Title
On the Applications of Cyclotomic Fields in Introductory Number Theory

Permalink
https://escholarship.org/uc/item/49c0561x

Journal
Berkeley Scientific Journal, 17(1)

ISSN
1097-0967

Author
Gaspard, Kabalan

Publication Date
2013

DOI
10.5070/BS3171016284

On the Applications of Cyclotomic Fields in Introductory Number Theory

Kabalan Gaspard

Mathematical Institute, Oxford University

Keywords: Number Theory, Cyclotomic Fields, Pell's Equation, Dirichlet, Kummer, Fermat's Last Theorem, Primary Units, Real Units.

Abstract. In this essay, we study and comment on two number theoretical applications of prime cyclotomic fields (cyclotomic fields obtained by adjoining a primitive p-th root of unity to $\mathbb{Q}$, where p is an odd prime). We begin by giving a simplified proof of Kummer's case of Fermat's Last Theorem obtained by linking different versions of the proof in different textbooks. We finally modernize Dirichlet's solution to Pell's Equation.

Throughout this paper, unless specified otherwise, $\zeta = \zeta_p := e^{\frac{2\pi\sqrt{-1}}{p}}$ where p is an odd prime. $K = \mathbb{Q}(\zeta)$ and $\mathcal{O}_K$ is the ring of integers of K. We assume knowledge of the basic properties of prime cyclotomic fields that can be found in most introductory algebraic number theory textbooks (e.g. [5, Section 13.2, pp. 193-9], [7, Section 3.2, pp. 64-9]), namely that:

- $Gal(K:\mathbb{Q})$ is isomorphic to $U(\mathbb{Z}/p\mathbb{Z})$ (the group of units of $\mathbb{Z}/p\mathbb{Z}$), which is cyclic and of order $p-1$.
- $\mathcal{O}_K = \mathbb{Z}[\zeta] = \langle 1, \zeta, ..., \zeta^{p-2}\rangle_{\mathbb{Z}}$, where $\{1, \zeta, ..., \zeta^{p-2}\}$ is a $\mathbb{Z}$-basis for $\mathcal{O}_K$.
- The only roots of unity in $\mathcal{O}_K$ (i.e. solutions in $\mathbb{C}$ to $x^n = 1$ for some $n \in \mathbb{N}$) are of the form $\pm\zeta^i$, with $i \in \mathbb{Z}$.

We also assume elementary knowledge of quadratic characters, quadratic reciprocity, and the Legendre symbol $\left(\frac{k}{p}\right)$.

1. Application to a Special Case of Fermat's Last Theorem

Fermat's well-known final theorem, proved by Andrew Wiles and Richard Taylor in 1994, states that

$$x^n + y^n = z^n$$

where $x, y, z, n \in \mathbb{Z}$ has no non-trivial solutions (x, y, z) for $n \geq 3$.

In fact, to prove this theorem, it suffices to prove that $x^p + y^p = z^p$ has no integral solutions for any positive odd prime p, since $x_0^n + y_0^n = z_0^n$ yields $x_1^p + y_1^p = z_1^p$ where p is an odd prime dividing n (which exists since $n \geq 3$) and $(x_1, y_1, z_1) = (x_0^{n/p}, y_0^{n/p}, z_0^{n/p})$. In other words, we can restrict our study to the case where n is an odd prime.

There is a very elegant proof of a special case of this theorem using cyclotomy. The main use of the concept here is that it allows us to transform a "sum of n-th powers" problem into a "divisibility" problem since we can now factor $x^p + y^p$ as $\prod_{i=0}^{p-1}(x + \zeta^i y)$.

In this section, we shall lay out said proof. We will suppose that (x_0, y_0, z_0) is a solution to $x^p + y^p = z^p$ for some odd prime p. Then

$$x_0^p + y_0^p = z_0^p \tag{1.1}$$

Without loss of generality, we can take x_0, y_0 and z_0 to be pairwise relatively prime, for if some integer d divides two of them, it must divide the third, and then $x_0^p + y_0^p = z_0^p$ is equivalent to $x_1^p + y_1^p = z_1^p$ where $(x_0, y_0, z_0) = d \cdot (x_1, y_1, z_1)$, with x_1, y_1, z_1 in $\mathbb{Z}$.

We shall now reduce the problem to a special case and suppose that p *does not divide the class number* h *of* $\mathcal{O}_K$, and that p does not divide $x_0 y_0 z_0$. From (1.1), we shall reach a contradiction. We will first need the following proposition:

Proposition 1. *Let $i \not\equiv j \pmod p$. Then the ideals $I := (x_0 + \zeta^i y_0)$ and $J := (x_0 + \zeta^j y_0)$ are relatively prime.*

Proof. Consider the ideal $I+J$. J contains the element $-(x_0+\zeta^j y_0)$, so $x_0+\zeta^i y_0-(x_0+\zeta^j y_0) = (\zeta^i-\zeta^j)y_0$ is in $I+J$. Likewise, since $\mathcal{O}_K = \mathbb{Z}[\zeta]$, we have that $-\zeta^j(x_0+\zeta^i y_0) = \zeta^j x_0+\zeta^{i+j} y_0$ is in I and that $\zeta^i(x_0+\zeta^j y_0) = \zeta^i x_0+\zeta^{i+j} y_0$ is in J. So $\zeta^i x_0+\zeta^{i+j} y_0-\zeta^j(x_0+\zeta^i y_0) = (\zeta^i-\zeta^j)x_0$ is in $I+J$. Now x_0 and y_0 are relatively prime rational integers, and so there exist a and b in $\mathbb{Z}$ such that $ax_0 + by_0 = 1$. So $a(\zeta^i - \zeta^j)x_0 + b(\zeta^i - \zeta^j)y_0$ is in $I + J$, implying that $(\zeta^i - \zeta^j)$ is in $I + J$.

