REGULAR PRIMES AND FERMAT'S LAST THEOREM

STEVEN SIVEK

1. Introduction

Fermat's Last Theorem is the well-known result that the equation $x^n + y^n = z^n$ has no nontrivial integer solutions for $n \geq 3$. Fermat gave only a proof in the case $n = 4$, and the theorem was proven for other specific values of n by mathematicians including Euler, Gauss, Dirichlet, and Lamé.

Given Fermat's solution in the case $n = 4$, it is easy to show that it suffices to prove the theorem in the case where x, y, z are pairwise relatively prime and n is an odd prime, and also that one may instead address the equation $x^p + y^p + z^p = 0$, since this is equivalent to $x^p + y^p = (-z)^p$. Early attempts often split the theorem into two cases: in the first, x, y, and z are not multiples of p, and in the second case, p divides one of them. The first proof of Fermat's theorem in any generality came from Sophie Germain, who proved the first case of the theorem for any p such that $2p + 1$ is also prime.

The first significant advance toward a proof in both cases was provided by Kummer, who proved the theorem to be true whenever p is a "regular" prime and invented much of the theory of ideals along the way. His proof made use of cyclotomic fields and the elegant identity

$$x^n + y^n = \prod_{i=0}^{n-1} (x + \zeta^i y), \tag{1.1}$$

where $\zeta = e^{2\pi i/n}$. Though Fermat's Last Theorem was eventually proven by Wiles et. al. in 1995 using a very different approach, Kummer's work is still significant because it motivated much of the earliest work in algebraic number theory.

The primes for which Kummer proved Fermat's Last Theorem are defined as follows:

Definition 1.1. Let p be prime, and let $K = \mathbb{Q}(e^{2\pi i/p})$ be the field obtained by adjoining a primitive pth root of unity to $\mathbb{Q}$. Then p is said to be *regular* if the order of the class group of K is not a multiple of p.

Kummer's result established the theorem for all primes less than 100 except the irregular primes 37, 59, and 67. It is still unknown whether or not there are infinitely many regular primes, though Jensen proved in 1915 that infinitely many primes (in fact, infinitely many primes congruent to 3 mod 4) are irregular. However, the regular primes are believed to have density $e^{-1/2}$ (just over 60 percent) in the set of all primes. This paper will give both cases of Kummer's proof and then develop a criterion for regularity which is much easier to use in practice.

The proof will proceed as follows: We will first prove case I by using unique factorization of ideals to show that $\langle x + \zeta y \rangle$ is a pth power of some ideal, which

must be principal by the regularity of p. This will enable us to write $x+\zeta y = u\alpha^p$ for some unit u, and upon writing $\alpha^p \equiv a \pmod{p}$ for some rational integer a we will derive a congruence relating x, y, and ζ modulo p which cannot have any solutions. In case II, we will assume that p divides z, write $z = p^e z_0$, and show that the more general $x^p + y^p = (1-\zeta)^{np} z_0^p$ cannot have any nontrivial solutions when $n \geq 1$. We will do this by taking a solution which minimizes n and showing that for such a solution, $\langle x+\zeta^i y\rangle$ is always two particular ideals times a pth power of an ideal. Then the quotient of two of these ideals is a pth power of a principal ideal, and so using two generators of these principal ideals and invoking a powerful lemma on units we will construct a solution with a strictly smaller value of n, causing a contradiction.

2. Cyclotomic fields

In this section we will gather some useful facts about cyclotomic fields, that is, fields of the form $\mathbb{Q}(\zeta_n)$ for some nth root of unity $\zeta_n = e^{2\pi i/n}$. We note without proof that the cyclotomic field $\mathbb{Q}(\zeta_n)$ has degree $\phi(n)$ over $\mathbb{Q}$, where ϕ denotes the Euler totient function. In fact, it is easy to show that $\mathbb{Q}(\zeta_n)$ is a Galois extension of $\mathbb{Q}$, and $\mathrm{Gal}(\mathbb{Q}(\zeta_n)/\mathbb{Q})$ is isomorphic to $(\mathbb{Z}/n\mathbb{Z})^*$, the group of units modulo n.

Now let p denote an odd prime, and let $\zeta = \zeta_p$ and $K = \mathbb{Q}(\zeta)$.

Proposition 2.1. *The roots of unity in K are precisely those η such that $\eta^{2p} = 1$.*

Proof. Note that $\zeta_{2p} = -\zeta_p^{(p+1)/2}$, so that ζ_{2p} and hence all of the $2p$th roots of unity are contained in K; in fact, since $\zeta_p = \zeta_{2p}^2$ we see that $K = \mathbb{Q}(\zeta_{2p})$. For the converse, we note that the torsion subgroup of the group of units U_K is known to be cyclic, so it is generated by some ζ_n. Since ζ_{2p} is in this cyclic group, we must have $2p \mid n$; but also $\mathbb{Q} \subset \mathbb{Q}(\zeta_n) \subset K$, so $[\mathbb{Q}(\zeta_n) : \mathbb{Q}]$ divides $[K : \mathbb{Q}]$, or $\phi(n) \mid \phi(2p)$. It is an easy exercise to show that this is only possible when $n = 2p$. □

Lemma 2.2. *If $1 \leq a \leq p-1$, then $1-\zeta$ and $1-\zeta^a$ are associates in the ring of integers $\mathcal{O}_K$.*

Proof. Let b be the inverse of a modulo p. Then $1-\zeta^a = (1-\zeta)(1+\zeta+\zeta^2+\cdots+\zeta^{a-1})$, and $1-\zeta = 1-\zeta^{ab} = (1-\zeta^a)(1+\zeta^a+\zeta^{2a}+\cdots+\zeta^{(b-1)a})$. All factors are algebraic integers, since $\mathbb{Z}[\zeta] \subset \mathcal{O}_K$, so $1-\zeta$ and $1-\zeta^a$ are multiples of each other in $\mathcal{O}_K$ and hence the two differ by a unit factor. □

Corollary 2.3. *The ideal $p\mathcal{O}_K$ factors into primes as $p\mathcal{O}_K = \langle 1-\zeta\rangle^{p-1}$.*

Proof. We know that $\prod_{i=0}^{p-1}(x-\zeta^i) = \frac{x^p-1}{x-1} = x^{p-1}+\cdots+x+1$, so setting $x=1$ yields $\prod_{i=0}^{p-1}(1-\zeta^i) = p$. If we pass to ideals, then $p\mathcal{O}_K = \prod_{i=0}^{p-1}\langle 1-\zeta^i\rangle$, and each ideal in the product is equal to $\langle 1-\zeta\rangle$ since $1-\zeta$ and $1-\zeta^i$ are associates. It now remains to be shown that $\langle 1-\zeta\rangle$ is prime.

Assume $p\mathcal{O}_K$ factors into primes as $\prod_{i=1}^{g}\mathfrak{p}_i^e$, where each $\mathfrak{p}_i$ has residue class degree f and all $\mathfrak{p}_i$ share the same e and f since K is Galois. Since $p\mathcal{O}_K$ is a perfect $(p-1)$st power, we have $p-1 \mid e$. Then we know that $p-1 = [K:\mathbb{Q}] = efg$, which is only possible if $g=1$ and $e = p-1$, so $p\mathcal{O}_K = \mathfrak{p}_1^{p-1}$ for a prime $\mathfrak{p}_1$. Unique factorization of ideals tells us that $\mathfrak{p}_1 = \langle 1-\zeta\rangle$, as desired. □

Proposition 2.4. *The ring of integers $\mathcal{O}_K$ is equal to $\mathbb{Z}[\zeta]$.*

Proof. Corollary 2.3 gives us $p\mathcal{O}_K = \langle(1-\zeta)^{p-1}\rangle$. We expand using the binomial theorem: $(1-\zeta)^{p-1} = \sum_{k=0}^{p-1}\binom{p-1}{k}(-\zeta)^k = \sum_{k=0}^{p-2}\left(\binom{p-1}{k}-(-1)^k\right)(-\zeta)^k$, since $\zeta^{p-1} = -\zeta^{p-2}-\cdots-\zeta-1$. But $\binom{p-1}{k} \equiv (-1)^k \pmod p$ for $0 \le k \le p-1$ by an easy induction: it is true for $k=0$, and for $1 \le k \le p-1$ we note that $\binom{p-1}{k} = \binom{p}{k}-\binom{p-1}{k-1} \equiv 0-(-1)^{k-1} \pmod p$. Hence we can write $\binom{p-1}{k}-(-1)^k = pa_k$ for some integers a_k, and then we have $(1-\zeta)^{p-1} = \sum_{k=0}^{p-2} pa_k(-\zeta)^k = py$, where $y = \sum a_k(-\zeta)^k$ is clearly a member of $\mathbb{Z}[\zeta]$.

