Bulletin of the *Transilvania* University of Braşov
Series III: Mathematics and Computer Science, Vol. 6(68), No. 1 - 2026, 131-158
https://doi.org/10.31926/but.mif.2026.6.68.1.10

AN ALTERNATIVE PROOF OF THE CATALAN CONJECTURE

Chang LIU[*,1] **and Preda MIHĂILESCU**[2]

Abstract

The original proof [8] of Catalan's conjecture uses real binomial power series expansions and annihilation of real class groups, based on a Theorem of Thaine. In this paper we provide a simplified approach to this proof, which uses Stickelberger annihilation of the minus part of the class group, and thus does not require the Theorem of Thaine.

2020 Mathematics Subject Classification: 11D61, 11R18, 11R27, 11D41.
Key words: diophantine equations, Stickelberger ideal, Catalan equation.

1 Introduction

Catalan's equation is

$$X^m - Y^n = 1; \quad X, Y \in \mathbb{N}; \quad m, n \geq 2, \tag{1}$$

and it was proved to have no solutions except for $(X, Y, m, n) = (3, 2, 2, 3)$, [8]. We recommend the reader interested in more historical details about the progress towards the proof of this fact, to read the books like [10], [2] or [9] written on the subject.

The purpose of the present paper is to provide a new proof of the above fact, based on a simplification brought to the arguments in [8]. First, we shall consider integer solutions $X, Y \in \mathbb{Z}$, and due to early results – see [9] – we may assume that $m, n > 3$ in (1). It will thus suffice to show that the equation

$$x^p - y^q = 1; \quad \text{with primes } p, q > 3 \text{ and integers } x, y \in \mathbb{Z} \setminus \{0, \pm 1\} \tag{2}$$

has no solutions. Allowing integer solutions has the advantage that if $(x, y; p, q)$ is a solution to (2), then so is $(-y, -x, q, p)$: if there is a solution to the pair (p, q)

[1]* *Corresponding author*, Mathematisches Institut der Universität Göttingen, e-mail: chang.liu@mathematik.uni-goettingen.de
[2]Mathematisches Institut der Universität Göttingen, e-mail: preda@uni-math.gwdg.de

of exponents, there is one also for the exponents in switched order, (q, p). This remark allows us to impose the condition $q > p$, without restriction of generality.

The work of Cassels [3] allows one to assume that if there are solutions to (2), then the following necessary conditions hold, among others:

$$\frac{x^p - 1}{x - 1} = pz^q; \quad x - 1 = p^{q-1}b^q \quad y = pbz. \tag{3}$$

Based on this result, Hyyrö proved the following lower bound: $|x| \geq p^{q-1} \cdot (q-1)^p$. We shall use in this paper the stronger lower bound

$$|x| > p^{\frac{3}{2}pq}, \tag{4}$$

proved in [6]. We shall show that (3) has no solutions with $x \in \mathbb{Z} \setminus \{0, \pm 1\}$, and odd primes $p, q > 3$, thus proving:

Theorem 1. *The equation (3) has no non-trivial solutions with $q > p > 3$.*

By the above, this implies the truth of the Catalan Conjecture, thus offering a new proof of the same.

1.1 Notations and overview of the approach

Let $\mathbb{K} = \mathbb{Q}[\zeta_p] = \mathbb{Q}[\zeta] = \mathbb{Q}[X]/(\Phi_p(X))$ be the $p-$th cyclotomic extension with, $\Phi_p(X) = \frac{X^p-1}{X-1}$, the $p-$th cyclotomic polynomial and ζ, a root thereof. Denote by $G = \text{Gal}\,(\mathbb{K}/\mathbb{Q})$ the Galois group of $\mathbb{K}/\mathbb{Q}$. We denote by $P := \{0, 1, 2, \ldots, p-1\}$ the set of least non-negative representatives of $\mathbb{F}_p$, and by $P^* := P \setminus \{0\}$ those of $\mathbb{F}_p^\times$. Similarly, we set $Q := \{0, 1, 2, \ldots, q-1\}$, and $Q^* := Q \setminus \{0\}$, for $\mathbb{F}_q$ and $\mathbb{F}_q^\times$, respectively. Let the automorphisms of G be denoted by $\sigma_c \in G : \zeta \mapsto \zeta^c$, for $c \in P^*$. We write, in multiplicative notation, $a^\sigma = \sigma(a)$ for all $a \in \mathbb{K}$ and $\sigma \in G$. We let $\jmath \in G$ be the complex conjugation and $\lambda = (1 - \zeta)$, which is an algebraic integer generating the unique totally ramified prime ideal above $\wp \supset p$ in $\mathbb{K}$. The fractional and principal ideals of $\mathbb{K}$ are $\mathcal{F}(\mathbb{K}), \mathcal{P}(\mathbb{K})$, respectively and the class group is $\mathcal{C}(\mathbb{K}) = \mathcal{F}/\mathcal{P}$. The Stickelberger Ideal $I \subset \mathbb{Z}[G]$ annihilates the class group; it is introduced in more detail in subsection 2.2 below.

Assuming that (3) has a non-trivial solution, we show that the G-orbit of $\alpha := \frac{x-\zeta}{1-\zeta} \in \mathbb{Z}[\zeta]$ is built of mutually coprime algebraic integers and the ideal $\mathfrak{A} = (\alpha, z)$ has order dividing q. By applying some $\theta \in I$ to $\mathfrak{A}$ we obtain identities of the type

$$(\beta/\overline{\beta})^q = \eta \left(\frac{1 - \zeta/x}{1 - \overline{\zeta}/x} \right)^\theta, \quad \eta \in \langle -\zeta \rangle,$$

with $\beta \in \mathbb{Z}[\zeta]$ being an integral element that depends on x and θ. If $G_\theta \in \mathbb{K}[[T]]$ is the formal binomial series $G_\theta(T) = (1 + \zeta T)^{\theta/q}$, it is absolutely convergent for $|T| < 1$, in particular at $T = -1/x$. Fixing some primitive $q-$th root of unity $\xi \in \mathbb{C}$, there is thus a *galois exponent* $\kappa(\theta) \in \mathbb{Z}/(q \cdot \mathbb{Z})$, such that

$$\beta/\overline{\beta} = \eta \xi^{\kappa(\theta)} G_{\theta(1-\jmath)}(-1/x), \quad \text{for some } \eta \in \langle -\zeta \rangle.$$

The approach of our proof is essentially based on linear algebra of power series, and will be described in more detail in the introduction of section 3, after we recall some classical results in the next chapter. We let $\mathbb{K}' = \mathbb{Q}[\xi]$ and $\mathbb{L} = \mathbb{K}' \cdot \mathbb{K}$, the galois group being $H = \text{ Gal } (\mathbb{K}'/\mathbb{Q}) \cong \text{ Gal } (\mathbb{L}/\mathbb{K})$.

2 Cyclotomy, classical and elementary facts.

We assume in the sequel that (x, z) is a solution to equation (3), for the prime exponents $p < q$.

2.1 Characteristic numbers and characteristic ideals

Lemma 1. *Let* $\alpha = \frac{x-\zeta}{1-\zeta}$ *and* $\mathfrak{A} = (\alpha, z)$*: the* characteristic number and ideal*, respectively.*

1. *Then* $\alpha \in \mathbb{Z}[\zeta]; \alpha \equiv 1 \bmod \lambda^2$.

2. *The conjugates of the characteristic number are mutually coprime:*

$$(\sigma_c(\alpha), \sigma_d(\alpha)) = (1), \quad \textit{for } 1 \leq c < d \leq p-1$$

3. *The conjugates of the* characteristic ideal $\mathfrak{A}$ *are mutually coprime and*

$$\mathfrak{A}^q \;=\; (\alpha), \quad \textit{and} \quad \mathbf{N}_{\mathbb{K}/\mathbb{Q}}(\mathfrak{A}) = (z). \tag{5}$$

Proof. For point 1: since $p = \prod_{i=1}^{p-1}(1-\zeta^i)$ and $p|(x-1)$, $(1-\zeta)|(x-1+1-\zeta)$, so $\alpha = 1 + \frac{x-1}{1-\zeta} \equiv 1 \bmod \lambda^2$ is integral and verifies the claimed congruence.

For point 2, let $I(c,d) = (\sigma_c(\alpha), \sigma_d(\alpha)) = (\frac{x-\zeta^c}{1-\zeta^c}, \frac{x-\zeta^d}{1-\zeta^d})$. We note that for any integers k and j not divisible by p, the number $\frac{1-\zeta^k}{1-\zeta^j}$ is a cyclotomic unit of the field $\mathbb{K}$ – e.g. [11], Proposition 8.1. Therefore, $\frac{x-\zeta^c}{1-\zeta^c} \cdot \frac{1-\zeta^c}{1-\zeta^d} = \frac{x-\zeta^c}{1-\zeta^d} \in I(c,d)$. Then $\frac{\zeta^c-\zeta^d}{1-\zeta^d} \in I(c,d)$. The ideal $I(c,d)$ thus contains a unit, and it must be $I(c,d) = (1)$.