Now

$$\begin{aligned}
(N(\zeta^i - \zeta^j))^2 &= \prod_{k=1}^{p-1}(\zeta^{ik} - \zeta^{jk})^2 \\
&= \prod_{k=1}^{p-1}(-\zeta^{-k(j-i)})(1 - \zeta^{k(j-i)})^2 \\
&= \prod_{k=1}^{p-1}(-\zeta^{-k})(1 - \zeta^{k})^2 \\
&= +\zeta^{-p\frac{p-1}{2}} \prod_{k=1}^{p-1}(1 - \zeta^{k})^2 \\
&= 1 \cdot \left(\sum_{k=1}^{p-1} 1\right)^2 \\
&= p^2.
\end{aligned}$$

So $N(\zeta^i - \zeta^j) = p$, implying that $N(I + J)$ divides p. If $N(I + J) = p$, then we would have p dividing $N(I)$ since I is contained in $I + J$. But then that would imply that p divides $\prod_{i=0}^{p-1}(x_0 + \zeta^i y_0) = x_0^p + y_0^p = z_0^p$, yielding that p divides z_0 since p is prime. But this contradicts our initial assumption. We must therefore have that $N(I + J)$ equals 1, and therefore that $I + J$ is $\mathcal{O}_K$. But $P \mid I$ and $P \mid J$, so $P \mid I + J$. So P must be all of $\mathcal{O}_K$, and so I and J are coprime. □

Now $x_0^p + y_0^p = z_0^p$ implies that $\prod_{i=0}^{p-1}(x_0 + \zeta^i y_0) = (z_0)^p$ as ideals. But $\{(x_0 + \zeta^i y_0) : 0 \leq i \leq p-1\}$ are pairwise coprime by Proposition 1. So by unique factorization of ideals, each of these ideals must be a p-th power. So in particular, taking $i = 1$, $(x_0 + \zeta y_0) = \mathfrak{I}^p$ for some ideal $\mathfrak{I}$. So since $(x_0 + \zeta y_0)$ is principal, $[\mathfrak{I}]$ has order dividing p in the ideal class group, but since p does not divide h by assumption, we must have that the order of $[\mathfrak{I}]$ is 1. So $\mathfrak{I}$ is principal. Let $\mathfrak{I} = (\alpha)$. Then $(x_0 + \zeta y_0) = (\alpha^p)$, and so $x_0 + \zeta y_0$ is associate to α^p. We write $x_0 + \zeta y_0 = u\alpha^p$ where u is a unit in $\mathcal{O}_K$. We now introduce the following lemma:

Lemma 1. *Let u be a unit in $\mathcal{O}_K$. Then $u = \overline{u}\zeta^t$ for some $t \in \mathbb{Z}$*

Proof. Define v by $u = v\overline{u}$. Conjugation is a Galois automorphism on $\mathcal{O}_K$ since $\overline{\zeta} = \zeta^{-1} = \zeta^{p-1}$. So $\overline{u}$ is also a unit, and so v is in $\mathcal{O}_K$. Now let $\sigma_1, ..., \sigma_{p-1}$ be the $(p-1)$ Galois automorphisms on $\mathcal{O}_K$ such that $\sigma_k(\zeta) = \zeta^k$, with $k \in \mathbb{Z}$. Then for all $1 \leq k \leq (p-1)$, we have that $\sigma_k v = \frac{\sigma_k u}{\sigma_k \overline{u}} = \frac{\sigma_k u}{\overline{\sigma_k u}}$ by the above remark. So $|\sigma_k v| = \sigma_k v \cdot \overline{\sigma_k v} = 1$. So $|\sigma_k v|^n = 1$ for any $n \in \mathbb{N}$.

Now consider the polynomial $f(x) = \prod_{k=1}^{p-1}(x - \sigma_k v)$. The coefficients of this polynomial are elementary symmetric polynomials in $\{\sigma_k v : 1 \leq k \leq p-1\}$, and so are invariant under action by $Gal(K:Q) = \{\sigma_k : 1 \leq k \leq p-1\}$. So $f(x)$ is in $\mathbb{Z}[x]$. But then the coefficient of x^k is $s_{(p-1)-k}$ where s_j is the j^{th} elementary symmetric polynomial. But by the previous paragraph, $\left|s_{(p-1)-k}\right| \leq \sum_{j=1}^{p-1-k} |\sigma_k v|^k \leq p-1-k$. So there are finitely many possible such $f(x)$ with integer coefficients since the coefficients are bounded. So there are finitely many possible roots since a polynomial of finite degree has a finite number of roots. But $|\sigma_k v^n| = 1$ for any $n \in \mathbb{N}$, so the $\{v^n : n \in \mathbb{N}\}$ satisfy the same argument. So we must have $v^n = v^{n'}$ for some n, n' in $\mathbb{Z}$. So $v^{n-n'} = 1$, and it follows that v is a root of unity in $\mathcal{O}_K$.

So by the basic properties of prime cyclotomic fields, we must have $v = \pm\zeta^t$ for some $t \in \mathbb{Z}$. Now set $\lambda := 1 - \zeta$, and consider congruence modulo λ. Then since $\dfrac{1-\zeta^k}{1-\zeta} = \sum_{i=1}^{k-1} \zeta^i$, which is in $\langle 1, \zeta, ..., \zeta^{p-2}\rangle_{\mathbb{Z}} = \mathcal{O}_K$, we get $\zeta^k \equiv 1 \pmod{\lambda}$ for all $k \in \mathbb{Z}$. So since $\overline{\zeta^k} = \zeta^{-k}$ is congruent to 1, which is itself congruent to ζ^k modulo λ, $\alpha \equiv \overline{\alpha} \pmod{\lambda}$ for all $\alpha \in \mathcal{O}_K$.

Namely, $u \equiv \overline{u} \equiv \pm u \pmod{\lambda}$ since $\overline{u} = \pm\zeta^{-t}u$. So if $v = -\zeta^t$, then $u \equiv -u \pmod{\lambda}$, which implies $2u \equiv 0 \pmod{\lambda}$. But this is impossible since $N(\lambda) = p$ does not divide $N(2u) = 2^{p-1}$, since p is odd. So $v = +\zeta^t$. □

Then by the above lemma we have $u = \zeta^t\overline{u}$ for some t in $\mathbb{Z}$. If $t \equiv 0 \pmod{p}$ then u is trivially real. Otherwise we set c such that $2c \equiv t \pmod{p}$ (this exists since p is a rational prime). Then $u = \zeta^{2c}\overline{u}$, implying $\zeta^{-c}u = \overline{\zeta^{-c}u}$ and so $\zeta^{-c}u$ is real. So in any case, there exists a rational integer c such that $u_0 := \zeta^{-c}u$ is real (and is obviously a unit).

