$ and assume that all polynomials of degree d over F have at most d roots. Now pick some $f \in F[x]$ of degree $d+1$. If there is no $\beta \in F$ satisfying $f(\beta) = 0$, then f has no roots in F, and we are done. Otherwise, f has a root $\alpha \in F$, so by Lemma 3.6, there is a $g \in F[x]$ of degree d such that $f(x) = (x-\alpha)g(x)$. Then, since the field F has no zero divisors, any root of f must be a root of either $(x - \alpha)$ or g. By inductive hypothesis, g has at most d roots, and the only root of $(x - \alpha)$ is α itself. Thus, f has at most $d + 1$ roots over F. Induction is complete. □

This lemma dealt with general polynomials over a general field. Now, by applying it to a very *specific* polynomial over a very *specific* field, we will be able to learn something interesting about our group, $\mathbb{Z}_p^\times$.

Corollary 3.11. *For any $d \in \mathbb{N}$, there are at most d distinct elements a in the group $\mathbb{Z}_p^\times$ satisfying $a^d = 1$.*

Proof. Consider the polynomial $f(a) = a^d - 1$ over the field $\mathbb{Z}_p$. From Lemma 3.10, we know that there are at most d roots of f in $\mathbb{Z}_p$. In particular, then, there are at most d roots in the set $\mathbb{Z}_p \setminus \{0\}$. Restated, this means there are at most d distinct numbers in $\mathbb{Z}_p^\times$ satisfying $a^d = 1$, which makes perfect sense even when we view $\mathbb{Z}_p^\times$ as a group rather than as a subset of a field. This is what we wished to show. □

Our next two proofs deal with the Euler φ function. For those unfamiliar with it, it is a function $\varphi : \mathbb{N} \to \mathbb{N}$ where $\varphi(n) = |\{a \in \mathbb{N} : a < n \text{ and } gcd(a, n) = 1\}|$. In other words, $\varphi(n)$ tells how many natural numbers less than n are relatively prime to n. Note that for all $n \in \mathbb{N}$, $\varphi(n) \geq 1$, since 1 is always relatively prime to n.

The following surprising identity of the φ function is important enough to mathematics in general that it merits the name "theorem," although we are only using it as a lemma here.

Theorem 3.12. *For all $n \in \mathbb{N}$, $\sum_{d|n} \varphi(d) = n$.*

Proof. For all $d \mid n$, define $S_d = \{x \leq d : gcd(x, d) = 1\}$. Note that this means $|S_d| = \varphi(d)$. Now let T be the set of ordered pairs $T = \bigcup_{d|n} \{d\} \times S_d$. Note that *this* means $|T| = \sum_{d|n} \varphi(d)$. Let $[n]$ denote the set $\{1, \ldots, n\}$. Finally, define the function $f : T \to [n]$ by $f(d, x) = x(n/d)$. Our claim is that f is bijective. This would imply $|T| = n$, which is exactly the equality we wish to prove.

First we must show that f is well-defined; that is, for all $d \mid n$ and $x \in S_d$, $x(n/d) \in [n]$. Since $d \mid n$, n/d is an integer. Also, $x \in S_d$ implies $1 \leq x \leq d$. Thus, $x(n/d)$ is an integer between 1 and n (inclusive), as desired.

We now show f is injective. Suppose $f(d_1, x_1) = f(d_2, x_2)$. Then

$$\begin{aligned} \frac{n}{d_1}x_1 = \frac{n}{d_2}x_2 &\Rightarrow d_2x_1 = d_1x_2 \\ &\Rightarrow x_1 \mid d_1x_2 \quad \text{and } x_2 \mid d_2x_1 \\ &\Rightarrow x_1 \mid x_2 \quad \text{and } x_2 \mid x_1 \end{aligned}$$

because x_1 and x_2 are relatively prime to d_1 and d_2, respectively. Thus, $x_1 = x_2$, and it follows that $(d_1, x_1) = (d_2, x_2)$.

Now, we prove f is surjective. Pick any $m \in [n]$. Now let $z = gcd(m, n)$, and define $d = n/z$ and $x = m/z$. We claim $x \in S_d$. First of all, we see $m \leq n$, so $x \leq d$. It remains to show that $gcd(x, d) = 1$. Now, since $z = gcd(m, n)$, there exist

$i, j \in \mathbb{Z}$ such that $mi + nj = z$. We then write $m = xz$ and $n = dz$ and substitute to find $(xz)i + (dz)j = z$, and thus $xi + dj = 1$. This shows x and d are relatively prime. Therefore, we can take

$$f(x, d) = \frac{n}{d}x = \frac{n}{n/z}\frac{m}{z} = m,$$

proving that f is surjective. Our proof of the φ identity is complete. □

We now prove an interesting relationship between our function N_p and the φ function.