For point 3, $z^q = \prod_{c \in P^*} \sigma_c(\alpha)$ by (3), so $\alpha | z^q$, and $z^q/\alpha = \prod_{c \in P^*, \sigma_c \neq \sigma_{id}} \sigma_c(\alpha)$. Hence, by point 2, we know $(\alpha, z^q/\alpha) = (1)$. For the characteristic ideal, this implies:

$$\mathfrak{A}^q = \left(\alpha^q, \alpha^{q-1}z, \ldots, \alpha z^{q-1}, \alpha \cdot (z^q/\alpha)\right) = (\alpha) \cdot (\alpha^{q-1}, \ldots, z^{q-1}, z^q/\alpha) = (\alpha),$$

hence $\mathfrak{A}^q = (\alpha)$ which means that the characteristic ideal is either principal or has order q. The norm $\mathbf{N}(\mathfrak{A}) = \prod_{c \in P^*}(\alpha^{\sigma_c}, z) = (\mathbf{N}(\alpha), z, z^2, \ldots, z^{p-1})$ by point 2. Since $\mathbf{N}(\alpha) = z^q$, we get the relation $\mathbf{N}(\mathfrak{A}) = (z)$. □

2.2 The Stickelberger's ideal

As mentioned above, the Stickelberger ideal is a subideal of the group ring $\mathbb{Z}[G]$ with the property of annihilating the class group. It is defined as the intersection of a fractional, principal ideal, with $\mathbb{Z}[G]$, as follows:

Definition 1. *The* Stickelberger element *is defined by*

$$\vartheta = \frac{1}{p}\sum_{c=1}^{p-1} c\sigma_c^{-1} \in \frac{1}{p}\mathbb{Z}[G] \tag{6}$$

and Stickelberger's ideal is defined as

$$I = \vartheta\mathbb{Z}[G] \cap \mathbb{Z}[G]. \tag{7}$$

We note that $\mathbf{N} = \mathbf{N}_{\mathbb{K}/\mathbb{Q}} \in I$ *and* $I \subset (1-\jmath)\mathbb{Z}[G] \oplus (\mathbf{N})$. *There exists a base for* $I^- = (1-\jmath)I$ *made of* $(p-1)/2$ *elements, called* Fueter elements, *e.g. [6], which are*

$$\begin{aligned}
\psi_n &= \vartheta(1+\sigma_n-\sigma_{n+1}) \\
&= \sum_{c\in P^*} n_c\sigma_c^{-1} \in \mathbb{Z}_{\geq 0}[G], \text{ for } n \in \left\{1,2,\ldots,\frac{p-1}{2}\right\} \qquad (8)\\
\text{with} \quad n_c &= \left(\left\lfloor\frac{(n+1)c}{p}\right\rfloor - \left\lfloor\frac{nc}{p}\right\rfloor\right) \\
\text{and} \quad n_c + n_{p-c} &= 1, \qquad (9)
\end{aligned}$$

From $n_c + n_{p-c} = 1$ *and* $n_c \geq 0$, *we note* $n_c = 0$ *or* $n_c = 1$, *and* $(1+\jmath)\cdot\psi_n = \mathbf{N}_{\mathbb{K}/\mathbb{Q}}$, *which means for every ideal* $\mathfrak{P} \subset \mathbb{Z}[\zeta]$, $\mathfrak{P}^{(1+\jmath).\psi_n} = \mathbf{N}(\mathfrak{P})$. *Note also that*

$$\psi_{p-(n+1)} - \psi_n = \vartheta(1+\jmath\sigma_{n+1}-\jmath\sigma_n-1-\sigma_n+\sigma_{n+1}) = \mathbf{N}-\mathbf{N} = 0, \tag{10}$$

which confirms the choice of indices $n \leq \frac{p-1}{2}$ *for the base of Fueter elements.*

We define $I_+ := \vartheta\mathbb{Z}[G] \cap \mathbb{Z}_+[G]$. *Combining the property of Stickelberger's ideal and the property of Fueter elements, we note that for every ideal* $\mathfrak{P} \subset \mathbb{Z}[\zeta]$ *and each* $\theta \in I_+$, *the ideal* $\mathfrak{P}^\theta \subset \mathbb{Z}[\zeta]$ *is generated by some* $\gamma \in \mathbb{Z}[\zeta]$, *which satisfies* $\gamma\cdot\overline{\gamma} = \mathbf{N}(\mathfrak{P})^{\varsigma_\theta}$, *for an integer* $\varsigma_\theta \in \mathbb{Z}$, *which we call the* relative weight *of* θ. *Note that the relative weight of each Fueter element is* 1. *Furthermore, we denote by* the absolute weight *of* $\theta = \sum_{c\in P^*} n_c\sigma_c^{-1} \in \mathbb{Z}[G]$ *the sum* $w(\theta) = \sum_c |n_c|$.

We define the Fermat quotient map $\phi : \mathbb{Z}[G] \to \mathbb{F}_p$ *such that* $\zeta^\theta = \zeta^{\phi(\theta)}$. *Explicitly,*

$$\phi\left(\sum_{c\in P^*} n_c\sigma_c^{-1}\right) = \sum_{c\in P^*} n_c/c \in \mathbb{F}_p. \tag{11}$$

We identify the value $\phi(\theta) \in \mathbb{F}_p$ *with its natural lift to* $\mathbb{N}$, *under the least non-negative remainder representation of* $\mathbb{F}_p$. *The* Fermat ideal *is* $I_0 = I \cap Ker\,(\phi)$: *this is the module of all Stickelberger elements* θ *such that* $\zeta^\theta = 1$.

We note the following useful property of the action of I on λ:

Lemma 2. *Let $M_\mu := (1-\zeta)^\mu$, for some $\mu \in I$, then $M_\mu = \pi^{\varsigma(\mu)} \cdot (\zeta)^{\phi(\mu)/2}$, where $\mathbb{Q}(\pi) \subset \mathbb{K}$ is the quadratic subfield of $\mathbb{K}$ and $\pi^2 = p \cdot \left(\frac{-1}{p}\right)$.*[3] *Moreover, $M_\mu^{\sigma-1} = \pm\zeta^{(\sigma-1)\phi(\mu)/2} \in \mu_{2p}$, for $\sigma \in G$.*

Proof. Since $\mathbf{N}_{\mathbb{K}/\mathbb{Q}}(1-\zeta) = p$, then $M_\mu \cdot \overline{M_\mu} = (1-\zeta)^{\mu(1+\jmath)} = p^{\varsigma(\mu)}$. Furthermore, $M_\mu/\overline{M_\mu} = (\frac{1-\zeta}{1-\zeta})^\mu = (-\zeta)^\mu$ and from the definition of the Fermat quotient map ϕ, $\zeta^\theta = \zeta^{\phi(\theta)}$, then it is equal to $\left(\frac{-1}{p}\right)^{\varsigma(\mu)} \zeta^{\phi(\mu)}$, so $M_\mu^2 = p^{\varsigma(\mu)} \cdot \left(\frac{-1}{p}\right)^{\varsigma(\mu)} \cdot \zeta^{\phi(\mu)}$. Hence, $M_\mu = \pi^{\varsigma(\mu)} \cdot \zeta^{\phi(\mu)/2}$. The last statement follows from $\sigma_c(\pi) = \left(\frac{c}{p}\right)\pi$. □

2.2.1 The β_0-map

Stickelberger's theorem implies that for $\theta \in I$, there is a principal ideal $\mathfrak{a}(\theta) = \mathfrak{A}^\theta$, and the principal ideal arising from the action of the elements of the Stickelberger ideal is generated by Jacobi numbers, which are products of Jacobi sums[6]. The product of Jacobi numbers by their complex conjugates is a rational integer, which is equal to the norm of $\mathfrak{A}$ raised to the relative weight of θ , i.e. $|\mathbb{Z}[\zeta]/\mathfrak{A}|^{\varsigma_\theta} = |z|^{\varsigma_\theta}$.

Remark 1. *Let $r \equiv 1 \pmod p$ be a rational prime which is totally split in $\mathbb{K}$, and let $\mathfrak{R} \subset \mathbb{Z}[\zeta]$ be a prime above r. By the above, for $\theta \in I_+$ with relative weight $\varsigma(\theta) = 1$, we have*

$$\mathfrak{R}^\theta = (\beta_r), \qquad \beta_r \in \mathbb{Z}[\zeta], \qquad \mathbf{N}_{\mathbb{K}/\mathbb{Q}}(\mathfrak{R}) = (r), \qquad \beta_r\bar{\beta}_r = r.$$

Moreover, β_r is a product of Jacobi sums of conductor r, and its norm implies that it is in fact a single Jacobi sum $j(\chi^a, \chi^b)$ for suitable a, b.