So $x_0 + \zeta y_0 = \zeta^c u_0 \alpha^p$ where u_0 is real. Working modulo p, we note that $\alpha^p \equiv \left(\sum_{i=0}^{p-2} a_i \zeta^i\right)^p \equiv \sum_{i=0}^{p-2} a_i^p \zeta^{ip} \equiv \sum_{i=0}^{p-2} a_i^p$ which is in $\mathbb{Z}$. So $\alpha^p \equiv \overline{\alpha^p} \pmod{p}$. It follows that $x_0 + \zeta y_0 = \zeta^c u_0 \alpha^p$. So taking conjugates we see that $x_0 + \zeta^{-1}y_0 \equiv \zeta^{-c}u_0\alpha^p \pmod{p}$, and multiplying by $\zeta^{\pm c}$ we get $\zeta^{-c}x_0 + \zeta^{1-c}y_0 \equiv u_0\alpha^p$ and $\zeta^c x_0 + \zeta^{c-1}y_0 \equiv u_0\alpha^p \pmod{p}$. Subtracting the latter congruence from the former yields

$$\zeta^{-c}x_0 + \zeta^{1-c}y_0 - \zeta^c x_0 - \zeta^{c-1}y_0 \equiv 0 \pmod{p} \tag{1.2}$$

Now an element of $\mathcal{O}_K = \mathbb{Z}[\zeta]$ is divisible by p if and only if all of the coefficients as a polynomial in ζ are divisible by p. p does not divide x_0 or y_0 since it doesn't divide $x_0 y_0 z_0$, so we must check the cases where one of $\{c, -c, 1-c, c-1\}$ is congruent to -1 modulo p or where two of $\{c, -c, 1-c, c-1\}$ are equal modulo p. These cases can be split as follows:

- $c \equiv 0 \pmod{p}$ (so that $c \equiv -c \pmod{p}$). Then p divides $y_0(\zeta - \zeta^{-1}) = y_0(\sum_{i=2}^{p-2} \zeta^i + 1)$, so p divides y_0 (even if $p = 3$), giving a contradiction.
- $c \equiv 1 \pmod{p}$ (so that $1 - c \equiv c - 1 \pmod{p}$). Then p divides $x_0(\zeta^{-1} - \zeta)$, so $p \mid x_0$ as in the previous case. We then have a contradiction.
- $c \equiv 2^{-1} \pmod{p}$ (so that $c \equiv 1 - c \pmod{p}$). Then p divides $(y_0 - x_0)\zeta^c + \zeta^{-c}(x_0 - y_0)$. So p divides $(x_0 - y_0)$. We then rewrite (1.1) as $x_0^p + (-z_0)^p = (-y_0)^p$ (since p is odd). Then with the same argument we will get $p \mid (x_0 + z_0)$. But (1.1) yields

$x_0^p + y_0^p - z_0^p \equiv 0 \pmod p$ and so $x_0 + y_0 - z_0 \equiv 0 \pmod p$. This yields $3x_0 \equiv 0 \pmod p$. We suppose for now that $p > 3$. Then this yields that p divides x_0. We then get yet another contradiction.

- Letting one of $\{c, -c, 1-c, c-1\}$ be congruent to -1 modulo p will yield one of the coefficients of the terms of (1.2) as $\pm(x_0 - y_0)$, giving the same contradiction as in the previous case.

We therefore obtain a contradiction in all cases. We have, however, supposed that $p > 3$. A general study of the case where $p = 3$ is done elegantly in section 10 of [4].

This is essentially a simplified version of the argument that is given in [5, Section 17.11, pp. 290-2], which features differently in [1, Section 3.1, pp. 160-3] and other number theory textbooks. However, while Ireland & Rosen use the fact that u_0 is real, Borevich & Shafarevich use instead the concept of a *primary* unit. We shall prove that the set of real units and the set of primary units are in fact identical in $\mathbb{Z}[\zeta]$ in the following paragraph, effectively

1.1. Remark: Primary elements in $\mathcal{O}_K$.

Definition 1. *Let $\alpha \in \mathcal{O}_K$ with α prime to p. Then α is* primary *if and only if α is congruent to a rational integer modulo* $(1-\zeta)^2$.

In fact, the definition of primary elements has historically been ambiguous in number theory. In [2], Dalawat shows that definitions of primary elements in $\mathcal{O}_K$ even differ by country ("p-primary", "*primaire*" and "*primär*") and, even though these definitions do form a chain of implications, they are not equivalent.

We also note that it is not true that if p an arbitrary odd prime and μ prime in $\mathcal{O}_K$, only one associate of μ is primary (for example, according to the above definition, both $\pm(4+3\omega)$ are primary in the ring of integers of $\mathbb{Q}(\omega)$ where $\omega = e^{\frac{2\pi\sqrt{-1}}{3}}$).

Proposition 2. *Let $\alpha \in \mathcal{O}_K$ (not necessarily prime) and suppose α prime to p in $\mathcal{O}_K$. Then there exists a k in $\mathbb{Z}$, unique modulo p, such that $\zeta^k\alpha$ is primary.*

Proof. Consider the ideal $P := (1-\zeta)$ in $\mathcal{O}_K$. Then the norm $N(P)$ of the ideal is $\prod_{i=1}^{p-1}(1-\zeta^i) = p$ by the fact that $Gal(K : \mathbb{Q})$ is isomorphic to $U(\mathbb{Z}/p\mathbb{Z})$. So P is a prime ideal and is thus of degree 1. So by Dedekind's Theorem in algebraic number theory, any element of $\mathcal{O}_K$ is the root of a monic polynomial of degree 1 in $\mathcal{O}_K/P$. So in the particular case of α, we have $\alpha - a_0 = 0$ in $\mathcal{O}_K/P$ for some a_0 in $\mathbb{Z}$. In other words, $\alpha \equiv a_0 \bmod (1-\zeta)$. Now set $\lambda := 1-\zeta$ for simplicity. So $\frac{\alpha - a_0}{\lambda} \in \mathcal{O}_K$ and so, by the same argument, $\frac{\alpha - a_0}{\lambda} \equiv a_1 \bmod \lambda$ for some a_1 in $\mathbb{Z}$. We stop repeating this here because multiplying the congruence by λ, we now have a congruence modulo λ^2, which is what we want to consider. More precisely, we now have $\alpha - a_0 \equiv a_1\lambda \bmod \lambda^2$, so $\alpha \equiv a_0 + a_1\lambda \bmod \lambda^2$.