Lemma 3.13. *Let p be prime and $d \mid (p-1)$. If $N_p(d) \neq 0$, then $N_p(d) = \varphi(d)$.*

Proof. Since $N_p(d) \neq 0$, we can pick some $a \in \mathbb{Z}_p^\times$ such that $ord(a) = d$. Then $\langle a \rangle = \{a, a^2, \ldots, a^d\}$ has d distinct elements. We claim that all $b \in \langle a \rangle$ satisfy $b^d = 1$. Indeed, this follows directly from Corollary 2.14, which states that raising any element of a (sub)group to the order of the (sub)group gives the identity. But from Corollary 3.11, in the entire group there can only be d solutions to the equation $b^d = 1$. Hence, the solutions to $b^d = 1$, of which the elements of order d are a subset, are precisely those elements in $\langle a \rangle$.

We desire to know how many of these elements have order exactly d. To that end, let $k \in \{1, \ldots, d\}$ so that $a^k \in \langle a \rangle$. We claim a^k has order d precisely when $gcd(d, k) = 1$. Since this statement is equivalent to the statement $N_p(d) = \varphi(d)$, if we can prove it, we are done. Suppose first that $gcd(d, k) = 1$. Pick any number $j \in \{1, \ldots, d\}$ and suppose $(a^k)^j = e$. Since $ord(a) = d$, this means $d \mid kj$. But d and k are relatively prime, so we must have $d \mid j$. This can only happen if $j = d$, which proves $ord(a^k) = d$. Now suppose $gcd(d, k) > 1$. Write $z = gcd(d, k)$, and define $l = d/z$. Then l is less than d, but

$$(a^k)^l = a^{kd/z} = (a^d)^{k/z} = e^{k/z} = e,$$

showing that a^k does not have order d. (Note that this was valid because z is a common divisor of d and k, and thus k/z is an integer.) We have shown that $N_p(d) = \varphi(d)$, as desired. □

With these tools, we are finally able to construct a proof of the cyclicity of $\mathbb{Z}_p^\times$, with which we conclude this section.

Theorem 3.14. *The group $\mathbb{Z}_p^\times$ is cyclic.*

Proof. We wish to show that $\mathbb{Z}_p^\times$ has at least one element of order $p-1$. In other words, we must prove $N_p(p-1) \neq 0$. Now, recall from Eq.(3.5) that $\sum_{d|(p-1)} N_p(d) = p-1$. Also, from Theorem 3.12, $\sum_{d|(p-1)} \varphi(d) = p-1$. Therefore,

$$\sum_{d|(p-1)} N_p(d) = \sum_{d|(p-1)} \varphi(d)$$

or

$$\sum_{d|(p-1)} \varphi(d) - N_p(d) = 0. \tag{3.15}$$

Now from Lemma 3.13, if $N_p(d) \neq 0$, then $\varphi(d) - N_p(d) = 0$. On the other hand, if $N_p(d) = 0$, then $\varphi(d) - N_p(d) > 0$, because $\varphi(d)$ is always positive. Therefore all of the terms in Eq.(3.15) are nonnegative, and the only way the sum can be zero is if each term is zero. That is to say, for all $d \mid (p-1)$, we must have $N_p(d) \neq 0$. In particular, $N_p(p-1) \neq 0$, and we are done. □

4. The Legendre Symbol

Our next topic of discussion is the Legendre symbol, which is used frequently in situations involving squares modulo some prime number. Our discussion follows that of Matthew Morrow in his lecture series of the 2011 University of Chicago VIGRE REU [7]. You will notice that in most of this section we will speak of *odd* primes; the number 2 is taken out as an easy special case. Also, as threatened in the previous section, we will be a bit cavalier about whether a natural number a represents itself or its equivalence class $[a]$ in $\mathbb{Z}_p^\times$. Hopefully, the meaning will be clear from the context.

We begin by introducing the concept of quadratic residues.

Definition 4.1. Let p be an odd prime and $a \in \mathbb{Z}$. We say a is a **quadratic residue mod p** if $gcd(a, p) = 1$ and there is an $x \in \mathbb{Z}$ such that $x^2 \equiv a \pmod{p}$.

It will turn out to be helpful to treat separately the case where a is a multiple of p; that is why in the definition we specify that a and p are relatively prime. Also, it is clear that if $a, b \in \mathbb{N}$ and $a \equiv b \pmod{p}$, then a is a quadratic residue mod p if and only if b is also a quadratic residue mod p. Therefore, we can completely describe the quadratic residues of any prime p by listing the quadratic residues in the group $\mathbb{Z}_p^\times = \{1, \ldots, p-1\}$. Here is an example of how to identify these numbers.