We now have $p\mathcal{O}_K = \langle py\rangle = \langle p\rangle\langle y\rangle$, so by unique factorization of ideals we may cancel $\langle p\rangle$ on both sides to get $\mathcal{O}_K = \langle y\rangle \subset \mathbb{Z}[\zeta]$. But clearly $\mathbb{Z}[\zeta] \subset \mathcal{O}_K$, as remarked earlier, so we conclude that $\mathcal{O}_K = \mathbb{Z}[\zeta]$ as desired. □

We remark that this proof is only valid for p prime, but in general it is still true that the cyclotomic field $\mathbb{Q}(\zeta_n)$ has ring of integers $\mathbb{Z}[\zeta_n]$ even when n is not prime.

Proposition 2.5. *Every unit $\eta \in \mathcal{O}_K$ is of the form $r\zeta^n$ for some real unit $r \in \mathcal{O}_K$.*

Proof. First note that $\overline{\eta} \in \mathcal{O}_K$ is also a unit, and so if we let $\rho = \frac{\eta}{\overline{\eta}}$ then $\rho \in \mathcal{O}_K$. Given any automorphism $\sigma \in \mathrm{Gal}(K/\mathbb{Q})$, we have $|\sigma(\rho)| = \left|\frac{\sigma(\eta)}{\sigma(\overline{\eta})}\right| = \frac{|\sigma(\eta)|}{\left|\overline{\sigma(\eta)}\right|} = 1$. If $f(x)$ is the minimal polynomial of ρ, and f has splitting field L, then $L \subset K$, so $|\tau(\rho)| = 1$ for every $\tau \in \mathrm{Gal}(L/\mathbb{Q})$. Thus the conjugates of ρ all lie on the unit circle. The same is clearly true for every ρ^e, $i \ge 1$, since we may replace η with the unit η^e and proceed identically.

Let ρ^e have minimal polynomial $g(x)$ with roots $\rho_1, \rho_2, \cdots, \rho_k$, all have which must have modulus 1. Then $g(x) = x^k + a_1x^{k-1} + \cdots + a_{k-1}x + a_k$, where each a_i is the sum of all i-wise products of the roots ρ_j; since $|\rho_j| = 1$ for all j, each i-wise product has modulus 1, and since there are $\binom{k}{i}$ such products the triangle inequality yields $|a_i| \le \binom{k}{i}$. But $\deg(g) \le p-1$, since g splits in a subfield of K, so g can only be one of finitely many polynomials. Since every ρ^e is a root of one of these polynomials, each of which has at most $p-1$ roots, we must have $\rho^a = \rho^b$ for some $a \ne b$, and so ρ is necessarily a root of unity. Then by proposition 2.1 we have $\rho = \pm\zeta^m$ for some m.

We now have $\eta = \pm\overline{\eta}\zeta^m$. Suppose that the sign is negative, and since $\eta \in \mathbb{Z}[\zeta]$ let $\eta = \sum c_i\zeta^i$ for some integers c_i. Then since $1-\zeta^i \in \langle 1-\zeta\rangle$ for all i, we have $\eta \equiv \sum c_i \equiv \overline{\eta} \pmod{1-\zeta}$, hence $\eta = -\overline{\eta}\zeta^k \equiv -\eta \pmod{1-\zeta}$. But then $2\eta \in \langle 1-\zeta\rangle$, hence $2 \in \langle 1-\zeta\rangle$. This is impossible, since $p \in \langle 1-\zeta\rangle$ is relatively prime to 2 and $\langle 1-\zeta\rangle$ is not the unit ideal, so $\eta = \overline{\eta}\zeta^m$. If m is odd, then replace m with $m+p$ so that $\eta = \overline{\eta}\zeta^{2n}$ for some integer n, and define $r = \eta\zeta^{-n}$; then r is a unit, and r is real since $\overline{r} = \overline{\eta}\zeta^n = r$, so $\eta = r\zeta^n$ and we are done. □

For the sake of convenience, throughout the rest of this paper we will let $\lambda = 1-\zeta$ and let $\mathfrak{l} = \langle\lambda\rangle$. Therefore $\mathfrak{l}$ is prime, $p\mathcal{O}_K = \mathfrak{l}^{p-1}$, and since $p \in \mathfrak{l}$ we note that $\mathfrak{l} \cap \mathbb{Z} = p\mathbb{Z}$.

3. Case 1: $p \nmid xyz$

Suppose that for some regular prime p, we have relatively prime integers x, y, and z which are not multiples of p such that $x^p + y^p + z^p = 0$, or $x^p + y^p = (-z)^p$. If $p = 3$, we note that $(3k \pm 1)^3 \equiv \pm 1 \pmod 9$, and since each of x, y, and z has the form $3k \pm 1$ we have $\pm 1 \pm 1 \pm 1 \equiv 0 \pmod 9$, which is impossible; hence we

need only consider $p \geq 5$. Using the factorization (1.1) and passing to ideals in the field $K = \mathbb{Q}(\zeta)$, where $\zeta = e^{2\pi i/p}$, we have

$$\langle z \rangle^p \quad = \quad \prod_{i=0}^{p-1} \langle x + \zeta^i y \rangle.$$

Proposition 3.1. *The ideals $\langle x+\zeta^i y\rangle$ are all relatively prime for $i = 0, 1, \ldots, p-1$.*

Proof. Suppose that some prime $\mathfrak{p}$ divides both $\langle x + \zeta^i y\rangle$ and $\langle x + \zeta^j y\rangle$, where $0 \leq i < j \leq p-1$. Then $\mathfrak{p}$ contains both generators $x + \zeta^i y$ and $x + \zeta^j y$, so it also contains $(x + \zeta^i y) - \zeta^{i-j}(x + \zeta^j y) = (1 - \zeta^{i-j})x$ and $\zeta^{-j}((x + \zeta^j y) - (x + \zeta^i y)) = (1 - \zeta^{i-j})y$. Since x and y are relatively prime, there exist integers m and n such that $mx + ny = 1$, and so $\mathfrak{p}$ must contain $m(1 - \zeta^{i-j})x + n(1 - \zeta^{i-j})y = 1 - \zeta^{i-j}$. It follows that $\mathfrak{l} = \langle 1 - \zeta^{i-j}\rangle \subset \mathfrak{p}$; but prime ideals in $\mathcal{O}_K$ are maximal, so $\mathfrak{p} = \mathfrak{l}$.

Since $\mathfrak{l} \mid \langle x+\zeta^i y\rangle$, we have $\mathfrak{l} \mid \langle z\rangle$, and so $z \in \langle z\rangle \cap \mathbb{Z} \subset \mathfrak{l} \cap \mathbb{Z}$. But $\mathfrak{l} \cap \mathbb{Z} = p\mathbb{Z}$, and this contradicts the assumption that z is not a multiple of p. Therefore the ideals $\langle x + \zeta^i y\rangle$ and $\langle x + \zeta^j y\rangle$ must be relatively prime. □

Since $\langle z\rangle^p$ is a perfect pth power, we see that each of the ideals $\langle x + \zeta^i y\rangle$ must also be a pth power. In particular, we may write $\langle x + \zeta y\rangle = I^p$ for some ideal I. But if I^p is principal, then I has order dividing p in the class group $\mathrm{Cl}(K)$. Since p is regular, p does not divide the order of $\mathrm{Cl}(K)$, so p cannot divide the order of any element; hence I belongs to the trivial class in $\mathrm{Cl}(K)$, or I is principal. Then $I = \langle \alpha\rangle$ for some $\alpha \in \mathcal{O}_K$, so $\langle x + \zeta y\rangle = \langle \alpha^p\rangle$, and we get $x + \zeta y = u\alpha^p$ for some unit $u \in \mathcal{O}_K$. By proposition 2.1, then, there exist a real unit r and integer k such that

$$x + \zeta y \quad = \quad r\zeta^k \alpha^p.$$

Lemma 3.2. *There exists an integer a such that $\alpha^p \equiv a \pmod{p}$.*

Proof. Since $\mathcal{O}_K = \mathbb{Z}[\zeta]$, we may write $\alpha = \sum c_i \zeta^i$, where the c_i are integers. Then $\alpha^p = \left(\sum c_i \zeta^i\right)^p \equiv \sum \left(c_i \zeta^i\right)^p \equiv \sum c_i^p \equiv \sum c_i \pmod{p}$. □