We want to eliminate the λ (i.e. degree 1) term by multiplying both sides by ζ^n for some $n \in \mathbb{Z}$. Notice that $\zeta = 1 - \lambda$. So modulo λ^2,

$$\begin{aligned}
\zeta^n\alpha &\equiv \zeta^n a_0 + a_1\zeta^n\lambda \\
&\equiv a_0(1-\lambda)^n + a_1\lambda(1-\lambda)^n \\
&\equiv a_0(1-n\lambda) + a_1\lambda(1-n\lambda)
\end{aligned}$$

since considering $(1-\lambda)^n$ as a polynomial in λ, λ^2 divides λ^i for $i \geq 2$. So

$$\zeta^n\alpha \equiv a_0 + (a_1 - na_0)\lambda \bmod \lambda^2$$

Now α prime to p, so if $a_0 \equiv 0 \pmod p$, then $a_0 \equiv 0 \pmod \lambda$, and so $\alpha \equiv 0 \pmod \lambda$, which is a contradiction. So $a_0 \not\equiv 0 \pmod p$, and so $a_1 - na_0 \equiv 0$ has a **unique** solution k modulo p. Now $(1-\zeta)$ divides $(1-\zeta^2)$, and $N(\frac{1-\zeta^2}{1-\zeta}) = \frac{N(1-\zeta^2)}{N(1-\zeta)} = 1$, so $(1-\zeta^2)$ is associate to $(1-\zeta)$. It follows that $(1-\zeta)^2$ divides p, and so k is (still, since $a_1 - na_0 \in \mathbb{Z}$) the **unique** integral solution modulo p to $a_1 - na_0 \equiv 0 \bmod \lambda^2$. Then $\zeta^k\alpha \equiv a_0 \bmod \lambda^2$, and therefore $\zeta^k\alpha$ is primary. □

Theorem 1. *Let u be a unit in $\mathcal{O}_K$. Then u is real $\Leftrightarrow$ u is primary in $\mathcal{O}_K$.*

Proof. Since $\mathcal{O}_K = \mathbb{Z}[\zeta] = \langle 1, \zeta, \ldots, \zeta^{p-2} \rangle_{\mathbb{Z}}$, we can write u as $\sum_{k=0}^{p-2} a_k \zeta^k$ for unique $a_0, \ldots, a_{p-2} \in \mathbb{Z}$. And so, noting that $\zeta^{p-1} = -\sum_{i=0}^{p-2} \zeta^i$, we see that $\zeta^{-t}u = \sum_{k=0}^{p-2} a_k \zeta^{k-t} = \sum_{k=0}^{p-2} (a_{k+t} - a_{(p-1)+t}) \zeta^k$ where a_k is defined to be $a_{(k \bmod p)}$ for all k in $\mathbb{Z} - \{0, \ldots, p-1\}$. Even when $t = 0$ and we have no ζ^{p-1} term, this still works since $a_{p-1} = 0$. So now $\sum_{k=0}^{p-2} (a_{p-k} - a_1) \zeta^k = \overline{u} = \zeta^{-t} u = \sum_{k=0}^{p-2} (a_{k+t} - a_{(p-1)+t}) \zeta^k$ by Lemma 1 and therefore, since this representation is unique, we get

$$a_{k+t} - a_{(p-1)+t} = a_{p-k} - a_1 \text{ for all } 0 \leq k \leq p-1 \tag{1.3}$$

Letting k_0 be the $\bmod\, p$ solution to $k + t \equiv p - k \pmod{p}$, we get $a_{k_0+t} = a_{p-k_0}$ and so (1.3) yields $a_{(p-1)+t} = a_1$. (1.3) then becomes

$$a_{k+t} = a_{p-k} = a_{-k} \text{ for all } 0 \leq k \leq p-1 \tag{1.4}$$

Since replacing k by $-(k+t)$ in (1.4) leaves the equation invariant, we get $\frac{p-1}{2}$ pairs of equal terms with distinct indices amongst $a_0, \ldots, a_{p-1}$ (the 'remaining' term being a_{k_0+t}). Let $b_1, \ldots, b_{\frac{p-1}{2}}$ be representatives of these distinct pairs, and let $b_{k_0+t} = a_{k_0+t}$ (we have simply selected and reordered the a_i's).

Now set $\lambda := 1 - \zeta$ as in the previous proof. By the proof of Proposition 2, there is a unique c modulo p such that $\zeta^c u$ is primary, and this c is the solution to $ax \equiv b \pmod{p}$ where $u \equiv a + b\lambda \pmod{\lambda^2}$. To find a and b, we define the polynomial $f(x) := \sum_{k=0}^{p-2} a_k x^k$. This notation yields that a and b are coefficients of the terms in 1 and x respectively in $f(1-x)$, since $u = f(\zeta) = f(1-\lambda)$. Making elementary use of the Binomial Theorem, we see that $f(1-x) = \sum_{k=0}^{p-2} a_k (1-x)^k = \sum_{k=0}^{p-2} a_k - \sum_{k=0}^{p-2} k a_k x + \ldots$ (we only need the first two terms since a and b are defined modulo λ^2). So c is the solution to

$$\left(\sum_{k=0}^{p-2} a_k \right) x \equiv -\sum_{k=0}^{p-2} k a_k \pmod{p} \tag{1.5}$$

Since $a_{p-1} = 0$, this is equivalent to

$$\left(\sum_{k=0}^{p-1} a_k \right) x \equiv -\sum_{k=0}^{p-1} k a_k \pmod{p} \tag{1.6}$$

Now

$$\begin{aligned} k_0 + t &\equiv p - k_0 \pmod{p} \\ \Rightarrow k_0 + t &\equiv -(k_0 + t) + t \pmod{p} \\ \Rightarrow (k_0 + t) &\equiv 2^{-1} t \pmod{p} \\ \Rightarrow b_{k_0+t} &= a_{k_0+t} = a_{2^{-1}t}. \end{aligned}$$

Finally, note that $i a_i + (t-i) a_{t-i} = t b_l$ since $a_i = a_{t-i} = b_l$ (with $1 \leq l \leq \frac{p-1}{2}$) by (1.4).