More precisely, if

$$\theta_{a,b} := \sum_{c=1}^{p-1} \left(\left\lfloor \frac{c(a+b)}{p} \right\rfloor - \left\lfloor \frac{ca}{p} \right\rfloor - \left\lfloor \frac{cb}{p} \right\rfloor \right) \sigma_c^{-1},$$

then one has

$$\big(j(\chi^a, \chi^b)\big) = \mathfrak{R}^{\theta_{a,b}}.$$

Iwasawa proved in [5] that every Jacobi number $\mathbf{J}$ verifies

$$\mathbf{J} \equiv 1 \bmod (1-\zeta)^2, \tag{12}$$

Let $\beta_0 \in \mathbb{Z}[\zeta]$ be the Jacobi number that generates the ideal. By the identity $(\alpha^\theta) = \mathfrak{A}^{q\theta} = \mathfrak{a}(\theta)^q$, we obtain

$$\alpha^\theta = \eta \cdot \beta_0^q, \tag{13}$$

[3] Here $\left(\frac{\cdot}{p}\right)$ denotes the Legendre symbol modulo p.

where η is a unit in $\mathbb{Z}[\zeta]$. Multiplying their complex conjugates on both sides of equation (13) and since

$$\alpha^{\theta(1+\jmath)} = \mathbf{N}_{\mathbb{K}/\mathbb{Q}}(\alpha)^{\varsigma_\theta} = z^{q\varsigma_\theta} = \beta_0^{q(1+\jmath)}$$

we get $\eta \cdot \bar{\eta} = 1$. It follows from Kronecker's Unit Theorem that $\eta \in \mu_{2p}$. Combined with the congruence in Point 1. of Lemma 1 and (12), we obtain that in fact $\eta = 1$. Hence $\alpha^\theta = \beta_0^q$.

We herewith define the maps $\beta_0 : I_+ \to \mathbb{Z}[\zeta]$ as follows. For $\theta \in I_+$, we let $\beta_0(\theta)$ be the unique Jacobi number defined above by the identity $\alpha^\theta = \beta_0^q$.

2.3 Formal binomial power series

This section basically follows the original paper proving the Catalan conjecture in Part 4.1 [8], except for the subscripts, which are slightly different. Therefore, some proofs will be omitted here.

2.3.1 Definition

The identity $\alpha^\theta = \beta_0(\theta)^q$ leads to a binomial series expansion which converges absolutely. In this section, we consider this series in detail.

For $\mu = \sum_{c \in P^*} n_c \sigma_c^{-1} \in \mathbb{Z}[G]$, we let $G_\mu(T) \in \mathbb{K}[[T]]$ be the formal binomial power series

$$\begin{aligned} G_\mu(T) &= (1+\zeta T)^{\mu/q} = \prod_{c \in P^*} (1+\sigma_{1/c}(\zeta)T)^{n_c/q} \\ &= \prod_{c \in P^*} \left(1 + \sum_{n \geq 1} \binom{n_c/q}{n} (\sigma_{1/c}(\zeta)T)^n\right) =: 1 + \sum_{n \geq 1} a'_n(\mu) T^n, \quad (14) \end{aligned}$$

and the coefficients $a'_n(\mu)$ arise from the multiplication and rearrangement of the terms in the product in (14). In particular, denoting $\eta_\mu := q \cdot a'_1 = \sum_{c \in P^*} n_c \zeta^{\sigma_{1/c}}$, we obtain

$$G_\mu(T) = 1 + \eta_\mu T/q + \sum_{n \geq 2} a'_n(\mu) T^n. \quad (15)$$

2.3.2 Coefficients

We investigate some properties of these coefficients on base of the coefficients in the binomial series. Note that $a'_n(\mu)^\sigma = a'_n(\mu\sigma), \quad (n = 0, 1, 2, \ldots)$, by definition, as formal power series. In addition, the common domain of convergence in $\mathbb{C}$ is $D = \{|T| < 1\} \subset \mathbb{C}$, so the evaluations of the two power series are equal in this domain.

Lemma 3. *The coefficients $a'_n(\mu) \in \mathbb{Z}[\zeta, 1/q]$ are q-integers, which means $a_n := a'_n \cdot q^k \in \mathbb{Z}[\zeta]$ for some k. More precisely, $k \geq E(n) = n + v_q(n!)$. In particular, $E(n) = n + v_q(n!) < n \cdot \frac{q}{q-1}$.*

See [8] for a proof.

Now, we investigate upper bounds for the coefficients. A power series $f(T)=\sum_{k=0}^{\infty} a'_k T^k$ with complex coefficients is *dominated* by the series $g(T)=\sum_{k=0}^{\infty} A_k T^k$ with non-negative real coefficients if $|a'_k| \leq A_k$ holds for $k=0,1,\ldots$; if this is the case, we write $g \gg f$. The relation of dominance is preserved by the addition and multiplication of power series.

Let r be a real number and s a complex number satisfying $|s| \leq 1$. Then, for the binomial series, we have

$$(1+sT)^r = \sum_{k=0}^{\infty} \binom{r}{k} s^k T^k$$
$$(1-T)^{-|r|} = \sum_{k=0}^{\infty} (-1)^k \binom{-|r|}{k} T^k.$$

The coefficients of the latter series are positive and $\left|\binom{r}{k}\right| \leq \left|\binom{-|r|}{k}\right|$. Since $|\zeta|=1$, it follows that $G_\mu(T) \ll (1-T)^{-h/q}$ where h is the absolute weight of μ. Combining with the previous result, we obtain the following bounds:

Lemma 4. *Let $a_n = q^{E(n)} a'_n$, where a'_n is the coefficients of $G_\mu(T)$ and $E(n) = n + v_q(n!)$, h is the absolute weight of μ, $h' = \lceil h/q \rceil$ Then*

$$|a'_n| \leq (-1)^n \binom{-h/q}{n} < \binom{h'+n-1}{n},$$

$$|a_n| \leq q^{E(n)} (-1)^n \binom{-h/q}{n} < \binom{h'+n-1}{n} q^{E(n)},$$

Finally, we consider the convergence of these binomial power series.

Lemma 5. *The series $G_\mu(t)$ are absolutely convergent in $\mathbb{C}$ when $|t|<1$.*

Proof. It follows that $G_\mu(t)$ is dominated by $(1-T)^{-h/q}$. For every non-negative integer m, we define the m -th partial sum by

$$G_\mu(t)|_m = 1 + \sum_{n=1}^{m} a'_n(\mu) t^n.$$

Using the common remainder estimates for Taylor series, the remainder $|R_m|$ is given by

$$|G_\mu(t) - G_\mu(t)|_m| \leq \binom{h'+m}{m+1} \cdot \frac{|t|^{m+1}}{(1-|t|)^{h'+m+1}},$$

as a consequence of Lemma 4. □

Remark 2. *Let $\theta = \theta_1 + \theta_2 \in I$ have non-negative coefficients in $\mathbb{Z}$ and assume that $\varsigma(\theta_1) = \varsigma(\theta_2) = k \geq 1$. We shall write in this particular case, which is the most frequent in the sequel,*

$$G_\theta(T) = \frac{(1+\zeta T)^{\theta_1/q}}{(1+\bar{\zeta}T)^{\theta_2/q}}. \tag{16}$$

With this, we note that

$$\beta(\theta) = z^k \cdot \eta \cdot \xi^{\kappa(\theta)} G_\theta(-1/x).$$

*Here $\eta = \eta(\theta) \in \langle\zeta\rangle$. We override the initial definition by twisting β by an adequate $p-$*th *root of unity – namely $\bar{\eta}$, so that we now have*

$$\beta(\theta) = z^k \cdot \xi^{\kappa(\theta)} G_\theta(-1/x);\ \beta \cdot \bar{\beta} = z^{2k}. \tag{17}$$

Let $\theta_i = \sum_{c \in P^*} n_{i,c} \sigma_c^{-1}$. We note that the value of η_μ in (15), for $\mu = \theta_1 - \jmath\theta_2$ is

$$\eta_\mu = \sum_{c \in P^*} (n_{1,c} - n_{2,p-c}) \zeta^{1/c}. \tag{18}$$

Consequently, the coefficient vector of η_μ in the normal power base of $\mathbb{Z}[\zeta]$ verifies

$$||\eta_\mu||_\infty \leq k = \varsigma(\theta_i). \tag{19}$$

3 Main proofs

Now, we assume $53 < p < q$. Let $I'(m) \subset I$ be the set of sums of m conjugates of Fueter elements, thus $\theta \in I'(m)$ if $\theta = \sum_{c \in P^*} n_c \sigma_c^{-1}$ with $n_c + n_{p-c} = m$ and $n_c \geq 0$; then

$$I'(m) \cap (1+\jmath)\mathbb{Z}[G] = (m/2)(\mathbf{N}) \quad \text{for } m \equiv 0 \bmod 2 \text{ and } \emptyset \text{ otherwise.}$$

We thus let $I(m) = I'(m)/(I'(m) \cap (\mathbf{N}))$ be the relevant minus part of $I'(m)$. Then the Fueter elements verify

$$\psi_n = \psi_{p-(n+1)} = \vartheta(\sigma_1 + \sigma_n + \sigma_{-(n+1)}) - \mathbf{N}, \quad 1 \leq n \leq \frac{p-1}{2}. \tag{20}$$

We shall be interested in an estimate of the number of *distinct* elements in $I(2)$. The precise number of distinct elements in $I(2)$ is an involved combinatorial problem, solved in subsection 3.1; for the solution of this problem, we made interesting use of AI, which are detailed in the Outlook.