Lemma 3.2 allows us to write $\zeta^{-k}(x + \zeta y) \equiv ra \pmod{p}$, where r is real and a is an integer, and taking complex conjugates yields $\zeta^k(x + \zeta^{-1} y) \equiv ra \pmod{p}$. Combining these two equations, we get $\zeta^{-k}(x + \zeta y) \equiv \zeta^k(x + \zeta^{-1} y)$, or

$$\zeta^k x + \zeta^{k-1} y - \zeta^{-k} x - \zeta^{-k+1} y \quad \equiv \quad 0 \pmod{p}. \tag{3.1}$$

Assume without loss of generality that $0 \leq k < p$. The proof now proceeds by a case-by-case analysis:

Case 1. Suppose $k = 0$. Then equation 3.1 reduces to $(\zeta^{-1} - \zeta)y \equiv 0 \pmod{p}$, or multiplying by ζ and factoring, $(1 - \zeta)(1 + \zeta)y \equiv 0 \pmod{p}$. But $1 + \zeta$ is a unit since $(1 + \zeta)(1 - \zeta) = (1 - \zeta^2)$ and $1 - \zeta$ and $1 - \zeta^2$ are known to be associates, so $\lambda y \equiv 0 \pmod{p}$. Therefore $\lambda y \in p\mathcal{O}_K = \mathfrak{l}^{p-1}$, and since $p > 2$ we have $y \in \mathfrak{l}^{p-2} \subset \mathfrak{l}$. We know that y is an integer, so $y \in \mathfrak{l} \cap \mathbb{Z} = p\mathbb{Z}$, and this cannot happen.

Likewise, when $k = 1$ we have $(\zeta - \zeta^{-1})x \equiv 0 \pmod{p}$, and by the same argument this implies $x \in p\mathbb{Z}$, which is impossible.

Case 2. Suppose that $k \neq 0, 1$ and $p \geq 5$. For some $\beta \in \mathbb{Z}[\zeta]$ we have $\beta p = \zeta^k x + \zeta^{k-1} y - \zeta^{p-k} x - \zeta^{p-(k-1)} y$. Since any $p - 1$ of the elements $1, \zeta, \ldots, \zeta^{p-1}$ form a basis of $\mathbb{Z}[\zeta]$, if the exponents k, $k - 1$, $p - k$, $p - k + 1$ are all incongruent

modulo p then we may pick a basis which includes these powers of ζ. Expressing β in terms of this basis and comparing coefficients, we see that p divides both x and y, which is impossible. Hence some of the exponents must be congruent modulo p: either $2k \equiv 0 \pmod{p}$, $2(k-1) \equiv 0 \pmod{p}$, or $p \equiv 2k-1 \pmod{p}$. The first two of these yield $k = 0$ and $k = 1$, which we have already eliminated, so we must have $2k \equiv 1 \pmod{p}$.

Returning to the congruence and multiplying both sides by ζ^k, we now have $\zeta^{2k}x + \zeta^{2k-1}y - x - \zeta y \equiv 0 \pmod{p}$, or $(\zeta - 1)x + (1 - \zeta)y \equiv 0 \pmod{p}$, or $\lambda(x - y) \in p\mathcal{O}_K = \mathfrak{l}^{p-1}$. Therefore $x - y \in \mathfrak{l}^{p-2} \subset \mathfrak{l}$, and since $x - y$ is an integer it must be in $\mathfrak{l} \cap \mathbb{Z}$: hence we have $x \equiv y \pmod{p}$.

We have now shown that if $x^p + y^p + z^p = 0$ and $p \nmid xyz$, then $p \geq 5$ and so $x \equiv y \pmod{p}$. But we may simply switch the roles of y and z at the beginning of this section to get $x \equiv z \pmod{p}$ as well, so $0 = x^p + y^p + z^p \equiv 3x^p \pmod{p}$. This can only happen when $x \equiv 0 \pmod{p}$, and so there are no solutions.

4. Case 2: $p \mid xyz$

Suppose that $x^p + y^p + z^p = 0$ and $p \mid xyz$, where $p \geq 3$ is regular. Assume without loss of generality that p divides z, and write $z = p^e z_0$, where p does not divide z_0; note that since x, y, and z_0 are not multiples of p in $\mathbb{Z}$, they are not multiples of λ in $\mathbb{Z}[\zeta]$. Then $\langle z \rangle = \mathfrak{l}^{e(p-1)}\langle z_0 \rangle$, so as in case 1, we have

$$\prod_{i=0}^{p-1} \langle x + \zeta^i y \rangle \quad = \quad \langle z \rangle^p = \mathfrak{l}^{pn}\langle z_0 \rangle^p, \tag{4.1}$$

where $n = e(p-1)$; we will prove that a slightly more general version of this has no solutions.

Proposition 4.1. *Equation 4.1 has no solutions, where $x, y, z_0 \in \mathbb{Z}[\zeta]$ are not multiples of λ and $n \geq 1$.*

Following the proof of proposition 3.1, we see that the greatest common divisor of any distinct ideals $\langle x + \zeta^i y \rangle$ and $\langle x + \zeta^j y \rangle$ is the ideal $\mathfrak{d} = \langle x, y \rangle$ times some power of $\mathfrak{l}$.

Let (x, y, z_0, n) be a solution of equation 4.1 such that n is minimal. Note that for all i and j we have $x + \zeta^i y \equiv x + \zeta^j y \pmod{\lambda}$, since the two sides differ by $\zeta^i y(1 - \zeta^{j-i})$ and $1 - \zeta^{j-i}$ is a multiple of $1 - \zeta = \lambda$. Therefore if λ divides $x + \zeta^i y$ for some i, it divides $x + \zeta^j y$ for all j. But $\mathfrak{l}$ is prime and $pn \geq 1$, so $\mathfrak{l}$ divides some ideal on the left hand side and therefore $\mathfrak{l} \mid \langle x + \zeta^i y \rangle$ for all i.

Suppose there exist some i and j which are distinct modulo p such that $\mathfrak{l}^2$ divides each of $\langle x + \zeta^i y \rangle$ and $\langle x + \zeta^j y \rangle$. Then both $x + \zeta^i y$ and $x + \zeta^j y$ are multiples of λ^2, so λ^2 divides their difference $y\zeta^i(1 - \zeta^{j-i})$: passing to ideals, we have $\mathfrak{l}^2 \mid \langle y \rangle \langle 1 - \zeta^{j-i} \rangle$. Since $j - i$ is not a multiple of p, $\langle 1 - \zeta^{j-i} \rangle = \mathfrak{l}$, hence by unique factorization of ideals $\mathfrak{l} \mid \langle y \rangle$. But y is not a multiple of λ, so this is impossible. We conclude that $\gcd(\langle x + \zeta^i y \rangle, \langle x + \zeta^j y \rangle) = \mathfrak{d}\mathfrak{l}$ for all distinct i and j.

We now claim that equation 4.1 has no solutions when $n = 1$. Assume such a solution exists; then $\langle x + \zeta^i y \rangle$ is always a multiple of $\mathfrak{l}$ but not of $\mathfrak{l}^2$, so $x + \zeta^i y \equiv \alpha_i \lambda \pmod{\lambda^2}$ for some $\alpha_i \in \mathbb{Z}[\zeta]$ which is not a multiple of λ. The α_i are all distinct modulo λ, since otherwise $x + \zeta^i y \equiv x + \zeta^j y \pmod{\lambda^2}$ and this is impossible as argued above. But $\mathbb{Z}[\zeta]/\lambda\mathbb{Z}[\zeta] \cong \mathbb{F}_p$, so there are only $p - 1$ nonzero residue classes modulo λ for the p values α_i. This is a contradiction, and so we must have $n \geq 2$.

As a consequence, we have $pn > p$, so $\mathfrak{l}^2 \mid \langle x + \zeta^k y \rangle$ for some unique k. Replacing y with $\zeta^k y$ in our original equation, we may assume that $k = 0$ without loss of generality. Then $\mathfrak{l}^{pn-(p-1)} \mid\mid \langle x + y \rangle$, and $\mathfrak{l} \mid\mid \langle x + \zeta^i y \rangle$ for $1 \leq i < p$.