(1.6) then becomes $\left(b_{k_0+t} + 2 \sum_{k=1}^{\frac{p-1}{2}} b_k \right) x \equiv -\left((2^{-1}t \bmod p) b_{k_0} + \sum_{k=1}^{\frac{p-2}{2}} t b_k \right) \pmod{p}$. It is clear that $c \equiv -2^{-1}t \pmod{p}$ is a solution to this congruence. By uniqueness of this solution, we see that

$$u \text{ is primary} \Leftrightarrow t \equiv 0 \pmod{p} \Leftrightarrow u \text{ is real.}$$

2. An Approach to Pell's Equation using Cyclotomy

Pell's Equation is

$$x^2 - dy^2 = 1, \quad x, y \in \mathbb{Z}$$

in x and y, where d is a positive integer. A nonpositive d trivially yields the single solution $(1, 0)$, and we can consider d to be square-free, since any square factor of d can be incorporated into y.

We will primarily consider the case where d is an odd prime p. The equation can then be solved using cyclotomy and quadratic residues. A partial solution was found by Dirichlet [3] using this method, building upon the work of Gauss. In this section, we build upon Dirichlet's work, explicitly writing the solution and using the machinery of Galois theory to streamline the approach. Again, we let p be an odd prime, define $p^* := (-1)^{\frac{p-1}{2}}p$, and let $i = \sqrt{-1}$. We start by introducing an important lemma.

Lemma 2. *Define* $\begin{cases} q_1(x) := 2 \prod\limits_{\substack{1\le k<p \\ \left(\frac{k}{p}\right)=1}} (x-\zeta^k) \\ q_{-1}(x) := 2 \prod\limits_{\substack{1\le k<p \\ \left(\frac{k}{p}\right)=-1}} (x-\zeta^k) \end{cases}$.

Then $\begin{cases} q_1(x) = f(x) + \sqrt{p^*}g(x) \\ q_{-1}(x) = f(x) - \sqrt{p^*}g(x) \end{cases}$ *where $f(x)$ and $g(x)$ are polynomials in $\mathbb{Z}[x]$.*

Proof. Note that the product of the 2 above polynomials (on the left-hand side) is $4 \prod\limits_{1\le k<p} (x-\zeta^k) = 4m_p(x)$, which has integer coefficients. The product is therefore fixed by any Galois automorphism in $Gal(K:\mathbb{Q})$. Now taking $\theta := \zeta^{\frac{p^2-1}{8}} \prod\limits_{k=1}^{\frac{p-1}{2}} (1-\zeta^k)^2$, we see that $\theta^2 = p^*$ since $(-1)^{\frac{p^2-1}{8}} \equiv \left(\frac{2}{p}\right) \pmod 2$, and θ is trivially in $\mathcal{O}_K$. So $\sqrt{p^*} \in \mathcal{O}_K$. Now an automorphism σ in the Galois group fixes p^* if and only if σ is a square. But this is if and only if σ fixes all (and only) the ζ^k such that k is a quadratic residue modulo p. So $\prod\limits_{\substack{1\le k<p \\ \left(\frac{k}{p}\right)=1}} (x-\zeta^k)$ is in $L[x]$ where $L := \mathbb{Q}(\sqrt{p^*})$. All the coefficients in $L[x]$ are of the form $a + b\sqrt{p^*}$ where a and b are both rational and $\frac{1}{2}\cdot$ an algebraic integer (allowing for the fact that $p^* \equiv 1 \pmod 4$). The coefficients of $2 \prod\limits_{\substack{1\le k<p \\ \left(\frac{k}{p}\right)=1}} (x-\zeta^k)$ are therefore rational algebraic integers and thus in $\mathbb{Z}$. We can now expand $q_1(x)$ and rewrite it as $q_1(x) = f(x) + \sqrt{p^*}g(x)$ where $f(x), g(x)$ are polynomials in $\mathbb{Z}[x]$.

A similar argument shows that $q_{-1}(x) \in L[x]$. Now let τ be the Galois automorphism in $Gal(K:Q)$ defined by $\tau(\sqrt{p^*}) = -\sqrt{p^*}$ (noting that $K:L:\mathbb{Q}$ is a tower of fields). Then by the above, and since τ^2 must fix $q_1(x)$, we must have that $\tau(\zeta^k) = \zeta^l$ where $\left(\frac{k}{p}\right)\left(\frac{l}{p}\right) = -1$. So since τ is a Galois automorphism over K, we must have $\tau(q_1(x)) = q_{-1}(x)$. This yields that $q_{-1}(x) = f(x) - \sqrt{p^*}g(x)$. □

To prove this result, Dirichlet simply referenced Article 357 of Gauss's great *Disquisitiones Arithmeticae*, in which the argument is quite long. Thankfully, using the above application of Galois Theory that was probably unavailable to Gauss and Dirichlet, the lemma is quickly proven.

We now go back to Pell's Equation. Setting $d = p$, Pell's Equation then becomes

$$x^2 - py^2 = 1 \tag{2.1}$$

By Lemma 2,

$$4m_p(x) = q_1(x)q_{-1}(x) = f(x)^2 - (p^*)g(x)^2$$

And so, replacing x by 1, we get

$$4p = x_1^2 - p^*y_1^2 \text{ where } x_1 := f(1) \text{ and } y_1 := g(1) \tag{2.2}$$

Since $f(x)$ and $g(x)$ are in $\mathbb{Z}[x]$ and x_1, y_1 in $\mathbb{Z}$, we can see that Lemma 2 relates to Pell's Equation insofar as it gives us a pair (x_1, y_1) that verifies an equation very similar to (2.1).