Example 4.2. Let us consider the prime 5. To figure out which numbers are quadratic residues mod 5, we simply square a representative of each of the five equivalence classes mod 5 and see what we get.

x:	0	1	2	3	4
x^2:	0	1	4	4	1

Thus, the quadratic residues mod 5 are 1 and 4. (Remember 0 is not counted.)

In this example, we saw that exactly half of the elements of $\mathbb{Z}_5^\times$ were quadratic residues mod 5. This is actually a general result.

Proposition 4.3. *Let p be an odd prime. Then there are exactly $\frac{p-1}{2}$ distinct quadratic residues mod p in $\mathbb{Z}_p^\times$.*

Proof. By definition, the quadratic residues mod p are exactly the elements in $\{k^2 : k \in \mathbb{Z}_p^\times\}$. Now, pick any $i, j \in \mathbb{Z}_p^\times$. Then

$$\begin{aligned} i^2 \equiv j^2 \pmod{p} &\iff p \mid (i^2 - j^2) = (i+j)(i-j) \\ &\iff p \mid (i+j) \text{ or } p \mid (i-j). \end{aligned}$$

But $i, j \in \{1, \ldots, p-1\}$, so $p \mid (i+j)$ iff $i + j = p$, and $p \mid (i-j)$ iff $i = j$. Thus for any $k \in \mathbb{Z}_p^\times$, there exists exactly one other number $l \in \mathbb{Z}_p^\times$, where $l = p - k$, such that $k^2 \equiv l^2 \pmod{p}$. Therefore, the set $\{k^2 : k \in \mathbb{Z}_p^\times\}$ has exactly $\frac{p-1}{2}$ distinct elements, as desired. □

The Legendre symbol is defined simply as a way to encode whether or not a given number is a quadratic residue modulo some prime p. It is a helpful shorthand to have when stating properties of quadratic residues.

Definition 4.4. Let p be an odd prime and $a \in \mathbb{Z}$. The **Legendre symbol of a on p** is defined as

$$\left(\frac{a}{p}\right) = \begin{cases} 0, & p \mid a \\ 1, & a \text{ is a quadratic residue mod } p \\ -1, & \text{otherwise.} \end{cases}$$

The next result is very helpful in proving properties of the Legendre symbol, and it is important enough that it has its own name. Unfortunately, it is also a bit tricky to prove. To crack it, we will need to use the main result of the previous section—the cyclicity of $\mathbb{Z}_p^\times$.

Proposition 4.5. (Euler's Lemma) *Let p be an odd prime and $a \in \mathbb{Z}$. Then*

$$\left(\frac{a}{p}\right) \equiv a^{\frac{p-1}{2}} \pmod{p}.$$

Proof. The case $p \mid a$ is trivial, so assume $a \in \mathbb{Z}_p^\times$ (where we are using a to represent its own equivalence class). Since $\mathbb{Z}_p^\times$ is cyclic, let g be a generator of $\mathbb{Z}_p^\times$, and write $a = g^r$. Consider the set $S = \{g^{2n} : n \in \{1, \ldots, \frac{p-1}{2}\}\}$; that is, the even powers of g. Each element of S is a quadratic residue mod p, since $g^{2n} = (g^n)^2$. Also, since g is a generator, S has $\frac{p-1}{2}$ distinct elements. But from Proposition 4.3, there can only be $\frac{p-1}{2}$ distinct quadratic residues mod p. Thus $a = g^r$ is a quadratic residue mod p if and only if r is even. In terms of the Legendre symbol,

$$\left(\frac{a}{p}\right) = (-1)^r.$$

We now wish to express -1 in terms of g; that is, we wish to find the number $m \in \{1, \ldots, p-1\}$ satisfying $g^m \equiv p-1 \pmod{p}$. Now, $g^m \equiv p-1 \pmod{p}$ implies $g^{2m} \equiv p^2 - 2p + 1 \equiv 1 \pmod{p}$, which means $(p-1) \mid 2m$, since $ord(g) = p-1$. But $m \in \{1, \ldots, p-1\}$, so this can only be true if $m = \frac{p-1}{2}$ or $m = p-1$. But the case $m = p-1$ is ruled out, because we already know $g^{p-1} = 1$, and if p is an odd prime then $1 \not\equiv p-1 \pmod{p}$. In conclusion, then,

$$\left(\frac{a}{p}\right) \equiv (g^{\frac{p-1}{2}})^r = (g^r)^{\frac{p-1}{2}} \equiv a^{\frac{p-1}{2}} \pmod{p}.$$

□

As a matter of mathematical interest, we will now use this to prove that the Legendre symbol is multiplicative. We will not need the following proposition to prove Fermat's two-squares theorem, but it is such an important property of the Legendre symbol that it would be negligent to omit it.