Write $\langle x + y \rangle = \mathfrak{d}\mathfrak{l}^{p(n-1)+1} I_0$ and $\langle x + \zeta^j y \rangle = \mathfrak{d}\mathfrak{l} I_j$ for $1 \leq j < p$. Dividing each side by $\mathfrak{d}^p \mathfrak{l}^{pn}$ in equation 4.1 gives us

$$\prod_{j=0}^{p-1} I_j \quad = \quad (\mathfrak{d}^{-1} \langle z_0 \rangle)^p.$$

Since each of the I_j must be relatively prime and the right hand side is a perfect pth power, it follows that each $I_j = J_j^p$ for some ideal J_j. Then for $i \geq 1$

$$\langle \frac{x + \zeta^i y}{x + y} \rangle \quad = \quad \frac{\mathfrak{d}\mathfrak{l} J_i^p}{\mathfrak{d}\mathfrak{l}^{p(n-1)+1} J_0^p} = \mathfrak{l}^{-p(n-1)} (J_i J_0^{-1})^p,$$

and since p does not divide the order of $\mathrm{Cl}(K)$ but the fractional ideal $(J_i J_0^{-1})^p$ must be principal, $J_i J_0^{-1}$ is principal.

Set $J_i J_0^{-1} = \langle \frac{a_i}{b_i} \rangle$, where $a_i, b_i \in \mathbb{Z}[\zeta]$, for $1 \leq i < p$; since $\mathfrak{l}$ does not divide J_0 or J_i, we may assume that λ does not divide a_i or b_i. Then for some unit ϵ_i we have $\frac{x+\zeta^i y}{x+y} = \epsilon_i \lambda^{-p(n-1)} \left(\frac{a_i}{b_i}\right)^p$, or clearing denominators,

$$\lambda^{p(n-1)} b_i^p (x + \zeta^i y) \quad = \quad \epsilon_i a_i^p (x + y).$$

When $i = 1$ we have $\lambda^{p(n-1)} b_1^p (x + \zeta y) = \epsilon_1 a_1^p (x + y)$, and when $i = 2$ we have $\lambda^{p(n-1)} b_2^p (x + \zeta^2 y) = \epsilon_2 a_2^p (x + y)$. Subtracting b_1^p times the second equation from $(1 + \zeta) b_2^p$ times the first, we get

$$\lambda^{p(n-1)} (b_1 b_2)^p \left[\zeta (x + y)\right] \quad = \quad (x + y) \left[\epsilon_1 (1 + \zeta)(a_1 b_2)^p - \epsilon_2 (a_2 b_1)^p\right].$$

Since $1 + \zeta$ is a unit, we define units $\epsilon_2' = \frac{-\epsilon_2}{\epsilon_1(1+\zeta)}$ and $\epsilon_3' = \frac{\zeta}{\epsilon_1(1+\zeta)}$ and divide both sides by the nonzero $\epsilon_1 (x + y)(1 + \zeta)$ to get

$$(a_1 b_2)^p + \epsilon_2' (a_2 b_1)^p \quad = \quad \epsilon_3' \lambda^{p(n-1)} (b_1 b_2)^p. \tag{4.2}$$

We wish to eliminate the ϵ_2' term, for then we will have constructed a solution to equation 4.1 with a smaller value of n than in the original solution. Since $n \geq 2$, λ^p divides $(a_1 b_2)^p + \epsilon_2' (a_2 b_1)^p$. But λ does not divide $a_2 b_1$, so $\langle a_2 b_1 \rangle$ is relatively prime to $\mathfrak{l}$; hence there is some $c \in \mathbb{Z}[\zeta]$ such that $c a_2 b_1 \equiv 1 \pmod{\lambda}$.

Lemma 4.2. *If $\alpha \equiv \beta \pmod{\lambda}$ for some $\alpha, \beta \in \mathbb{Z}[\zeta]$, then $\alpha^p \equiv \beta^p \pmod{\lambda^p}$.*

Proof. We have $\alpha = \beta + q\lambda$ for some $q \in \mathbb{Z}[\zeta]$, so $\alpha^p = (\beta + q\lambda)^p = \beta^p + (q\lambda)^p + \sum_{i=1}^{p-1} \binom{p}{i} \beta^i (q\lambda)^{p-i}$. Each term in the summation is a multiple of $p\lambda$ and hence a multiple of λ^p, and $(q\lambda)^p$ is clearly also a multiple of λ^p, so $\alpha^p \equiv \beta^p \pmod{\lambda^p}$ as desired. □

From this lemma it follows that $c^p (a_2 b_1)^p \equiv 1 \pmod{\lambda^p}$, and so $0 \equiv c^p (a_1 b_2)^p + \epsilon_2' c^p (a_2 b_1)^p \equiv (c a_1 b_2)^p + \epsilon_2' \pmod{\lambda^p}$, or $\epsilon_2' \equiv (-c a_1 b_2)^p \pmod{\lambda^p}$. But also $-c a_1 b_2 \equiv d \pmod{\lambda}$ for some integer d, since $\mathbb{Z}[\zeta]/\lambda\mathbb{Z}[\zeta] \cong \mathbb{F}_p$, so by another application of the lemma $\epsilon_2' \equiv d^p \pmod{\lambda^p}$. Since p is an associate of λ^{p-1}, it divides λ^p, and so $\epsilon_2' \equiv d^p \pmod{p}$, or simply $\epsilon_2' \equiv d \pmod{p}$.

We now invoke a deep result known as Kummer's Lemma:

Lemma 4.3. *Let p be a regular prime, let $u \in \mathbb{Z}[\zeta]$ be a unit, and suppose there exists $m \in \mathbb{Z}$ such that $u \equiv m \pmod{p}$. Then $u = \epsilon^p$ for some unit ϵ.*

Thus we can write $\epsilon_2' = \eta^p$, where $\eta \in \mathbb{Z}[\zeta]$ is a unit. Equation 4.2 now becomes

$$(a_1b_2)^p + (\eta a_2 b_1)^p = \epsilon_3' \lambda^{p(n-1)} (b_1 b_2)^p.$$

Let $x' = a_1b_2$, $y' = \eta a_2 b_1$, and $z_0' = b_1 b_2$. Factoring the left hand side and passing to ideals, we have

$$\prod_{i=0}^{p-1} \langle x' + \zeta^i y' \rangle = \mathfrak{l}^{p(n-1)} \langle z_0' \rangle^p,$$

and we know that none of x', y', and z_0' are multiples of λ. Since $n - 1 \geq 1$, we have a solution of 4.1 of the form $(x', y', z_0', n-1)$, contradicting the minimality of n. Therefore there cannot be any solutions, and in particular the equation $x^p + y^p + z^p = 0$ does not have any solutions where p divides z.

Combining this with case I, we have now proven (modulo Kummer's lemma)

Theorem 4.4. *The equation $x^p + y^p + z^p = 0$ does not have any nontrivial solutions in integers x, y, z when p is a regular prime.*

5. Bernoulli numbers

The goal of this section is to define a sequence of numbers which provide a very useful criterion for the regularity of a prime p.

Given $n \geq 0$, let $B_n(t)$ be the unique polynomial of degree n (up to a constant) such that $\int_x^{x+1} B_n(t)dt = x^n$ for all x; we may easily see that it is unique because differentiation yields $B_n(x+1) - B_n(x) = nx^{n-1}$, and specifying a value of $b_n = B_n(0)$ guarantees that this difference equation has a unique solution. Furthermore, we have $\int_x^{x+1} B_n'(t)dt = \frac{d}{dx}\int_x^{x+1} B_n(t)dt = nx^{n-1} = n\int_x^{x+1} B_{n-1}(t)dt$, and so $B_n'(t) = nB_{n-1}(t)$.

From this differential equation it follows by an easy induction that $B_n(t) = t^n + \binom{n}{1} b_1 t^{n-1} + \cdots + \binom{n}{n} b_n$. But since $(n+1)0^n = \int_0^1 B_{n+1}'(t)dt = B_{n+1}(1) - B_{n+1}(0)$, or $B_{n+1}(0) = B_{n+1}(1)$, if we set $B_0(t) = b_0 = 1$ we have

$$\binom{n+1}{0} b_0 + \binom{n+1}{1} b_1 + \cdots + \binom{n+1}{n} b_n = 0. \tag{5.1}$$