Since p is prime, the above equation shows that p must divide x_1. So defining $\xi_1 := \frac{x_1}{p}$, we can rewrite equation (2.2) as $4p = p^2\xi_1^2 - p^*y_1^2$, and so, dividing by p,

$$p\xi_1^2 - (-1)^{\frac{p-1}{2}}y_1^2 = 4 \tag{2.3}$$

We now analyze $q_1(x)$ and $q_{-1}(x)$ to obtain some insight as to the values x_1 and y_1. $x^2 \equiv (p-x)^2 \pmod p$, so all quadratic residues are in $\{x^2 \pmod p : 1 \le x \le \frac{p-1}{2}\}$. We can therefore reorder the terms in $q_1(x)$ and write it as $2\prod\limits_{k=1}^{\frac{p-1}{2}}(x-\zeta^{k^2})$, and so $q_1(1) = 2\prod\limits_{k=1}^{\frac{p-1}{2}}(1-\zeta^{k^2})$, which can much more easily be computed than $2 \prod\limits_{\substack{1\le k<p \\ \left(\frac{k}{p}\right)=1}} (1-\zeta^k)$. Since $4p = q_1(1)q_{-1}(1)$, we can also easily compute $q_{-1}(1)$.

Now, the value of p^* depends on the value of p modulo 4 so we will consider the two cases separately for simplicity.

Case 1: $p \equiv 1 \pmod 4$.

Then (2.3) becomes $p\xi_1^2 - y_1^2 = 4$ (or, to emphasize the similarity to Pell's Equation, $y_1^2 - p\xi_1^2 = -4$).

We then have two subcases.

Case 1.1: If $p \equiv 1 \pmod 8$, then $y_1^2 - \xi_1^2 \equiv 4 \pmod 8$. Trivially y_1 and ξ_1 must either be both odd or both even. But $1^2 \equiv 3^2 \equiv 5^2 \equiv 7^2 \equiv 1 \pmod 8$, so if y_1 and ξ_1 were both odd we would have $y_1^2 - \xi_1^2 \equiv 0 \pmod 8$, yielding a contradiction. It follows that y_1 and ξ_1 are both even, and we can thus define the new integers $y_2 := \frac{y_1}{2}$ and $\xi_2 := \frac{\xi_1}{2}$. Then $y_2 - p\xi_2^2 = -1$. We can now use the fact that $(\sqrt{p})^2$ is an integer to get rid of the minus sign in front of 1. The equation $y_2^2 - p\xi_2^2 = -1$ yields $(y_2 - \sqrt{p}\xi_2)(y_2 + \sqrt{p}\xi_2) = -1$, and so $(y_2 - \sqrt{p}\xi_2)^2(y_2 + \sqrt{p}\xi_2)^2 = 1$. But $(y_2 \pm \sqrt{p}\xi_2)^2 = a \pm b\sqrt{p}$ for yet another pair of integers (a, b). Taking $(x, y) = (a, b)$, we have solved (2.1). Summarizing, we get a solution from

$$(a, b) \quad = \quad \left(\frac{1}{4}(g(1)^2 + \frac{f(1)^2}{p}) \;,\; \frac{f(1)g(1)}{2p}\right)$$

where we can directly compute $f(1)$ and $g(1)$ from $q_1(1)$ and $q_{-1}(1)$.

Case 1.2: If $p \equiv 5 \pmod 8$, then $y_1^2 + 3\xi_1^2 \equiv 4 \pmod 8$. Given that the only quadratic residues modulo 8 are $0, 1, 4$, we must have (y_1^2, ξ_1^2) congruent to one of $(1, 1), (0, 4)$ or $(4, 0)$ modulo 8.

We now use the fact that $8^2 = 2^{2 \cdot 3} = 4^3$ and consider $(y_1 + \sqrt{p}\xi_1)^3 = y_2 + \sqrt{p}\xi_2$ where y_2 and ξ_2 are the integers defined as $(y_1^3 + 3p\xi_1^2 y_1)$ and $(p\xi_1^3 + 3y_1^2\xi_1)$ respectively. We then see that $y_2^2 - p\xi_2^2 = (y_1^2 - p\xi_1^2)^3 = -4^3$.

But $y_2 = y_1(y_1^2 + 3p\xi_1^2)$, which is congruent to $y_1(y_1^2 - \xi_1^2)$ modulo 8. If $(y_1^2, \xi_1^2) \equiv (1, 1) \pmod 8$ then $y_2 \equiv 0 \pmod 8$. If $(y_1^2, \xi_1^2) \equiv (0, 4)$ or $(4, 0) \pmod 8$ then $y_2 \equiv 4 \cdot 4,\ 0 \cdot 4$ or $\pm 2 \cdot 4$, all congruent to 0 modulo 8. So in any case $y_2 \equiv 0 \pmod 8$.

Likewise $\xi_2 = \xi_1(p\xi_1^2 + 3y_1^2)$, and as above simple arithmetic modulo 8 shows that $\xi_2 \equiv 0 \pmod 8$.

So 8 divides both y_2 and ξ_2. So writing $y_3 := \frac{y_2}{8}$ and $\xi_3 := \frac{\xi_2}{8}$, we get that y_3 and ξ_3 are integers with $(y_3^2 - p\xi_3^2) = \frac{-4^3}{8^2} = -1$. As in Case 1.1, writing $(y_3 \pm \sqrt{p}\xi_3)^2 = a \pm b\sqrt{p}$ where a and b are integers gives a solution $(x, y) = (a, b)$ of (2.1). Summarizing, we get a solution from

$$(a, b) = \begin{pmatrix} \frac{1}{64}((g(1)^3 + \frac{3f(1)^2g(1)}{p})^2 + p(\frac{f(1)^3}{p^2} + 3\frac{g(1)^2f(1)}{p})^2) \;, \\ \frac{1}{32}(g(1)^3 + 3\frac{f(1)^2g(1)}{p})(\frac{f(1)^3}{p^2} + 3\frac{g(1)^2f(1)}{p}) \end{pmatrix}$$

Case 2: $p \equiv 3 \pmod 4$.