In the main Proposition 3 we show that if $J \subset I(2)$ is such that $|\kappa(J)| = 1$, then the size of J is upper bounded by $|J| < \frac{2(p-1)}{3} + 1$. An application of the pigeonhole principle to the formula (31) for distinct elements in $I(2)$ implies that there are subsets $J \subset I(2)$ with $|J| > p$, say. This contradicts the Proposition 3 and completes the proof.

3.1 Estimating the number of distinct elements in $I(n)$

For $t \in J$, let $\beta(t), G_t(-1/x)$ have the usual meaning introduced in Remark 2. We shall index various vectors by the elements in J, for instance $(\beta(t))_{t\in J}$.

Definition 2. *We say a pair $(\sigma,\tau) \in G^2$ is* complementary*, if $\tau = \jmath\sigma$. Accordingly, (σ_a, σ_b) is complementary, if $a + b \equiv 0 \bmod p$. Note that complementary pairs (σ_a, σ_b) have the property that $\vartheta(\sigma_a + \sigma_b) = \mathbf{N}$.*

For our purpose, it will be sufficient to find estimates for $n \leq 2$. The next simple lemmata will explain the approach taken.

Lemma 6. *Let $\theta \in I(1)$. Then there is a $\sigma \in G$ and $n \in P^*, n \leq \frac{p-1}{2}$ such that $\theta = \sigma\psi_n$, with $\psi_n = \vartheta(1 + \sigma_n - \sigma_{n+1})$ the respective Fueter element.*

Proof. In view of Remark 1, letting $\mathfrak{R}$ be like in the Remark and fixing a primitive character $\chi : \mathbb{F}_r^\times \to \langle\zeta\rangle$, we have a one-to-one correspondence between $\theta \in I(1)$ and Jacobi sums of conductor r with $|j(\chi^a, \chi^b)| = r^{\varsigma(\theta)/2}$. Then

$$\mathfrak{R}^{\sigma_a^{-1}\theta} = j(\chi, \chi^{b/a}) = \mathfrak{R}^{\psi_{b/a}}.$$

This holds for all totally split primes r and we conclude that $\theta = \sigma_a\psi_{b/a}$. □

We shall recalibrate elements $\theta \in I(m)$ as follows; since

$$\psi_n = \vartheta(1 + \sigma_n - \sigma_{n+1}) = \vartheta(1 + \sigma_n + \sigma_{-(n+1)}) - \mathbf{N},$$

we conclude that

$$\theta = \vartheta \cdot S_\theta - m\mathbf{N}, \quad \text{with } S_\theta \in \mathbb{Z}_{\geq 0}[G]. \tag{21}$$

Lemma 7. *Let $m \in \mathbb{Z}_{>0}$ and $\theta_i = \vartheta S_{\theta_i} - m\mathbf{N} \in I(m), i = 1, 2$, with S_θ like above. Then $\theta_1 = \theta_2$ if and only if there exists an integer $n \geq 0$ such that, after deleting exactly the same number n of complementary pairs $\{\sigma_j, \sigma_{p-j}\}$ from S_{θ_1} and from S_{θ_2} (the specific j's may differ), the reduced sums must coincide.*

Proof. Define the difference

$$\Delta := S_{\theta_1} - S_{\theta_2} = \sum_{k\in P^*} v(k)\sigma_k.$$

We shall use the spectral decomposition in orthogonal idempotents over $\mathbb{C}$ in order to prove that Δ only depends on the odd and the trivial characters. This implies the claim.

Let $\mu : G \to \mathbb{C}$ be a primitive character of order $p-1$ and $\chi \in \langle\mu\rangle$. By classical results in cyclotomy – see for instance [11], subsection 6.3 for local versions – we

have the following decomposition of unity and relation between the Stickelberger element and the generalized Bernoulli numbers:

$$\begin{aligned} e_\chi &= \frac{1}{|G|}\sum_{g\in G}\chi(g^{-1})g, \quad e_\chi\sigma_k = \chi(k)e_\chi; \\ 1 &= \sum_\chi e_\chi; \quad \text{and} \\ \vartheta e_\chi &= B_{1,\bar\chi}e_\chi, \end{aligned} \tag{22}$$

where $B_{1,\bar\chi}$ are the generalized Bernoulli numbers – [11] p. 31-32. In particular, $\vartheta e_\chi = 0$ for even characters except for the trivial one for which $\vartheta e_\chi = \frac{1}{2}\cdot\mathbf{N}$.

We apply these facts to $\vartheta\Delta$, finding

$$\vartheta\Delta = \sum_\chi e_\chi B_{1,\bar\chi}\Delta = \sum_{\chi \text{ odd}} e_\chi B_{1,\bar\chi}\Delta + \frac{1}{2}\mathbf{N}\Delta.$$

Plainly, $\vartheta\Delta = 0$ if and only if $e_\chi\Delta = 0$ for all odd characters, and in addition $\vartheta(1+\jmath)\Delta = 0$: indeed, the contribution of even characters reduces to that of the trivial character, which vanishes if and only if $\vartheta(1+\jmath)\Delta = 0$. The last relation implies Δ vanishes in all odd eigenspaces, and $\theta_1 = \theta_2$, after eventually removing on both sides the same number of complementary pairs. Removal of the complementary pairs thus cleans the plus part contribution in both elements, reducing them to $\theta_i \in I^-; i = 1, 2$. This completes the proof. □

3.1.1 Counting the elements in $I(1)$

Given the limited size of the Latin alphabet and of printed indices, we shall use only in this subsection the letters x, y as indices – thus with a deeply different meaning from the main variables x, y connected by the Catalan equation. Suppose $\sigma_y\psi_x = \sigma_a\psi_b$, where $x, b \in \{1, 2, \ldots, \frac{p-1}{2}\}$ and $y, a \in P^*$. Then, the norms cancel in the representation (21) and thus:

$$\vartheta(\sigma_y + \sigma_{yx} + \sigma_{-y(x+1)}) = \vartheta(\sigma_a + \sigma_{ab} + \sigma_{-a(b+1)}).$$

One verifies, using the ranges of x, y, that none of the three possible pairs in the triple $\{\sigma_y, \sigma_{yx}, \sigma_{-y(x+1)}\}$ is a complementary pair. Hence, the two sets

$$\{\sigma_y, \sigma_{yx}, \sigma_{-y(x+1)}\} \quad \text{and} \quad \{\sigma_a, \sigma_{ab}, \sigma_{-a(b+1)}\}$$

must coincide. After investigating all permutations, we see that only three describe elements in $I(1)$, and these are

$$\theta = \sigma_y\psi_x = \sigma_{yx}\psi_{1/x} = \sigma_{-y(x+1)}\psi_{-1/(x+1)}, \tag{23}$$

with all subscripts taken in $\mathbb{F}_p^\times$. If $1/x \notin \{1, 2, \ldots, \frac{p-1}{2}\}$, we replace $1/x$ by $p - 1 - 1/x$; the same adjustment applies to $-1/(x+1)$. Consequently, each $\theta \in I(1)$ can be expressed in at most three ways, as a conjugate of some Fueter elements.

Lemma 8. *For every element $\sigma_y\psi_x \in I(1)$, there are three distinct expressions as above, except when $x = 1$ or $x = \frac{p-1}{2}$.*

Proof. We have already observed that each $\theta \in I(1)$ admits at most three expressions as a conjugate of a Fueter element. If fewer than three occur, then the subscripts in one expression coincide with those in another one. A short verification shows that this may happen precisely when $x = 1$ or $x = \frac{p-1}{2}$, and in both cases there are exactly two expressions. Moreover, for all $y \in G$ one has $\sigma_y\psi_1 = \sigma_{-2y}\psi_{(p-1)/2}$ and the triple $\{\sigma_y, \sigma_{yx}, \sigma_{-y(x+1)}\}$ has two identical elements when $x = 1$ or $x = \frac{p-1}{2}$. □

We now count the size of $I(1)$. There are

$$M = (p-1)\cdot\frac{p-1}{2} = \frac{(p-1)^2}{2}$$

pairs (y, x). By Lemma 8, the map $\mathfrak{E} : P^* \times P^* \to I(1)$ given by $(y, x) \mapsto \sigma_y\psi_x$ is threefold for $x \notin \{1, \frac{p-1}{2}\}$ and twofold otherwise. Therefore,

$$|I(1)| = M - \frac{2}{3}\left(\frac{p-1}{2} - 2\right)(p-1) - (p-1) = \frac{p^2-1}{6}. \tag{24}$$

We herewith have proved:

Lemma 9. *For any $\theta \in I(1) = \vartheta S_\theta - \mathbf{N}$, S_θ is invariant under the transformations $(y, x) \mapsto (yx, 1/x)$, $(y, x) \mapsto (-y(x+1), -1/(x+1))$ and $(y, x) \mapsto (y, p-1-x)$. Hence, $\theta_1 = \theta_2 \iff S_{\theta_1} = S_{\theta_2}$. Moreover, the number of elements in $I(1)$ is given by (24).*

3.1.2 Counting the elements in $I(2)$

The elements of $I(2)$ are conjugates of some element of the form $\psi_a + \sigma_i\psi_b$. We investigate when two elements of $I(2)$ are identical. Assume $\sigma_r\psi_a + \sigma_s\psi_b = \sigma_x\psi_c + \sigma_y\psi_d$, so that

$$\begin{aligned} &\vartheta\big(\sigma_r + \sigma_{ra} + \sigma_{-r(a+1)} + \sigma_s + \sigma_{sb} + \sigma_{-s(b+1)}\big) \\ &= \vartheta\big(\sigma_x + \sigma_{xc} + \sigma_{-x(c+1)} + \sigma_y + \sigma_{yd} + \sigma_{-y(d+1)}\big) \quad . \end{aligned} \tag{25}$$

Let S_1, S_2 denote the cofactors of ϑ on the left and on the right hand side in the above identity. By Lemma 7, we have $S_1 - S_2 \in (1+\jmath)\,\mathbb{Z}[G]$.