The numbers b_n which satisfy this recurrence (and are the constant terms of the polynomials $B_n(t)$) are called the *Bernoulli numbers*. It is clear from the recurrence that the Bernoulli numbers are all rational. Furthermore, one can easily see that the problem of finding sums of consecutive nth powers reduces easily to knowing values of the Bernoulli numbers, since by the definition of the polynomials $B_n(t)$ we have $c^n + (c+1)^n + \cdots + d^n = \int_c^{d+1} B_n(t)dt = \frac{1}{n+1}\int_c^{d+1} B_{n+1}'(t)dt$, or

$$\sum_{k=c}^{d} k^n = \frac{B_{n+1}(d+1) - B_{n+1}(c)}{n+1}. \tag{5.2}$$

Finally, from equation 5.1 we get $\frac{b_n}{n!} = \frac{-1}{(n+1)!}\sum_{k=0}^{n-1}\binom{n+1}{k}b_k = \sum_{k=0}^{n-1}\frac{-1}{(n+1-k)!}\frac{b_k}{k!}$. Multiplying by x^n and summing over all n, we have

$$\begin{aligned}
\sum_{n=1}^{\infty}\frac{b_n x^n}{n!} &= \frac{-1}{x}\sum_{n=1}^{\infty}\sum_{k=0}^{n-1}\frac{x^{n+1-k}}{(n+1-k)!}\frac{b_k x^k}{k!} \\
-x\sum_{n=1}^{\infty}\frac{b_n x^n}{n!} &= \sum_{n=0}^{\infty}\sum_{k=0}^{n+1}\frac{x^{n+1-k}}{(n+1-k)!}\frac{b_k x^k}{k!} - \left(\sum_{n=0}^{\infty}\frac{b_n x^{n+1}}{n!} + \sum_{n=0}^{\infty}\frac{b_{n+1}x^{n+1}}{(n+1)!}\right) \\
&= \left[\left(\sum_{i=0}^{\infty}\frac{x^i}{i!}\right)\left(\sum_{j=0}^{\infty}\frac{b_j x^j}{j!}\right) - 1\right] - x\sum_{n=0}^{\infty}\frac{b_n x^n}{n!} - \left(\sum_{n=0}^{\infty}\frac{b_n x^n}{n!} - 1\right)
\end{aligned}$$

or, if $f(x) = \sum_{k=0}^{\infty}\frac{b_k x^k}{k!}$, then $-x(f(x)-1) = (e^x f(x) - 1) - xf(x) - (f(x)-1)$. Solving for $f(x)$, we conclude:

$$\sum_{n=0}^{\infty}\frac{b_n x^n}{n!} = \frac{x}{e^x - 1}. \tag{5.3}$$

As a corollary, note that $\frac{x}{2} + \sum\frac{b_n x^n}{n!} = \frac{x}{2}(1 + \frac{2}{e^x-1}) = \frac{x}{2}\left(\frac{e^x+1}{e^x-1}\right)$. The quantity in parentheses is an odd function, since $\frac{e^{-x}+1}{e^{-x}-1} = \frac{1+e^x}{1-e^x}$, so $\frac{x}{2} + \sum\frac{b_n x^n}{n!}$ is even. Therefore $b_1 = -\frac{1}{2}$ and $b_n = 0$ for all odd $n \geq 3$.

The criterion Kummer found for regularity is the following, which we will prove in the following sections:

Claim 5.1. An odd prime p is regular if and only if p divides the numerators of the Bernoulli numbers $b_2, b_4, \ldots b_{p-3}$.

This claim, together with recurrence 5.1, give us a criterion which can easily be checked to determine whether a prime is regular. We note the following useful fact:

Proposition 5.2. *The quantity* $(n+1)!b_n$ *is an integer for all* $n \geq 0$.

Proof. Clearly the assertion is true for $n = 0$; we will prove it for $n \geq 1$ by induction. Multiplying equation 5.1 by $n!$, we get $(n+1)!b_n = -\sum_{i=0}^{n-1}\binom{n+1}{i}(n!b_i)$. If $(i+1)!b_i$ is an integer for all $i < n$, then $n!b_i$ must also be an integer. But then the entire right hand side is an integer, and so $(n+1)!b_n$ is an integer as well. □

Therefore if p is prime, then p does not divide the denominators of the numbers b_i (when written in lowest terms) for $i \leq p-2$, since $(p-1)!b_i$ is an integer and $p \not| (p-1)!$. In particular, we may treat these b_i as integers mod p and use equation 5.1 to compute $b_2, b_4, \ldots, b_{p-3} \pmod p$; the prime p is regular if and only if any of these are zero.

Assuming the truth of claim 5.1 and the ability to do arithmetic mod p in constant time, we now have an algorithm which can determine if p is regular: for each even i, $2 \leq i \leq p-3$, compute the $(i+1)$st row of Pascal's triangle mod p and use this to calculate the sum $\sum_{j=0}^{i-1}\binom{i+1}{j}b_j \pmod p$, then multiply by $-(i+1)^{-1}$ to get b_i. Each of these steps takes linear time (except the $O(1)$ multiplication), since the coefficients $\binom{i+1}{j}$ can each be computed from the $(i-1)$st row in constant time using the recurrence $\binom{i+1}{j} = \binom{i-1}{j-2} + 2\binom{i-1}{j-1} + \binom{i-1}{j}$. We conclude that this algorithm determines the regularity of p in $O(p^2)$ time.

6. The class number of $\mathbb{Q}(\zeta_p)$

Fix an odd prime p and let $K = \mathbb{Q}(\zeta_p)$. For a choice of primitive root γ modulo p such that $\gamma^{p-1} \not\equiv 1 \pmod{p^2}$, we define $\gamma_i \equiv \gamma^i \pmod{p}$ such that $0 < \gamma_i < p$ for all i, and then we construct a polynomial $\phi(x) = \sum_{k=0}^{p-2} \gamma_k x^k$. Of course, this requires us to know that such γ exists:

Proposition 6.1. *There exists some γ, $1 < \gamma < p$, such that γ is a primitive root modulo both p and p^2.*

Proof. Let g be a primitive root mod p with $1 < g < p$, and let h and H be the smallest positive inverses of g mod p and p^2, respectively. Then $gh < p^2$ and $gH > p^2$, so we must have $H = h + pd$ for some integer d, $0 < d < p$. If g is not a primitive root mod p^2, then we must have $g^{p-1} \equiv 1 \pmod{p^2}$, since the order of g mod p^2 is a proper divisor of $\phi(p^2) = p(p-1)$ but a multiple of its order mod p (namely, $p-1$). But $(gH)^{p-1} \equiv 1 \pmod{p^2}$, so $H^{p-1} \equiv 1 \pmod{p^2}$ as well. Since $H = h + pd$, we have $1 \equiv (h+pd)^{p-1} \equiv h^{p-1} + \binom{p-1}{1} h^{p-2} pd \pmod{p^2}$ by the binomial theorem, so $h^{p-1} \equiv 1 + (h^{p-2}d)p \not\equiv 1 \pmod{p^2}$. Since h has order $p-1$ mod p, its order mod p^2 is a multiple of $p-1$ but a divisor of $p(p-1)$, hence it must be $p(p-1)$; that is, h is the desired primitive root. □

Let $\alpha = \zeta = \zeta_p$ and $\beta = \zeta_{p-1}$ be primitive pth and $(p-1)$st roots of unity, respectively. Let $\mu = \frac{p-1}{2}$, and consider the automorphism $\sigma : \alpha \mapsto \alpha^\gamma$. Note that $\sigma^\mu \alpha = \alpha^{\gamma^\mu} = \alpha^{-1}$, since $\gamma^\mu \equiv -1 \pmod{p}$.

Lemma 6.2. *For any rational integers $x_1, x_2, \cdots, x_\mu$ which sum to zero, the product $\pm\alpha^k(1-\sigma\alpha)^{x_1}(1-\sigma^2\alpha)^{x_2}\cdots(1-\sigma^\mu\alpha)^{x_\mu}$ is a unit in $\mathbb{Z}[\zeta]$.*

Proof. Call this product $P \in K$. Since $\sigma^i\alpha = \zeta^{k_i}$for some k_i, we know that $1 - \sigma^i\alpha$ and $1-\zeta$ are associates; let ϵ_i denote their quotient. Hence $P = \pm\alpha^k(1-\zeta)^{x_1+x_2+\cdots+x_\mu}\prod \epsilon_i^{x_i}$. But $\sum x_i = 0$, so $P = \pm\alpha^k \prod \epsilon_i^{x_i}$, and all of the terms on the right hand side are units in $\mathbb{Z}[\zeta]$. □

If we let U_K denote the group of units of $\mathbb{Z}[\zeta]$ and $H \subset U_K$ the subgroup of units of the form $\pm\alpha^k \prod (1-\sigma^i\alpha)^{x_i}$ as in the lemma, then H is a subgroup of finite index.

Kummer showed, via some extensive analytic calculations, the following result:

Theorem 6.3. *K has class number $h = h_1 h_2$, where $h_1 = \frac{|\phi(\beta)\phi(\beta^3)\cdots\phi(\beta^{p-2})|}{(2p)^{\mu-1}}$ and $h_2 = [U_K : H]$.*

See [4] for a detailed proof, or [1] for a proof which characterizes h_2 diffferently as the class number of the field $\mathbb{Q}(\alpha + \alpha^{-1})$.