Write $l := \frac{p-1}{2}$. Trivially l is an odd integer. We see that $f(x)$ can be written as $\frac{1}{2}(q_1(x) + q_{-1}(x)) = \prod_{\substack{1 \le k < p \\ (\frac{k}{p}) = 1}} (x - \zeta^k) + \prod_{\substack{1 \le k < p \\ (\frac{k}{p}) = -1}} (x - \zeta^k)$, which is of degree l. We shall find a relation amongst the coefficients of f by comparing $f(\zeta)$ and $f(\overline{\zeta})$, which is equal to $\overline{f(\zeta)}$ since $f(x) \in \mathbb{Z}[x]$. Trivially $(\frac{1}{p}) = 1$, so $\prod_{\substack{1 \le k < p \\ (\frac{k}{p}) = 1}} (\zeta - \zeta^k) = 0$ and so $f(\zeta) = \prod_{\substack{1 \le k < p \\ (\frac{k}{p}) = -1}} (\zeta - \zeta^k)$. Also note that $(\frac{-1}{p}) = (-1)^{\frac{p-1}{2}} = -1$, and so $(\frac{k}{p}) = -(\frac{-k}{p})$ for all $1 \le k \le p-1$. So $f(\zeta) = \prod_{\substack{1 \le k < p \\ (\frac{k}{p}) = 1}} (\zeta - \zeta^{-k})$.

By the same line of reasoning, $f(\overline{\zeta}) = f(\zeta^{-1}) = \prod_{\substack{1 \le k < p \\ (\frac{k}{p}) = 1}} (\zeta^{-1} - \zeta^k)$. So

$$\frac{f(\zeta)}{f(\zeta^{-1})} \quad = \quad \prod_{\substack{1 \le k < p \\ (\frac{k}{p}) = 1}} \frac{(\zeta - \zeta^{-k})}{(\zeta^{-1} - \zeta^k)} = (-1)^l \prod_{\substack{1 \le k < p \\ (\frac{k}{p}) = 1}} \zeta^{1-k}$$

since there are precisely l quadratic residues modulo p

$$= \quad -\zeta^l \prod_{\substack{1 \le k < p \\ (\frac{k}{p}) = 1}} \zeta^{-k}$$

$$= \quad -\zeta^l$$

since $\sum_{\substack{1 \le k < p \\ (\frac{k}{p}) = 1}} k = p\frac{p-1}{2} + 0$ since the Legendre symbol is a quadratic character modulo p and since $\left(\frac{0}{p}\right) = 0$.

So $f(\zeta) = -\zeta^l f(\zeta^{-1})$. So writing $f(x) = a_l x^l + a_{l-1}x^{l-1} + ... + a_1 x + a_0$, this yields $a_l\zeta^l + a_{l-1}\zeta^{l-1} + ... + a_1\zeta + a_0 = -a_0\zeta^l - a_1\zeta^{l-1} - ... - a_{l-1}\zeta - a_l$, i.e.

$$\sum_{k=0}^{l} a_k\zeta^k = \sum_{k=0}^{l}(-a_k)\zeta^{l-k} \tag{2.4}$$

But $\{1, \zeta, ..., \zeta^l\}$ is a subset of the $\mathbb{Z}$-basis $\{1, \zeta, ..., \zeta^{p-2}\}$ of $\mathcal{O}_K$ for $p \geq 3$, and so is $\mathbb{Z}$-linearly independent. So $a_{l-k} = -a_l$ for all $0 \leq k \leq l$. We can therefore rewrite $f(x)$ as $2(x^l - 1) + b_1 x(x^{l-2} - 1) + b_2 x^2(x^{l-4} - 1) + ... + b_{\frac{l-1}{2}} x^{\frac{l-1}{2}}(x - 1) = \sum_{k=0}^{\frac{l-1}{2}} b_k x^k(x^{l-2k} - 1)$ for some integers b_k for all $0 \leq k \leq \frac{l-1}{2}$.

Replacing x by $i := \sqrt{-1}$, we see that $x^k(x^{l-2k} - 1)$ depends on whether p is congruent to 3 or 7 modulo 8.

Let $p \equiv 3 \pmod 8$. Then $l \equiv 1 \pmod 4$ and simple calculation yields

$$i^k(i^{l-2k} - 1) = \begin{cases} 1 - i & \text{if } k \equiv 1, 2 \pmod 4 \\ -(1 - i) & \text{if } k \equiv 0, 3 \pmod 4 \end{cases}$$

If $p \equiv 7 \pmod 8$ then $l \equiv 3 \pmod 4$, and the same type of calculation yields

$$i^k(i^{l-2k} - 1) = \begin{cases} 1 + i & \text{if } k \equiv 3, 2 \pmod 4 \\ -(1 + i) & \text{if } k \equiv 0, 1 \pmod 4 \end{cases}$$

Setting $i^* := \begin{cases} -i & \text{if } p \equiv 3 \pmod 8 \\ +i & \text{if } p \equiv 7 \pmod 8 \end{cases}$, we see that $f(i) = \sum_{k=0}^{\frac{l-1}{2}} \pm b_k(1 + i^*) = y_2(1 + i^*)$ where y_2 is in $\mathbb{Z}$.