If $S_1 = S_2$, then we focus on the first three terms $S_1' = \sigma_r + \sigma_{ra} + \sigma_{-r(a+1)}$ in S_1: they must correspond to three terms in S_2, which we also split in $S_2' = \sigma_x + \sigma_{xc} + \sigma_{-x(c+1)}$ and $S_2'' = \sigma_y + \sigma_{yd} + \sigma_{-y(d+1)}$. If two of the terms of S_1' are in S_2', then so must be the third, because the sum of the indices in S_1' adds up to 0, and the same must hold for S_1' in relation to S_2. So in this case, we have $\sigma_r\psi_a = \sigma_x\psi_c$. If however, two of the terms of S_1' are in S_2'', the same argument shows that $\sigma_r\psi_a = \sigma_y\psi_d$, so in the case $S_1 = S_2$, the two terms on the left and right hand side are identical.

We claim that $S_i; i = 1, 2$ contain at most one pair of complementary elements. Indeed, if there are two complementary pairs in S_1 and S_2, then S_1 must contain three complementary pairs, because the sum of the subscripts of these six terms is zero. Hence $\sigma_r\psi_a + \sigma_s\psi_b = \mathbf{N}$. This confirms the claim.

The sums S_1' and $S_1'' = \sigma_s + \sigma_{sb} + \sigma_{-s(b+1)}$ contain no pairs of complementary elements, so assuming there is one such pair in S_1, it is built of one term in each of the two partial sums. Choose one element from $\Sigma_1' = \{\sigma_r, \sigma_{ra}, \sigma_{-r(a+1)}\}$ and one element from $\Sigma_1'' = \{\sigma_s, \sigma_{sb}, \sigma_{-s(b+1)}\}$ so that these two form a complementary pair. Using Lemma 8 and (23), one verifies with a case by case investigation, that any choice $(u, v) \in \Sigma_1' \times \Sigma_1''$ as a complementary pair leads, under application of the three representations in (23) to one and the same solution, by redefining the indices. We conclude:

Lemma 10. *Notations being like above, if $S_1 = S_2$, the two sums are equal termwise. Otherwise, there is at most one pair of complementary elements in S_1, and for each possible choice of this pair $(u, v) \in \Sigma_1' \times \Sigma_1''$, we obtain various representations of one and the same element in $I(2)$. The same holds for the right hand side.*

The next lemma investigates the various ways in which the sum $\sigma_r\psi_a + \sigma_s\psi_b$ can be represented as a sum of two elements in $I(1)$ – thus the various possibilities for S_2 – in the case when there is one complementary pair in S_1. By Lemma 10, we can choose complementary pairs on both sides arbitrarily. For convenience, apply σ_r^{-1} to both sides of the equality in the assumption prior to (25). Setting $i := r^{-1}s$, $j := r^{-1}x$ and $k := r^{-1}y$, we obtain $\psi_a + \sigma_i\psi_b = \sigma_j\psi_c + \sigma_k\psi_d$. We deduce in the following all the possible representations of $\theta = \psi_a + \sigma_i\psi_b$, as a sum of two elements in $I(1)$, under the additional assumption that Σ_1 contains a pair of complementary elements.

Lemma 11. *Let $a, b \in \left\{1, 2, \ldots, p-2\right\}$. Define $i := \frac{a}{b+1}$. Then the following identity holds:*

$$\psi_a + \sigma_i\psi_b = \sigma_{-(i+1)}\psi_{-\frac{b+1}{a+b+1}} + \sigma_{i+1}\,\psi_{\frac{ab}{a+b+1}} = \sigma_{-ib-1}\psi_{-\frac{b+1}{ab+b+1}} + \sigma_{ib+1}\psi_{\frac{a}{ab+b+1}}, \quad (26)$$

with all subscripts taken modulo p. Moreover, these are the only possible representations of $\theta = \psi_a + \sigma_i\psi_b$ as sum of two elements in $I(1)$, with a pair of complementary elements in $\Sigma' \times \Sigma''$.

Proof. We assume that (25) holds and use the notation introduced above. The choice of a complementary pair on each side in (25) leaves out on either side four additional terms which are identical, as a set. In view of Lemma 10, we may choose any pair of complementary values in $\Sigma' \times \Sigma_1''$. Let the complementary pairs be $(\sigma_a, \sigma_{-i(b+1)})$ and (σ_j, σ_k). We are left to compare two sets of 4 elements

If $\{\sigma_{jc}, \sigma_{-j(c+1)}\} = \{\sigma_1, \sigma_{-(a+1)}\}$, then it follows that $\sigma_a = \sigma_j$, by completing the triples belonging to the respective elements in $I(1)$, hence $\psi_a = \sigma_j\psi_c$. It also holds that if $\{\sigma_{kd}, \sigma_{-k(d+1)}\} = \{\sigma_1, \sigma_{-(a+1)}\}$ then $\psi_a = \sigma_k\psi_d$, for the same reason.

If $\{\sigma_{jc}, \sigma_{-j(c+1)}\} = \{\sigma_1, \sigma_i\}$, then, by completing the triples, we obtain $\sigma_j = \sigma_{-(1+i)}$. Consequently,

$$c = \begin{cases} -\dfrac{b+1}{a+b+1}, & \text{if } jc = 1, \\ -\dfrac{a}{a+b+1}, & \text{if } jc = i. \end{cases}$$

Since $\psi_x = \psi_{p-1-x}$, it follows that

$$\psi_{-\frac{b+1}{a+b+1}} = \psi_{-\frac{a}{a+b+1}}.$$

Moreover, because (σ_j, σ_k) is a complementary pair, we have $k = 1 + i$ and the remaining pair forces $kd = ib$. Therefore,

$$d = \frac{ib}{k} = \frac{ib}{1+i} = \frac{\frac{a}{b+1}\,b}{\frac{a+b+1}{b+1}} = \frac{ab}{a+b+1}.$$

Hence, this yields precisely the second identity in (26).

If instead $\{\sigma_{kd}, \sigma_{-k(d+1)}\} = \{\sigma_1, \sigma_i\}$, the same identity follows upon interchanging the summand:

$$\sigma_{-(i+1)}\psi_{-\frac{b+1}{a+b+1}} \;+\; \sigma_{i+1}\psi_{\frac{ab}{a+b+1}} \;=\; \sigma_{i+1}\psi_{\frac{ab}{a+b+1}} \;+\; \sigma_{-(i+1)}\psi_{-\frac{b+1}{a+b+1}}.$$

The remaining case $\{\sigma_{jc}, \sigma_{-j(c+1)}\} = \{\sigma_1, \sigma_{ib}\}$ or $\{\sigma_{kd}, \sigma_{-k(d+1)}\} = \{\sigma_1, \sigma_{ib}\}$ is treated analogously and gives the third identity in (26). □

We now estimate the size of $I(2)$. Without accounting for duplications, there are $M = \binom{\frac{p^2-1}{6}+1}{2}$ pairs, with repetition, of elements in $I(1)$, as follows from (24). First, exactly $U := \frac{p^2-1}{12} = |I(1)|/2$ of these add up to $\mathbf{N}$: this happens precisely for pairs of conjugate elements in $I(1)$, and each pair is counting only once; the total number is thus equal to $|I(1)/\{1, \jmath\}|$.

The remaining duplications are governed by Lemma 11. From now on we thus require $\psi_a + \sigma_i\psi_b \neq \mathbf{N}$. Assume first that the three representations in Lemma 11 are pairwise distinct. Then the six terms in (26) are mutually distinct. Under these assumptions, for $a, b \in \Big\{1, 2, \ldots, p-2\Big\}$,

$$b \notin \Big\{1,\ a-1,\ (a-1)^{-1},\ \tfrac{-(2a+1)}{a+1},\ \tfrac{-(a+1)}{2a+1},\ -(a+1)^{-1},\ -(a+1)\Big\} \quad \text{and} \quad a \neq p-2$$

the entries of

$$L := \{\,1,\ i,\ ib,\ -(a+1)\,\}, \qquad i := \frac{a}{b+1},$$

are pairwise distinct, and moreover $-1, 0 \notin L$.