In order to determine whether a prime p is regular, it now suffices to check that p does not divide h_1 or h_2. We will provide a criterion for determining whether $p \mid h_1$ in this section and show in the next section that this is the only needed criterion because $p \mid h_2$ only if $p \mid h_1$.

Let $P = \prod_{i=0}^{\mu-1} \phi(\beta^{2i+1})$, so that $h_1 = \frac{|P|}{(2p)^{\mu-1}}$. Since $\phi(x) = \sum_{k=0}^{p-2} \gamma_k x^k$ and $\gamma\gamma_k \equiv \gamma_{k+1} \pmod{p}$, we have $(\gamma x - 1)\phi(x) = \sum_{k=1}^{p-2}(\gamma\gamma_k - \gamma_{k+1})x^k + \gamma\gamma_{p-2}x^{p-1} - \gamma_0$. Then $\gamma\gamma_{p-2} \equiv \gamma^{p-1} \equiv 1 \pmod{p}$, and $\gamma_0 = 1$, so $(\gamma x - 1)\phi(x) = p\psi(x) + x^{p-1} - 1$ for some polynomial $\psi(x)$ with integer coefficients. Applying this for

$x = \beta, \beta^3, \ldots, \beta^{p-2}$, and noting that $(\beta^i)^{p-1} - 1 = 0$, for each i we get

$$(\gamma\beta - 1)\cdots(\gamma\beta^{p-2} - 1)P \quad = \quad p^\mu \psi(\beta)\psi(\beta^3)\cdots\psi(\beta^{p-2}).$$

Now, $\beta, \beta^3, \ldots, \beta^{p-2}$ are the roots of $\frac{x^{p-1}-1}{x^{(p-1)/2}-1} = x^{(p-1)/2}+1$, so $\prod(x-\beta^i) = x^\mu+1$, where i runs through $1, 3, \ldots, p-2$. Substituting $x = \gamma^{-1}$ and multiplying through by γ^μ, we get $\prod(1-\gamma\beta^i) = 1+\gamma^\mu$, hence $\prod(\gamma\beta^i - 1) = (-1)^\mu(\gamma^\mu+1)$. Therefore

$$(-1)^\mu(\gamma^\mu + 1)P \quad = \quad p^\mu \psi(\beta)\psi(\beta^3)\cdots\psi(\beta^{p-2}).$$

Since γ is a primitive root modulo p and $2\mu = p-1$, we have $\gamma^\mu \equiv -1 \pmod p$, or $p \mid \gamma^\mu + 1$. But $\gamma^\mu + 1$ divides $\gamma^{p-1} - 1$, and this is not a multiple of p^2. It follows that $p^{\mu-1}$ divides P, and p divides $h_1 = \frac{|P|}{(2p)^{\mu-1}}$ if and only if it divides $\psi(\beta)\psi(\beta^3)\cdots\psi(\beta^{p-2})$. Next, since β and γ are both primitive $(p-1)$st roots of unity modulo p, it follows that $\psi(\beta)\psi(\beta^3)\cdots\psi(\beta^{p-2}) \equiv \psi(\gamma)\psi(\gamma^3)\cdots\psi(\gamma^{p-2}) \pmod p$. Therefore $p \nmid h_1$ if and only if p does not divide any of the $\psi(\gamma^i)$.

Lemma 6.4. *p divides $\psi(\gamma^n)$, $1 \le n < p-2$, if and only if $\sum_{k=1}^{\gamma-1} B_{n+1}\left(\frac{k}{\gamma}\right) \equiv (\gamma-1)B_{n+1}(p) \pmod p$.*

Proof. Note first that the Bernoulli polynomials involved are well-defined modulo p, since their coefficients are Bernoulli numbers b_k, $k < p-1$, and by proposition 5.2 we know that p does not divide the denominators of any of these numbers.

Define c_j such that $pc_j = \gamma\gamma_j - \gamma_{j+1}$ for $j < p-1$ and $pc_{p-1} = \gamma\gamma_{p-2} - 1$, so that $\psi(x) = \sum_{k=1}^{p-1} c_k x^k$. Note that $0 \le c_j < \gamma$ for all j, since γ_{j+1} is the least positive integer congruent to the positive $\gamma\gamma_j$ modulo p and $pc_j < \gamma p$. In fact, it is clear that if $\frac{kp}{\gamma} < \gamma_j < \frac{(k+1)p}{\gamma}$, then $c_j = k$. We may write $\psi(\gamma^n) \equiv \sum_{k=1}^{p-1} c_k \gamma_k^n \pmod p$, and since each of the γ_k is congruent to a unique integer between 1 and $p-1$, we have

$$\psi(\gamma^n) \quad \equiv \quad \sum_{k=1}^{p-1} \lfloor \frac{\gamma k}{p} \rfloor k^n \pmod p.$$

Rearranging terms and applying equation 5.2, we get

$$\begin{aligned} \psi(\gamma^n) \quad &\equiv \quad \sum_{k=1}^{\gamma-1} \left(\lceil \frac{kp}{\gamma} \rceil^n + \cdots + (p-1)^n \right) \pmod p \\ &\equiv \quad \frac{1}{n+1} \sum_{k=1}^{\gamma-1} \left(B_{n+1}(p) - B_{n+1}\left(\lceil \frac{kp}{\gamma} \rceil \right) \right). \end{aligned}$$

But $n+1$ is invertible modulo p, so $\psi(\gamma^n) \equiv 0 \pmod p$ iff the summation is a multiple of p, and thus we may drop the $\frac{1}{n+1}$ term.

Finally, note that the numbers $\gamma\lceil\frac{p}{\gamma}\rceil, \gamma\lceil\frac{2p}{\gamma}\rceil, \ldots, \gamma\lceil\frac{(\gamma-1)p}{\gamma}\rceil$ are congruent to some permutation of $1, 2, \ldots, \gamma-1 \pmod p$. This is true because they are all distinct modulo p (since γ is invertible) but $kp \le \gamma\lceil\frac{kp}{\gamma}\rceil < kp + \gamma$. Therefore the numbers $B_{n+1}\left(\lceil\frac{kp}{\gamma}\rceil\right)$, $1 \le k \le \gamma-1$, are congruent to the numbers $B_{n+1}\left(\frac{i}{\gamma}\right)$, $1 \le i \le \gamma-1$, and we are done. □

Recall from the previous section that $(\gamma x)^m = \int_{\gamma x}^{\gamma x+1} B_m(t)dt$, hence $\gamma^m x^m = \int_x^{x+1/\gamma} \gamma B_m(\gamma u)du$ by a simple substitution. But also $\gamma^m x^m = \gamma^m \int_x^{x+1} B_m(t)dt$,

which is in turn equal to $\gamma^m \left[\int_x^{x+1/\gamma} \left(B_m(t) + B_m(t+\frac{1}{\gamma}) + \ldots B_m(t+\frac{\gamma-1}{\gamma})\right) dt\right]$. Since this is true for any x, we conclude that $\gamma B_m(\gamma t) = \gamma^m[B_m(t) + \cdots + B_m(t+\frac{\gamma-1}{\gamma})]$. Finally, letting $t=0$, we get $\gamma b_m = \gamma^m[b_m + B_m(\frac{1}{\gamma}) + \cdots + B_m(\frac{\gamma-1}{\gamma})]$, or

$$(\gamma^{1-m}-1)b_m \quad = \quad \sum_{k=1}^{\gamma-1} B_m(\frac{k}{\gamma}).$$

Combining this with the lemma, we see that for $1 \leq n < p-2$, p divides $\psi(\gamma^n)$ if and only if $(\gamma^{-n}-1)b_{n+1} \equiv (\gamma-1)B_{n+1}(p) \pmod p$. But $B_{n+1}(p) \equiv B_{n+1}(0) = b_{n+1} \pmod p$, so this condition is equivalent to $(\gamma^{-n}-\gamma)b_{n+1} \equiv 0 \pmod p$. We know that $\gamma^{n+1} \not\equiv 1 \pmod p$ since γ is a primitive root and $n+1 < p-1$, so we may cancel out the term $\gamma^{-n}-\gamma$ to conclude that for $n < p-2$, p divides $\psi(\gamma^n)$ if and only if p divides the numerator of b_{n+1}.