Proposition 4.6. *Let p be an odd prime and $a, b \in \mathbb{Z}$. Then*

$$\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right).$$

Proof. We apply Euler's lemma:

$$\left(\frac{ab}{p}\right) \equiv (ab)^{\frac{p-1}{2}} = a^{\frac{p-1}{2}} b^{\frac{p-1}{2}} \equiv \left(\frac{a}{p}\right)\left(\frac{b}{p}\right) \pmod{p}.$$

But both sides can only take on the values -1, 0, or 1, and $p \geq 3$, so they are congruent mod p if and only if they are equal. □

Now, we reach the last proposition of this section. For reasons which we will not delve into here, it is often called the first supplementary law to quadratic reciprocity. This result is the key to the first step in our proof of Fermat's two-squares theorem, as we will see quite soon.

Proposition 4.7. *Let p be an odd prime. Then*

$$\left(\frac{-1}{p}\right) = \begin{cases} 1, & p \equiv 1 \pmod 4 \\ -1, & p \equiv 3 \pmod 4. \end{cases}$$

Proof. We apply Euler's lemma:

$$\left(\frac{-1}{p}\right) \equiv (-1)^{\frac{p-1}{2}} \pmod p.$$

But, again, both sides can only take on the values -1 or 1, and $p \geq 3$, so they are congruent mod p if and only if they are equal. To conclude, we simply note that the exponent $(p-1)/2$ is even when $p \equiv 1 \pmod 4$, and odd when $p \equiv 3 \pmod 4$. □

5. A Congruence Condition for Primes of the Form $x^2 + y^2$

We are finally ready to prove Fermat's theorem on sums of two squares. Our proof is adapted from a solution written by Michael Bennett [2].

Theorem 5.1. *Let p be a prime. There exist $x, y \in \mathbb{Z}$ such that $p = x^2 + y^2$ if and only if $p = 2$ or $p \equiv 1 \pmod 4$.*

Proof. Let us first prove the simpler forward implication. We will do this by showing that if $p \equiv 3 \pmod 4$, then for any $x, y \in \mathbb{Z}$, $p \not\equiv x^2 + y^2 \pmod 4$. Now, if $x \equiv 0$ or $x \equiv 2 \pmod 4$, then $x^2 \equiv 0 \pmod 4$. Similarly, if $x \equiv 1$ or $x \equiv 3 \pmod 4$, then $x^2 \equiv 1 \pmod 4$. Hence, for any $x, y \in \mathbb{Z}$, $x^2 + y^2 \equiv 0$, 1, or 2 $\pmod 4$. This proves that if $p \equiv 3 \pmod 4$, then it cannot be expressed as a sum of two squares.

We now prove the reverse implication. Clearly $2 = 1^2 + 1^2$, so for $p = 2$ the theorem holds. Now assume $p \equiv 1 \pmod 4$. Then by Proposition 4.7, the number -1 is a quadratic residue mod p. Thus, we can find an $x \in \{1, \ldots, p-1\}$ such that $x^2 \equiv -1 \pmod p$. In other words, there is a $k \in \mathbb{N}$ such that $kp = x^2 + 1^2$. Furthermore,

$$kp = x^2 + 1 \leq (p-1)^2 + 1 = p^2 - 2p + 2 < p^2,$$

since $p > 1$. So we can say $k < p$.

Now define $m = min\{n \in \mathbb{N} : (\exists x, y \in \mathbb{Z})(np = x^2 + y^2)\}$. Since k is in this set, $m \leq k < p$. To finish the proof, we wish to show that $m = 1$.

Let u and v be the integers satisfying $mp = u^2 + v^2$. Consider the set

$$S = \left\{ \lfloor -\frac{m}{2} + 1 \rfloor, \ldots, \lfloor \frac{m}{2} \rfloor \right\},$$

which is a set of representatives of $\mathbb{Z}$ modulo m. Now let a and b be the unique integers in S satisfying $a \equiv u \pmod m$ and $b \equiv v \pmod m$. Since $m \mid (u^2 + v^2)$, we know $m \mid (a^2 + b^2)$, so we can write $lm = a^2 + b^2$ for some $l \in \mathbb{Z}$. Note that l is nonnegative, since m is positive. Because we chose a and b so that $|a|, |b| \leq \frac{m}{2}$, we know

$$lm = a^2 + b^2 \leq \left(\frac{m}{2}\right)^2 + \left(\frac{m}{2}\right)^2 = \frac{m^2}{2} < m^2,$$

and thus $l < m$.