We claim that the G-orbit of $\psi_a + \sigma_i\psi_b$ has $p-1$ distinct elements. Suppose that there exists $\sigma_r \neq \sigma_1$ such that

$$\sigma_r\big(\psi_a + \sigma_i\psi_b\big) = \psi_a + \sigma_i\psi_b,$$

then $\{1, -(a+1), i, ib\} = \{r, -r(a+1), ri, rib\}$. Taking the product of the four elements on each side yields $r^4 \equiv 1 \bmod p$. Hence, $r \equiv -1 \bmod p$ or $r^2 \equiv -1 \bmod p$. If $r \equiv -1 \bmod p$, then

$$2(\psi_a + \sigma_i\psi_b) = (\sigma_1 + \sigma_{-1})(\psi_a + \sigma_i\psi_b) = (\sigma_1 + \sigma_{-1})\psi_a + (\sigma_1 + \sigma_{-1})\sigma_i\psi_b = 2\mathbf{N},$$

hence $\psi_a + \sigma_i\psi_b = \mathbf{N}$, contradicting our assumption. If $r^2 \equiv -1 \bmod p$, the set identity

$$\{1, -(a+1), i, ib\} = \{r, -r(a+1), ri, rib\}$$

implies that $\{1, -(a+1), i, ib\} = \{r^2, -r^2(a+1), r^2 i, r^2 ib\} = \{-1, (a+1), -i, -ib\}$, so σ_{-1} stabilizes $\psi_a + \sigma_i\psi_b$, again a contradiction. Therefore, for all $\sigma_j \in G \setminus \sigma_1$,

$$\sigma_j\big(\psi_a + \sigma_i\psi_b\big) \neq \psi_a + \sigma_i\psi_b. \tag{27}$$

For counting purposes, we therefore declare two ordered pairs (a, b) and (a', b') to represent the same G-orbit in $I(2)$, if there exists $r \in P^*$ such that

$$\big\{\sigma_1,\ \sigma_i,\ \sigma_{ib},\ \sigma_{-(a+1)}\big\} = \big\{\sigma_r,\ \sigma_{ri'},\ \sigma_{ri'b'},\ \sigma_{-r(a'+1)}\big\}, \qquad i' := \frac{a'}{b'+1}. \tag{28}$$

The counting problem reduces to a combinatorial orbit problem as follows. We consider the *scaling action* of $\mathbb{F}_p^\times$ on 4-sets by

$$r \cdot \{a, b, c, d\} := \{ra, rb, rc, rd\} \qquad (r \in \mathbb{F}_p^\times).$$

For $x, y \in D := \mathbb{F}_p \setminus \{-1, 0, 1\}; r \in \mathbb{F}_p^*$, we let $Q(x, y; r) = \{r, rx, ry, -r(1 + x + y)\}$ and let $\mathfrak{Q}$ be the set of all possible 4-sets arising in this way. The scaling by r defines an equivalence relation on $\mathfrak{Q}$, and we want to determine the number of equivalence classes. Equivalently, we are counting the $\mathbb{F}_p^\times$-orbits of the sets $Q(x, y)$.

Definition 3. *A representative $T = \{u, x, y, z\} \in \mathfrak{Q}$ is called* normalized *if $1 \in T$.*

Lemma 12. *Fix (x, y) as above and let $z = -(1 + x + y)$. Then the only scalars $r \in \mathbb{F}_p^\times$ that send the orbit of $Q(x, y; 1)$ to a normalized representative are*

$$r \in V := \Big\{1,\ \frac{1}{x},\ \frac{1}{y},\ \frac{1}{z}\Big\},$$

and these four choices are distinct. Consequently, each orbit contains exactly four normalized representatives. Moreover, these four representatives are mutually distinct.

Proof. If $Q(x, y; r) \in \mathfrak{Q}$ is normalized, then $1 \in Q(x, y; r) = \{r, rx, ry, rz\}$. Thus $1 \in \{r, rx, ry, rz\}$, which forces $r \in \{1, 1/x, 1/y, 1/z\}$. Since $x, y, z \in D$ are pairwise distinct and not in $\{-1, 0, 1\}$, these four scalars are distinct. Hence, there are exactly four normalized representatives in the orbit.

For the second statement, consider that for two distinct $r, r' \in \{1, 1/x, 1/y, 1/z\}$ we have

$$\{r, rx, ry, -r(1+x+y)\} = \{r', r'x, r'y, -r'(1+x+y)\}. \tag{29}$$

By dividing by $r' \neq 0$ on both sides and setting $s = r/r'$ we obtain the identity

$$\{s, sx, sy, -s(1+x+y)\} = \{1, x, y, -(1+x+y)\},$$

in which the right hand side is normalized. The element 1 on the right must equal one of s, sx, sy, sz, thus $s \in V$. We have thus reduced the problem to the case in which $r' = 1$ in (29), so

$$\{1, x, y, -(1+x+y)\} = \{r, rx, ry, -r(1+x+y)\}.$$

In view of (27), the G-orbit of $\psi_a + \sigma_i \psi_b$ has $p-1$ distinct elements, so the above identity cannot hold under the conditions on a, b. □

As a consequence, we find:

Proposition 1. *There are* $\dfrac{p^2 - 12p + 35}{24}$ *equivalence classes of* $\{1, x, y, z\}$. *Thus,* $W := \dfrac{p^2 - 12p + 35}{24}(p-1)$ *elements of* $I(2)$ *admit three distinct expressions.*

Proof. Count ordered triples $(x, y, z) \in \mathbb{F}_p^3$ with $x + y + z + 1 = 0$. Since z is determined by $z = -1 - x - y$, it suffices to count ordered pairs $(x, y) \in \mathbb{F}_p^2$. Let $T := \{-1, 0, 1\} \subset \mathbb{F}_p$ and $D := \mathbb{F}_p \setminus T$ ($p \geq 5$).

Imposing $x, y, z \in D$ and applying inclusion–exclusion to the events $x \in T$, $y \in T$, $z \in T$ yields

$$p^2 - 3 \cdot 3p + \big(|x, y \in T| + |x, z \in T| + |y, z \in T|\big) - |x, y, z \in T|$$

Since $|x, y \in T| = 9$, and for fixed $x \in T$ and $z \in T$ the value $y = -1 - x - z$ is uniquely determined, so $|x, z \in T| = |y, z \in T| = 9$. Moreover $|x, y, z \in T| = \#\{(x, y) \in T^2 : \ -1 - x - y \in T\} = 6$.

Next, we exclude collisions among x, y, z. If $x = y$, then $z = -1 - 2x$. There are $p - 3$ choices with $x \in D$, and exactly one value $x = -\frac{1}{2}$ gives $z \in T$, hence there are $p - 4$ admissible pairs with $x = y$ and $x, y, z \in D$. Similarly, there are $p - 4$ admissible pairs with $x = z$ and with $y = z$. Any two collision conditions force $x = y = z$, which has one solution in $\mathbb{F}_p$. Therefore

$$|\{x = y\} \cup \{x = z\} \cup \{y = z\}| = 3(p-4) - 3 \cdot 1 + 1 = 3p - 14.$$

Subtracting gives

$$W_{\text{ord}} = p^2 - 9p + 21 - (3p - 14) = p^2 - 12p + 35.$$

Each unordered triple corresponds to $3! = 6$ ordered ones, so

$$W_{\text{unord}} = \frac{p^2 - 12p + 35}{6}.$$

By Lemma 12, each orbit contains exactly four pairwise distinct normalized representatives, hence dividing by 4 yields the claimed equivalence classes count. Therefore, the number of equivalence classes equals

$$\frac{W_{\text{unord}}}{4} = \frac{p^2 - 12p + 35}{24}.$$

□

The next step is to count the cases in Lemma 11 with exactly two coinciding representations, which arise when two of the three expressions in (26) agree. In order to avoid confusion with the previous case, we simplify the subscripts. The result is as follows.

Proposition 2. *Let $a, b \in \{1, 2, \ldots, \frac{p-1}{2}\}$ and $r \in P^*$. Then*

$$\sigma_r\psi_a + \sigma_r\psi_{a+1} \;=\; \sigma_r\psi_1 + \sigma_{2r}\psi_b, \tag{30}$$

with all subscripts taken modulo p, where

$$a \in \{2, 3, \ldots, (p-3)/2\}, \qquad b \equiv \frac{a}{2} \pmod{p}.$$

Hence there are $R := \frac{p-5}{2}(p-1)$ elements of $I(2)$ that admit exactly two distinct expressions.