Now,

$$\begin{aligned} g(x) &= \frac{1}{2\sqrt{p^*}}(q_1(x) - q_{-1}(x)) \\ &= \frac{1}{\sqrt{p^*}}\left(\prod_{\substack{1\leq k<p \\ (\frac{k}{p})=1}}(x - \zeta^k) - \prod_{\substack{1\leq k<p \\ (\frac{k}{p})=-1}}(x - \zeta^k)\right) \end{aligned}$$

And so

$$\begin{aligned} g(\zeta) &= -\frac{1}{\sqrt{p^*}}\left(\prod_{\substack{1\leq k<p \\ (\frac{k}{p})=1}}(\zeta - \zeta^{-k})\right) \\ \text{and } g(\zeta^{-1}) &= \frac{1}{\sqrt{p^*}}\left(\prod_{\substack{1\leq k<p \\ (\frac{k}{p})=1}}(\zeta^{-1} - \zeta^k)\right) \end{aligned}$$

A similar line of reasoning as for $f(x)$ gives us that $g(\zeta) = +\zeta^l g(\zeta^{-1})$. Following the same steps as for $f(x)$, we find that, writing $g(x)$ as $\frac{1}{\sqrt{p^*}}\sum_{k=0}^{l} a_k x^k$, we get $a_{l-k} = +a_l$ for all $0 \leq k \leq l$. We can therefore similarly rewrite $g(x)$ as $\sum_{k=0}^{\frac{l-1}{2}} b_k x^k(x^{l-2k} + 1)$, where all the b_k are integers (remembering that $g(x) \in \mathbb{Z}[x]$ by Lemma 2). A similar argument shows that $g(i) = \sum_{k=0}^{\frac{l-1}{2}} \pm b_k(1 - i^*) = \xi_2(1 - i^*)$ where ξ_2 is in $\mathbb{Z}$.

Now $l \equiv 3 \pmod 4$, so $q_1(i)q_{-1}(i) = 4m_p(i) = 4(1 + i + ... + i^l) = 4 \cdot ((1 + i - 1 - i) + (1 + i - 1 - i) + ... + (1 + i - 1)) = 4i$.

So $f(i)^2 - p^*g(i)^2 = f(i)^2 + pg(i)^2 = 4i$, and so $y_2^2(1+i^*)^2 + p\xi_2(1-i^*)^2 = 2y_2^2 i^* - 2p\xi_2^2 i^* = 4i$. Dividing by $2i^* = \pm 2i$, this yields

$$\begin{aligned} y_2^2 - p\xi_2^2 &= \pm 2 \\ \text{i.e. } (y_2 + \sqrt{p}\xi_2)^2(y_2 - \sqrt{p}\xi_2)^2 &= 4 \end{aligned}$$

Now y_2 and ξ_2 are odd, otherwise we would have $y_2^2 - p\xi_2^2 \equiv y_2^2 + \xi_2^2 \equiv 0 \not\equiv \pm 2 \pmod 4$. So the coefficients of $(y_2 + \sqrt{p}\xi_2)^2 = (y_2^2 + p\xi_2^2) + 2y_2\xi_2\sqrt{p}$ are even. We can thus define the integers $a := \frac{(y_2^2 + p\xi_2^2)}{2}$ and $b := y_2\xi_2$ and get

$$a^2 - pb^2 = \frac{(y_2 + \sqrt{p}\xi_2)^2(y_2 - \sqrt{p}\xi_2)^2}{2 \cdot 2} = \frac{4}{4} = 1$$

This solves the equation, where

$$(a,b) \;=\; \left(\frac{i^*}{4}(pg(i)^2 - f(i)^2)\ ,\ \frac{1}{2}g(i)f(i)\right)$$

where we can directly compute $f(i)$ and $g(i)$ from $q_1(i)$ and $q_{-1}(i)$.

To apply this method to the general case of Pell's Equation (where d is square-free but not necessarily prime), we factor d as $\prod_{k=1}^{r} p_k$ where the p_k's are rational primes. So it suffices to study the case where $d = pq$ for primes p and q and deduce the general case by induction. We will not describe said case in depth here since this paper mainly focuses on prime cyclotomic fields, but we remark that taking $\mathbb{Q}(\zeta_{pq})$, $m_{pq}(x) = m_p(x)m_q(x)\frac{(x^{pq}-1)/(x-1)}{((x^p-1)/(x-1))\cdot((x^q-1)/(x-q))} = \frac{(x^{pq}-1)(x-1)}{(x^p-1)(x^q-1)}$ which can be shown to be irreducible by a similar method as the simple proof for showing that $\sum_{k=0}^{p-1} x^k$ is the minimal polynomial of ζ_p in $\mathbb{Z}[x]$. Following the same reasoning as in the case where $d = p$, we can write $4m_{pq}(x) = f(x)^2 \pm pqg(x)^2$ where $f(x), g(x) \in \mathbb{Z}[x]$. The rest of the problem is solved in a similar fashion as well.

Using some interesting approximation methods and quadratic number fields, Ireland & Rosen ([5, Proposition 17.5.2, pp. 277-8]) show that $x^2 - dy^2 = 1$ has *infinitely many solutions* for any square-free integer d (including $d = 2$), and that every solution has the form $\pm(x_n, y_n)$ (with $n \in \mathbb{Z}$) where $x_n + \sqrt{d}y_n = (x_1 + \sqrt{d}y_1)^n$ for some solution (x_1, y_1) commonly known as the *fundamental* solution.

Acknowledgment *Many thanks to Professor Dan Segal, All-Souls College, Oxford, for his advice.*

References

[1] Borevich, Z. I., and Shafarevich I. R., Number Theory, Academic Press, New York, 1966.

[2] C. S. Dalawat, Primary units in cyclotomic fields, *Annales des sciences mathématiques du Québec* to appear, 2011.

[3] G. L. Dirichlet, Sur la manière de résoudre l'équation $t^2 - pu^2 = 1$ au moyen des fonctions circulaires, *Journal für die reine und angewandte Mathematik* 17, pp. 286-290, 1837.

[4] V. Flynn, *Algebraic Number Theory Lecture Notes*, University of Oxford, Oxford Mathematical Institute, Oxford, UK, 2011. Lecture Notes.

[5] K. Ireland and M. Rosen, A Classical Introduction to Modern Number Theory, Springer-Verlag, New York, 1990.

[6] S. Lang, Algebraic Number Theory, Springer-Verlag, New York, 1986.

[7] I. Stewart and D. Tall, Algebraic Number Theory and Fermat's Last Theorem, A K Peters, Massachusetts, 2002.

[8] L. C. Washington, Introduction to Cyclotomic Fields, Springer-Verlag, New York, 1982.