We now handle the case $n = p-2$ separately: in fact, we assert that p never divides $\psi(\gamma^{p-2})$. Write $\gamma^{p-1} - 1 = p\nu$, noting that ν is not a multiple of p. Then $(\gamma \cdot \gamma^{p-2} - 1)\phi(\gamma^{p-2}) = p\psi(\gamma^{p-2}) + (\gamma^{(p-2)(p-1)} - 1)$. Dividing by p, we have $\nu\phi(\gamma^{p-2}) = \psi(\gamma^{p-2}) + \frac{\gamma^{(p-2)(p-1)}-1}{p}$, where the last term is an integer. In fact, since $\gamma^{p-1} \equiv p\nu + 1 \pmod{p^2}$, we may raise both sides to the $(p-2)$nd power to get $\gamma^{(p-2)(p-1)} \equiv (p-2)p\nu + 1 \pmod{p^2}$, so $\frac{\gamma^{(p-2)(p-1)}-1}{p} \equiv (p-2)\nu \pmod p$. Therefore

$$\nu\phi(\gamma^{p-2}) \quad \equiv \quad \psi(\gamma^{p-2}) - 2\nu \pmod p.$$

But $\phi(\gamma^{p-2}) = \sum_{k=0}^{p-2} \gamma_k(\gamma^{p-2})^k \equiv \sum_{k=0}^{p-2} \gamma^{(p-1)k} \equiv \sum_{k=0}^{p-2} 1^k \equiv p-1 \pmod p$, so we get $\psi(\gamma^{p-2}) \equiv \nu \pmod p$, and thus p does not divide $\psi(\gamma^{p-2})$.

Combining the information we have about each $\psi(\gamma^k)$, we conclude the following:

Theorem 6.5. *p divides h_1 if and only if p divides the numerator of one of the Bernoulli numbers $b_2, b_4, \ldots, b_{p-3}$.*

7. h_2 and Kummer's lemma

In order to determine whether p is regular, we must also consider whether or not p divides h_2. It turns out that this is in fact irrelevant. More precisely, we will prove the following theorem (and obtain Kummer's Lemma as a corollary):

Theorem 7.1. *If p divides h_2, then p divides h_1.*

Suppose $p \mid h_2$, and recall that $h_2 = [U_K : H]$. Since p is prime, U_K/H contains an element of order p; that is, there is some unit ϵ not of the form $\pm\zeta^k \prod_{i=1}^{\mu}(1-\sigma^i\zeta)^{x_i}$, $\sum x_i = 0$, such that ϵ^p has that form (i.e. $\epsilon^p \in H$). Fix a choice of sign, k, and all values of x_i so that $\epsilon^p = \pm\zeta^k \prod(1-\sigma^i\zeta)^{x_i}$. Then we cannot have $p \mid x_i$ for all i: otherwise, we would have $\epsilon^p = \pm\zeta^k\eta^p$, where $\eta = \prod(1-\sigma^i\zeta)^{y_i}$ and $x_i = py_i$, so $(\epsilon\eta^{-1})^p = \pm\zeta^k$ and thus $\pm\zeta^k = \pm 1$. We would then have $\epsilon^p = (\pm\eta)^p$, so $\epsilon = \pm\zeta^j \prod(1-\sigma^i\zeta)^{y_i}$ for some j, which contradicts the assumption that $\epsilon \notin H$.

Since we can write $\epsilon = a_0 + a_1\zeta + \cdots + a_{p-1}\zeta^{p-1}$ for some integers a_i, we have $\epsilon^p \equiv \sum(a_i\zeta^i)^p \equiv \sum a_i \pmod p$, so there is an integer c such that $\epsilon^p \equiv c \pmod p$. Therefore (absorbing a factor of -1 into c if necessary)

$$\zeta^k(1-\sigma\zeta)^{x_1}(1-\sigma^2\zeta)^{x_2}\cdots(1-\sigma^\mu\zeta)^{x_\mu} \quad \equiv \quad c \pmod p \tag{7.1}$$

for some integers k, c, x_i such that $\sum x_i = 0$ and the x_i are not all multiples of p.

Define the rational function $f(t) = t^k \prod_{i=1}^{\mu}(1-\sigma^i t)^{x_i}$, where $\sigma^i t = t^{\gamma^i}$. Equation 7.1 gives us $f(t) = c + p \cdot q(t) + (1 + t + \cdots + t^{p-1})r(t)$ for some polynomials $q(t)$ and $r(t)$ with integer coefficients. Taking logarithmic derivatives, we get

$$\frac{k}{t} - \sum_{i=1}^{\mu} x_i \frac{\gamma^i \sigma^i t}{t(1-\sigma^i t)} \quad = \quad \frac{p \cdot q'(t) + (1 + \cdots + t^{p-1})r'(t) + \left(\sum_{i=0}^{p-2}(i+1)t^i\right) r(t)}{c + p \cdot q(t) + (1 + t + \cdots + t^{p-1})r(t)}$$

and if we multiply through by ζ and then set $t = \zeta$ we get

$$k - \sum_{i=1}^{\mu} \frac{x_i \gamma^i \sigma^i \zeta}{1 - \sigma^i \zeta} \quad = \quad \frac{\zeta p \cdot q'(\zeta) + \zeta \left(\sum_{i=0}^{p-2}(i+1)\zeta^i\right) r(\zeta)}{f(\zeta)}.$$

Note that $f(\zeta) = \pm\epsilon^p$, so the right hand side is in $\mathbb{Z}[\zeta]$, and $\sum_{i=0}^{p-2}(i+1)\zeta^i = \frac{p\zeta^{p-1}}{\zeta - 1}$ (which one may compute by calling it S and noting that $\zeta S - S = (p-1)\zeta^{p-1} - \sum_{i=0}^{p-2}\zeta^i = p\zeta^{p-1}$). Thus if we let η denote the unit $-(\pm\epsilon^p)^{-1}$, we have

$$k - \left(\sum_{i=1}^{\mu} x_i \gamma^i \sigma^i\right) \frac{\zeta}{1-\zeta} \quad \equiv \quad \eta \left(\frac{p}{1-\zeta}\right) r(\zeta) \pmod{p}.$$

But $\frac{1}{1-\zeta} = \frac{-\zeta}{p}\left(\frac{p\zeta^{p-1}}{\zeta - 1}\right) = \frac{-1}{p}\left(\zeta + 2\zeta^2 + \cdots + (p-1)\zeta^{p-1}\right)$, so if we write $\gamma_i \equiv \gamma^i \pmod{p}$ as before and rearrange the terms $j\zeta^j$, noting that $\gamma_i \zeta^{\gamma_i} = \gamma_i \zeta^{\gamma^i} = \gamma_i \sigma^i \zeta$, we get $\frac{\zeta}{1-\zeta} = \frac{1}{1-\zeta^{-1}} = \sigma^\mu\left(\frac{1}{1-\zeta}\right) = \frac{-1}{p}\sigma^\mu\left(\sum_{i=1}^{p-1}\gamma_i\sigma^i\right)\zeta$. Writing the operator in parentheses as $\phi(\sigma)$, where $\phi(x) = \sum \gamma_i x^i$ as before, we have

$$k + \left(\sum_{i=1}^{\mu} x_i \gamma^i \sigma^i\right) \frac{\sigma^\mu(\phi(\sigma))(\zeta)}{p} \quad \equiv \quad \eta \left(\frac{p}{1-\zeta}\right) r(\zeta) \pmod{p}.$$

Note on the left that σ^μ commutes with the operator in parentheses, since that operator is a linear combination of powers of σ. Also, if we write $\eta \cdot r(\zeta) = x \cdot (1-\zeta) - d$ for some $x \in \mathbb{Z}[\zeta]$ and $d \in \mathbb{Z}$ (recall that we can do this because $\mathbb{Z}[\zeta]/(1-\zeta)\mathbb{Z}[\zeta] \cong \mathbb{F}_p$), the right hand side becomes $px - \frac{p}{1-\zeta}d \equiv \frac{p}{1-\zeta}(-d) \pmod{p}$. But again $\frac{p}{1-\zeta} = -(\phi(\sigma))(\zeta)$, so

$$k + \sigma^\mu\left(\sum_{i=1}^{\mu} x_i \gamma^i \sigma^i\right) \frac{(\phi(\sigma))(\zeta)}{p} \quad \equiv \quad d \cdot (\phi(\sigma))(\zeta) \pmod{p},$$

Finally, we apply the operator $\gamma\sigma - 1$ to both sides and recall that $(\gamma x - 1)\phi(x) = p\psi(x) + (x^{p-1} - 1)$, and so $(\gamma\sigma - 1)(\phi(\sigma))(\zeta) = (p\psi(\sigma) + (\sigma^{p-1} - 1))\zeta = (p\psi(\sigma))(\zeta)$, to get

$$(\gamma k - k) + \sigma^\mu\left(\sum_{i=1}^{\mu} x_i \gamma^i \sigma^i\right)(\psi(\sigma))(\zeta) \quad \equiv \quad dp(\psi(\sigma))(\zeta) \equiv 0 \pmod{p}.$$

Let $F(\sigma) = (\sum_{i=1}^{\mu} x_i \gamma^i \sigma^i)\psi(\sigma)$. Since $\sigma^\mu(F(\sigma))\zeta \equiv k - \gamma k \pmod{p}$, and $k - \gamma k$ is an integer, we may apply $(\sigma^\mu)^{-1}$ to get $(F(\sigma))\zeta \equiv k - \gamma k \pmod{p}$. Furthermore, $F(\sigma)$ commutes with σ^j, so $(F(\sigma))(\sigma^j\zeta) \equiv \sigma^j(k - \gamma k) \equiv k - \gamma k \pmod{p}$ for all j, so we see that $F(\sigma)$ carries all powers of ζ to the integer $K = k - \gamma k$ modulo p.