Proof. Comparing the supports S of the two sides in (30), and deleting the unique complementary pair, we see that the same four terms remain, which are

$$\{\sigma_r,\ \sigma_{ra},\ \sigma_{r(a+1)},\ \sigma_{-r(a+2)}\}.$$

Since these are distinct expressions, the equality $\psi_1 \neq \psi_a$ forces $a \neq 1$. If $a = \frac{p-1}{2}$, then $\psi_{a+1} = \psi_{a-1}$, which has already been counted at $a = \frac{p-3}{2}$. One checks that for every $\sigma_j \in G \setminus \{\sigma_1\}$, $\sigma_j(\psi_a + \psi_{a+1}) \neq \psi_a + \psi_{a+1}$. Varying $r \in \mathbb{F}_p^\times$ therefore yields $R = \frac{p-5}{2}(p-1)$. □

Finally, we can determine $|I(2)|$. By the inclusion–exclusion principle,

Theorem 2.

$$|I(2)| = M - U - 2W - R = \frac{p^4 - 6p^3 + 40p^2 - 66p + 31}{72} = \frac{(p-1)^2((p-2)^2 + 27)}{72}. \tag{31}$$

3.2 Linear systems

We assume that $(x, y; p, q)$ is a non-trivial solution to (2). Recall the power series defined for sums of totally positive elements $t = t_1 + t_2$, with $\varsigma(t_1) = \varsigma(t_2)$:

$$G_t = (1 + \zeta T)^{(t_1 - \jmath t_2)/q} = 1 + \sum_{n \geq 1} \frac{a_n(t)}{q^{E(n)}} T^n = S_N(t; T) + R_N(t; T),\ N \in \mathbb{N},$$

for which the map $G_t \mapsto G_{\sigma t}$ acts by $\sigma \in G$ on the coefficients $a_n(t)$, but not on the sum of the series evaluated at $-1/x$. For $\varsigma(t) = 2$, the remainders verify, by Lemma (5), the uniform bound

$$|R_N(t; -1/x)| \leq \frac{1}{(|x|-1)^{N+1}}; \quad \forall N \geq 0. \tag{32}$$

Lemma 13. *Let $J \subset I(2)$ be a subset with $n = |J| < p-1$ and consider the infinite vectors $A(t) = (a_k(t))_{k=0}^{\infty}$, in which the $a_k(t)$ are related to the power series $G(t) = (1+\zeta T)^{(t-\mathbf{N})/q}$. Then the system $\{A(t) \ : \ t \in J\}$ is $\mathbb{K}$-linear independent.*

Proof. Since $J \subset I(2)$, it follows from (16) that $k = 1$ in the sequel. Assuming that the claim is false, there are $x_t \in \mathbb{K}$ such that $\sum_{t\in J} x_t A(t) = 0$. Since

$$(1+\zeta T)^{(t-\mathbf{N})/q} = \sum_{k=0}^{\infty} \frac{A(t)_k}{q^{E(k)}}, \quad \forall t \in J,$$

we conclude that

$$\Gamma(T) := \sum_{t\in J} x_t (1+\zeta T)^{(t-\mathbf{N})/q} = 0, \tag{33}$$

as a formal power series $\Gamma(T) \in \mathbb{K}[[T]]$; the evaluations converge in $\mathbb{C}$ for all complex T with $|T| < 1$. Let $\Omega = \{\omega_i \ : i = 1, 2, \ldots, n\}$ be an enumeration of the functions $\left((1+\zeta T)^{(t-\mathbf{N})/q}\right)_{t\in J}$. Embedding $\mathbb{K}$ in $\mathbb{L} = \mathbb{K}[\xi]$ makes $\mathbb{F} := \mathbb{L}(\Omega)$ into an abelian Kummer extension over $\mathbb{L}(T)$. Let $m = v_q\left([\mathbb{F} : \mathbb{L}(T)\right) < n$ be the number of multiplicatively independent elements of Ω; after eventual renumbering, we may assume that $\omega_1, \omega_2, \ldots, \omega_m$ are multiplicatively independent and their adjunction to $\mathbb{L}(T)$ generates the field $\mathbb{F}$. We may also assume without loss of generality that $\Gamma(T)$ is the shortest vanishing linear combination of elements in Ω over $\mathbb{L}$, so in particular, all $x_t \neq 0$. We suppose first that $m > 1$ and let $\tau_i \in \ \mathrm{Gal}\ (\mathbb{F}/\mathbb{L}(T)); i = 1, 2, \ldots, m$ be the automorphism defined by $\omega_j \mapsto \xi^{\delta_{i,j}}\omega_j$, with $\delta_{i,j}$ the Kronecker δ-symbol. For all $\tau \in \ \mathrm{Gal}\ (\mathbb{F}/\mathbb{L}(T))$, it follows from (33) that

$$\sum_{i=1}^{n} x_i \tau(\omega_i) = 0, \tag{34}$$

under the map of subscripts $t \mapsto i$ if $(1+\zeta T)^{(t-\mathbf{N})/q} = \omega_i$. In particular,

$$\sum_{j=1}^{q}\sum_{i=1}^{n} x_i \tau_1^j(\omega_i) = 0.$$

In the above sum, the coefficient of ω_1 vanishes, while $\omega_j; j = 2, \ldots, m$ being fixed by τ_1. The sum can be rewritten as

$$\sum_{j=2}^{m} q x_j \omega_j + \sum_{j=m+1}^{n} x_j c_j \omega_j = 0; \quad c_j \in \mathbb{Z}[\xi].$$

Since $m > 1$, the first sum above contains at least one term, and we thus obtained a shorter non trivial vanishing linear combination, in contradiction with our assumption.

If $m = 1$, then all ω_i are multiplicatively dependent on a single radical and, by Kummer theory, there is for each $i = 2, 3, \ldots, n$ an $e_i \in \mathbb{Z} \setminus q\mathbb{Z}$ and $w_i \in \mathbb{L}(T)$, such that $\omega_i = w_i \omega_1^{e_i}$; we also let $w_1 = e_1 = 1$. Then $e_i \not\equiv e_j \bmod q$ for $i \neq j$. Otherwise, $\omega_i/\omega_j \in \mathbb{L}(T)$ and consequently $t_i \equiv t_j \bmod q$. Since $J \subset I(2)$, the coefficients in $t = \sum_{c \in P^*} n_c(t)\sigma_c^{-1}$, for $t \in J$, are $n_c(t) \in \{0, 1, 2\}$ so $t_i \equiv t_j \bmod q$ is impossible for $q \geq 3$. The e_i being pairwise different, the ω_i are part of an $\mathbb{L}(T)$-base for the vector space $\mathbb{F}/\mathbb{L}(T)$. In particular, in this case, we should have $n < q$. This completes the proof (see also [1], Prop 2.1.8, page 56) □

Proposition 3. *Let $J \subset I(2)$ be a subset on which κ is constant; say $\kappa(J) = \{u\}$. Then $|J| \leq \lceil \frac{2(p-1)}{3} \rceil$.*

Proof. Let $J \subset I(2)$, with $p - 1 \geq |J| = n > \frac{p-1}{2}$ be such that $\kappa(J) = \{u\}$ and consider the indeterminate vector $\vec{\ell} = (\ell_t)_{t \in J} \in (\mathbb{Z}[\zeta])^{|J|}$. We use the power series in subsection 2.3 and the strategy described in Remark 2 to raise a contradiction to the assumption that $|J| \geq \lceil \frac{2(p-1)}{3} \rceil + 1$. The vector $\vec{\ell}$ will produce a linear combination of $\beta(t) : t \in J$ with vanishing leading coefficients, leading to contradictory bounds on x. For this, we set up the linear system:

$$\mathcal{S}_n : \begin{array}{rcll} \sum_{t \in J} \ell_t \beta(t) & = & A, & \text{and} \\ \sum_{t \in J} \ell_t a_k(t) & = & 0, & k = 0, 1, \ldots, n-2. \end{array} \tag{35}$$

We assume first that the linear system is regular and $A \in \mathbb{Z}[\zeta]$ is the determinant of the system matrix M. By the Cramer rule, the solutions are $\ell_t = \frac{\det(M_t)}{A}$, where the matrices M_t are obtained by replacing in M the column of index t be the column of constants, which is $(A, 0, \ldots, 0)$. Since the entries of M are integral, developing $\det(M_t)$ along the replaced column, we see that $M_t \in A \cdot \mathbb{Z}[\zeta]$ and thus $\ell_t \in \mathbb{Z}[\zeta]$. More precisely, we have $\ell_t = \det(M_t')$, with M_t' the minor of M obtained by removing the first line and the column of index t. By Lemma (4), $|a_k(t)| \leq q^{k+e(k)}$, with $e(k) = v_q(k!)$. We use the Hadamard bound for estimating first $|A|$:

$$|A| \leq \sqrt{n}^n \cdot z \cdot \prod_{k=0}^{n-2} q^k = \sqrt{n}^n \cdot z \cdot q^{\binom{n-1}{2}} < \sqrt{n}^n \cdot z \cdot q^{\binom{n}{2}} =: B. \tag{36}$$

Using the same bounds, we see that

$$|\det(M_t')| < \Lambda := \sqrt{n-1}^{n-1} q^{\binom{n}{2}} < \frac{B}{z\sqrt{n}}, \tag{37}$$

so we may conclude that $\mathcal{S}_n$ has a solution with $||\vec{\ell}|| < \Lambda$, where the infinite norm used here is $||\vec{x}|| = \max_j(|x_j|)$.