Now we write

$$(lm)(mp) = (a^2 + b^2)(u^2 + v^2) = (au + bv)^2 + (av - bu)^2, \tag{5.2}$$

which we obtained by rearranging the terms in the product $(a^2+b^2)(u^2+v^2)$. Now we notice that $au+bv \equiv a^2+b^2 \equiv 0 \pmod{m}$ and $av-bu \equiv ab-ba \equiv 0 \pmod{m}$. This means $\frac{au+bv}{m} \in \mathbb{Z}$ and $\frac{av-bu}{m} \in \mathbb{Z}$, so we can divide Eq.(5.2) by m^2 to obtain

$$lp = \left(\frac{au+bv}{m}\right)^2 + \left(\frac{av-bu}{m}\right)^2,$$

which expresses lp as a sum of square numbers. But since $0 \leq l < m$ and m is the minimum positive multiple of p which can be expressed as a sum of square numbers, it must be that $l = 0$. This in turn implies $a = b = 0$, which means $m \mid u$ and $m \mid v$. Then we can rewrite the equation $mp = u^2 + v^2$ as $p = m[(\frac{u}{m})^2 + (\frac{v}{m})^2]$, which means $m \mid p$. This can only be true if $m = 1$ or $m = p$, but we already know that $m < p$. Thus, $m = 1$, as desired. □

It might seem like the amount of background preparation in this paper was disproportionate to the goal, especially given that the only previous result that we use in this proof is Proposition 4.7, and its role may seem rather minor. On the contrary, though, it established a crucial first step and provided the bound that allowed us to finish the proof. It is one thing to know that we can find a k such that $p \mid (k^2 + 1)$; it is another thing entirely to know *why* such a k can be found, and this is what the background theory explains.

By finding an expression for $(\frac{-2}{p})$ in terms of congruence conditions for p, our method could be generalized slightly to find which primes can be expressed in the form $x^2 + 2y^2$. However, there are limitations to the strategy we used. As mentioned in the introduction, the general case of primes of the form $x^2 + ny^2$ has been solved, and the solution lies in the realm of class field theory. As a matter of fact, using the theory of Gaussian integers, one can devise a rather shorter and cleaner proof of Theorem 5.1 than the one we used above. The article "Sums of Two Squares" by Pete Clark provides one example [3]. For more information on the generalizations of this problem, the interested reader is advised to obtain a copy of the aforementioned book *Primes of the Form* $x^2 + ny^2$, by D. A. Cox [5].

Acknowledgments. It is a pleasure to thank my mentors, Nick Ramsey, Matthew Wright, and Eric Astor, for their guidance and encouragement throughout the development of this paper. I would also like to thank Dr. Matthew Morrow and Dr. Laszlo Babai for introducing me to group theory and the Legendre symbol in such a clear and interesting way. Finally, I owe great thanks to Dr. Peter May for organizing the University of Chicago VIGRE REU, without which opportunity this paper would never have been possible.

References

[1] Babai, Laszlo. "Apprentice Program: Linear Algebra and Combinatorics." Lecture series, Univ. Chicago Math Dept. VIGRE REU, June–July 2011. <http://people.cs.uchicago.edu/∼laci/REU11/appr.html>.

[2] Bennett, Michael. "Homework 2 Solutions." Univ. British Columbia, Math 313, Section 201, Spring 2009. <http://www.math.ubc.ca/∼bennett/Math313/HW2Sol.pdf>.

[3] Clark, Pete L. "Sums of Two Squares." Course notes, Univ. Georgia, Math 4400/6400, 2009. <http://www.math.uga.edu/∼pete/4400twosquares.pdf>.

[4] Conrad, Keith. "Cyclicity of $(\mathbf{Z}/(p))^{\times}$." <http://www.math.uconn.edu/∼kconrad/blurbs/grouptheory/cyclicFp.pdf>.
[5] Cox, David A. Primes of the Form $x^2 + ny^2$. New York: Wiley, 1989.
[6] "Fermat's theorem on sums of two squares." Wikipedia.org. 5 July 2011. 1 August 2011 <http://en.wikipedia.org/wiki/Fermat's_theorem_on_sums_of_two_squares>.
[7] Morrow, Matthew. "Number Theory: Reciprocity and Polynomials." Lecture series, Univ. Chicago Math Dept. VIGRE REU, June 2011. <http://math.uchicago.edu/∼mmorrow/REU.pdf>.