We now consider σ as an operator on the $\mathbb{F}_p$-vector space $\mathbb{Z}[\zeta]/p\mathbb{Z}[\zeta]$. Note that this space has dimension $\leq p - 1$, since the elements ζ^i $(0 \leq i \leq p-2)$ span it.

Proposition 7.2. *The operator σ has eigenvectors $v_j = \zeta + \gamma^j\sigma\zeta + \gamma^{2j}\sigma^2\zeta + \cdots + \gamma^{(p-2)j}\sigma^{p-2}\zeta$, where $\sigma v_j = \gamma^{-j}v_j$, for $0 \le j \le p-2$. Since the eigenvalues $\gamma^0, \gamma^{-1}, \ldots, \gamma^{-(p-2)}$ are all distinct modulo p, $\mathbb{Z}[\zeta]/p$ has dimension $p-1$ and the elements v_j form a basis.*

Proof. Note that $\sigma v_j = \sum_{i=0}^{p-2}\gamma^{ij}\sigma^{i+1}\zeta$, so $\gamma^j\sigma v_j = \sum_{i=0}^{p-2}\gamma^{j(i+1)}\sigma^{i+1}\zeta = v_j$, or $\sigma v_j = \gamma^{-j}v_j$. It is easy to check that $v_j \neq 0$ for all j, so the proposition follows. □

As a result, note that for any polynomial $f(x)$, we have $f(\sigma)v_j = f(\gamma^{-j})v_j$ in this vector space. In particular, $F(\sigma)v_j \equiv \left(\sum x_i\gamma^{i(1-j)}\right)\psi(\gamma^{-j})v_j \pmod p$. But since $F(\sigma)\zeta^i \equiv K \pmod p$ for all i, we also have $F(\sigma)v_j \equiv \sum_{i=0}^{p-2}F(\sigma)(\gamma^{ij}\sigma^i\zeta) \equiv \sum_{i=0}^{p-2}\gamma^{ij}K \pmod p$. Since $\sum\gamma^{ij} = \frac{\gamma^{(p-1)j}-1}{\gamma^j-1} \equiv 0 \pmod p$, we have $F(\sigma)v_j \equiv 0 \pmod p$ for $1 \le j \le p-2$.

Since $F(\gamma^{-j})v_j \equiv \left(\sum_{i=1}^{\mu}x_i\gamma^{i(1-j)}\right)\psi(\gamma^{-j})v_j \equiv 0 \pmod p$, either $\psi(\gamma^{-j}) \equiv 0 \pmod p$ or $\sum x_i\gamma^{i(1-j)} \equiv 0 \pmod p$ for each of $i = 1, 2, \ldots, p-2$. Suppose that $p \nmid h_1$; then from the previous section we know that $\psi(\gamma^i) \not\equiv 0 \pmod p$ for $i = 1, 3, \ldots, p-2$, or equivalently $\psi(\gamma^{-j}) \not\equiv 0 \pmod p$ for $j = 1, 3, \ldots, p-2$. Hence for these μ values of j we have

$$x_1\gamma^{1-j} + x_2\gamma^{2(1-j)} + \cdots + x_\mu\gamma^{\mu(1-j)} \equiv 0 \pmod p.$$

Thus we have a system of μ equations in the μ unknowns $x_1, x_2, \ldots, x_\mu$, and the coefficient matrix $(\gamma^{i(1-j)})$ is invertible because it can be easily transformed into a Vandermonde matrix, so the only solution is $x_i \equiv 0 \pmod p$ for $1 \le i \le \mu$. But recall the construction of the unit ϵ which we used to define the x_i at the beginning of this section: we assumed that $p \mid h_2$ and showed that the x_i cannot all be multiples of p. Since this is clearly the case here, we have a contradiction, so we conclude that $p \mid h_1$ and thus our proof of theorem 7.1 is complete.

Corollary 7.3. *A prime p is regular if and only if p does not divide the numerators of the Bernoulli numbers $b_2, b_4, \ldots, b_{p-3}$.*

The corollary follows by observing that p divides $h = h_1h_2$ if and only if p divides h_1, and applying theorem 6.5.

We may now prove Kummer's Lemma on units quite easily:

Theorem 7.4. *[Kummer's Lemma] Let p be an odd regular prime, and let $u \in \mathbb{Z}[\zeta]$ be a unit. Then $u \equiv c \pmod p$ for some rational integer c if and only if u is the pth power of some unit of $\mathbb{Z}[\zeta]$.*

Proof. It is clear that for integers a_i we have $\left(\sum a_i\zeta^i\right)^p \equiv \sum a_i \pmod p$, so perfect pth powers are always congruent to rational integers modulo p.

Suppose $u \equiv c \pmod p$. Since $|U_K/H| = h_2$, we have $u^{h_2} \in H$, and so $u^{h_2} = \pm\zeta^k\prod_{i=0}^{\mu}(1-\sigma^i\zeta)^{x_i}$ for some k and integers x_i which sum to zero, hence $\pm\zeta^k\prod(1-\sigma^i\zeta)^{x_i} \equiv c^{h_2} \pmod p$. We have just shown that if p does not divide h_1, then $p \mid x_i$ for all i, so $u^{h_2} = \pm\zeta^k\eta^p$ where $\eta = \prod(1-\sigma^i\zeta)^{x_i/p}$ is a unit. Furthermore, p does not divide h_2, so $ah_2 + bp = 1$ for some $a, b \in \mathbb{Z}$; then $u = u^{ah_2}u^{bp} = (\pm\zeta^k)^a\eta^{ap}u^{bp}$, so $u = \zeta^{ak}(\pm\eta^au^b)^p$. Therefore $\pm c(\eta^{-a}u^{-b})^p \equiv \zeta^{ak} \pmod p$, and so $\zeta^{ak} \equiv c' \pmod p$ for a rational integer c'. But raising both sides to the pth power gives $1 \equiv (c')^p \equiv c' \pmod p$, so $\zeta^{ak} \equiv 1 \pmod p$, or p divides $1-\zeta^{ak}$. If $ak \not\equiv 0 \pmod p$, then ak is an associate of $1-\zeta$, hence p divides $1-\zeta$. This is impossible

since p and $(1-\zeta)^{p-1}$ are associates, so we conclude that $p \mid ak$, hence $\zeta^{ak} = 1$. It follows that $u = (\pm\eta^a u^b)^p$, as desired. □

This result completes the proof of Fermat's Last Theorem for regular primes. In other words, the equation $x^p + y^p = z^p$ has no nontrivial solutions whenever p is a regular prime, and we have shown that p is regular if and only if p does not divide the numerators of $b_2, b_4, \ldots, b_{p-3}$.

References

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

[2] Conrad, Keith, "Fermat's Last Theorem for Regular Primes," available http://www.math.uconn.edu/~kconrad/blurbs/.

[3] Conrad, Keith, "Kummer's Lemma," available http://www.math.uconn.edu/~kconrad/blurbs/.

[4] Edwards, Harold M. *Fermat's Last Theorem: A Genetic Introduction to Algebraic Number Theory*, Springer-Verlag, New York, 1977.

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

[6] Ribenboim, Paulo, *13 Lectures on Fermat's Last Theorem*, Springer-Verlag, New York, 1979.

[7] Stein, William, *Introduction to Algebraic Number Theory* (course notes), available http://modular.fas.harvard.edu/129-05/notes/129.pdf.

[8] Stewart, Ian and David Tall, *Algebraic Number Theory and Fermat's Last Theorem*, A K Peters, Natick, Massachusetts, 2002.

[9] Wong, Erick Bryce <erick@cecm.sfu.ca>, "Re: Primitive roots of p^2," article <b81skf$7rb$1@morgoth.sfu.ca> in Usenet newsgroup sci.math, 4/21/2003.