The bounds above have been derived using uniform rational bounds on the absolute values of the entries in M. It follows that they are Galois invariant, and

we may state that $|\sigma A| < B$ for all $\sigma \in G$. For $\sigma = 1$, however, the power series development offers much stronger bounds. We deduce from the system (35), that the linear combination identified with A by the second equation, has vanishing leading terms, and thus:

$$|A| = \left| \sum_{t \in J} \ell_t z R_{n-2}(t; -1/x) \right| < \frac{z\Lambda n}{(|x|-1)^{n-1}} < \frac{2B\sqrt{n}}{|x|^{n-1}}. \tag{38}$$

Since complex conjugation is continuous on $\mathbb{C}$, we also deduce from (38), that

$$|\overline{A}| = |\jmath A| < \frac{2B\sqrt{n}}{|x|^{n-1}}.$$

We assumed that $\mathcal{S}_n$ is regular, so $A \neq 0$ and thus $|\mathbf{N}(A)| \geq 1$. Assembling the results, we find

$$1 \leq |\mathbf{N}(A)| < \frac{4nB^{p-1}}{|x|^{2(n-1)}}.$$

Inserting the inequality $z^q < |x|^{p-1}$, rearranging factors above, we find

$$|x|^{2(n-1)-(p-1)^2/q} < 4n \cdot n^{\frac{n(p-1)}{2}} \cdot q^{(p-1)\binom{n}{2}} < 4n \cdot n^{\frac{n(p-1)}{2}} \cdot x^{\frac{4\binom{n}{2}}{3q}}. \tag{39}$$

by using equation 4 and also $q < (p-1)^2$ proved in [7]. Comparing exponents of x and since $4n \cdot n^{\frac{n(p-1)}{2}} < n^{\frac{np}{2}} < |x|^{\frac{n}{3q}}$, we find after dividing by $n-1$:

$$2 < \frac{(p-1)^2}{q(n-1)} + \frac{n}{3q(n-1)} + \frac{2n}{3q}.$$

Since $q > p$ is prime, we have $q \geq p+2$. Set $y := \dfrac{n-1}{p-1} \in (0,1]$, i.e. $n = (p-1)y + 1$. Then

$$\frac{(p-1)^2}{q(n-1)} \leq \frac{p-1}{(p+2)y},$$

$$\frac{n}{3q(n-1)} = \frac{1}{3q}\Big(1 + \frac{1}{(p-1)y}\Big) \;\leq\; \frac{1}{3(p+2)} + \frac{1}{3(p+2)(p-1)} \cdot \frac{1}{y},$$

$$\frac{2n}{3q} \leq \frac{2}{3(p+2)}\big((p-1)y+1\big) = \frac{2}{3} \cdot \frac{p-1}{p+2}\, y + \frac{2}{3(p+2)}.$$

Based on these estimates, we obtain

$$\frac{(p-1)^2}{q(n-1)} + \frac{n}{3q(n-1)} + \frac{2n}{3q} \leq \frac{p-1}{p+2}\left(\frac{1}{y} + \frac{2}{3}y\right) + \frac{1}{p+2} + \frac{1}{3(p+2)(p-1)} \cdot \frac{1}{y}. \tag{40}$$

For $y \in \left[\frac{2}{3}, 1\right]$ (i.e. $n \geq \frac{2}{3}(p-1)+1$), we have $\frac{1}{y} \leq \frac{3}{2}$ and the function $g(y) := \dfrac{1}{y} + \dfrac{2}{3}y$ is decreasing with $g(2/3) = \frac{35}{18}$. Hence, for all such y,

$$\frac{p-1}{p+2}\left(\frac{1}{y}+\frac{2}{3}y\right)+\frac{1}{p+2}+\frac{1}{3(p+2)(p-1)}\cdot\frac{1}{y}$$
$$\leq \frac{p-1}{p+2}\cdot\frac{35}{18}+\frac{1}{p+2}+\frac{1}{2(p+2)(p-1)} < 2,$$

which leads to a contradiction. Therefore (39) cannot hold whenever

$$n \geq \frac{2}{3}(p-1)+1.$$

For discussing the singular case, we introduce the following notations:

$$A_k = (a_k(t))_{t\in J};\ k\geq 0; \quad A_{-1}=(\beta(t))_{t\in J}; \tag{41}$$
$$V = [A_k; k=0,1,\ldots,n-2]_{\mathbb{K}}. \tag{42}$$

Let $\mathrm{rk}(M)=r<n$, for M the same matrix as above. We assume first that $A_{-1}\notin V$ and let[4]

$$A_{j_0}=A_0, A_{j_1}=A_1, A_{j_2},\ldots,A_{j_{r-2}}; \quad 2\leq j_2<j_3\ldots<j_{r-2}\leq n-2$$

form a base for the vector space V. We build a regular matrix $M_1\in \mathrm{GL}(\mathbb{K},r)$ by setting the row vectors $A_{-1}; A_{j_i}, i=0,1,\ldots,r-2$ and extracting r linear independent columns; let these columns correspond to the subset $J'\subset J$ of indices. The choice is possible since row and column ranks are equal in M and M_1 will be a square submatrix of maximal rank. Let $A=\det(M_1)\neq 0$ and define the vector of constants $Y=(A,0,\ldots,0)\in\mathbb{K}^r$. Let $\vec{\ell}$ be the unique solution of the system $M_1\vec{\ell}=Y$. Since $A_k\in V$ for every $k=0,1,\ldots,n-2$, we deduce from the construction that

$$\sum_{t\in J'}\ell(t)a_k(t) = 0; \quad k=0,1,\ldots,n-2, \quad \text{while} \tag{43}$$
$$\sum_{t\in J'}\ell(t)\beta(t) = A.$$

We use the Hadamard inequalities to compute the bounds on $|A|$ and $||\vec{\ell}||$. Leaving details to the reader, the results are

$$|A| < B := r^{r/2}zq^b; \quad b=\binom{n}{2};$$
$$||\vec{\ell}|| < \Lambda := (r-1)^{\frac{r-1}{2}}q^{\binom{n}{2}}.$$

Like before, we compare bounds on the norm of A, finding

$$1\leq |\mathbf{N}(A)| < \frac{4rB^{p-1}}{x^{2(n-1)}}$$

[4]One verifies from the definition that A_0 and A_1 are always linear independent.

The bound B being in this case slightly better than in the regular case, we obtain in this case too, a contradiction, showing that the premise implies $n \leq \lceil \frac{2(p-1)}{3} \rceil$.

Assuming now that $A_{-1} \in V$, we build a regular matrix $M_2 \in \text{ GL } (\mathbb{K}, r+1)$ as follows. Let $j_i; i = 0, 1, \ldots, r-1$ be the smallest indices $\leq n-2$, in ascending order, such that $A_{j_i}; i = 0, 1, \ldots, r-1$ span V. By Lemma 13, there are some $N \geq n-1$ such that $A_N \notin V$: otherwise, the infinite vectors of the coefficients span a space of dimension $r < n$, which contradicts the Lemma. Choose the least N such that $A_N \notin V$. Using these vectors as rows, extract a set of $r+1$ linear independent columns and build thus the matrix M_2. The linear system will be

$$\begin{aligned} \sum_{t \in J'} \ell(t) a_k(t) &= 0; \quad \text{for } k = j_i, i = 0, 1, \ldots, r-1, \text{ and} \\ \sum_{t \in J'} \ell(t) a_N(t) &= A = \det(M_2). \end{aligned}$$

The bounds are now different in size from the previous cases, since the row A_{-1} – of high absolute value – does not occur in M_2. One verifies that

$$\begin{aligned} |A| &< B := (r+1)^{\frac{r+1}{2}} q^b; \quad b = N(\frac{q}{q-1}) + 1 + \binom{n-1}{2} - \binom{n-r+1}{2}; \\ ||\vec{\ell}|| &< \Lambda := r^{\frac{r}{2}} q^{1+\binom{n-1}{2} - \binom{n-r+2}{2}}. \end{aligned}$$

The norm double estimates are now

$$1 \leq |\mathbf{N}(A)| < \frac{4r^2 \Lambda^2 B^{p-3}}{x^{2(N+1)}} < \frac{B^{p-1}}{x^{2(N+1)}}$$

Notice in this case that the power of x in the denominator depends on N rather than n, thus improving our upper bound. Since $r \leq n-1$ and $N \geq n-1$, the above inequality fails when $n \geq \frac{2}{3}(p-1) + 1$, so we reach a contradiction again, thus completing the proof. □

For some $J \subset I(2)$, we let supp $(u) = \{t \in J \ : \ \kappa(t) = u\}$, for $u \in Q$, the *support* of $u \in Q$, among the elements of J. Let $U = \{u \in Q \ : \ \text{supp } (u) \neq \emptyset\}$ and $h = |U| \leq q$. The supp (u) are random sets, being determined by the random behavior of $\kappa(t)$. However, since $I(2)$ is large enough, the Proposition 3 allows us to complete the proof of Theorem 1:

Proof. Let $J \subset I(2)$ and $h_u = |\kappa^{-1}(u)|$. We have $|J| = \sum_{u \in U} h_u$; by Proposition 3, $h_u < \frac{2(p-1)}{3} + 1$ and thus $|J| < \frac{2q(p+2)}{3} =: N$. This holds for all $J \subset I(2)$, leading to a contradiction to the estimate for $|I(2)|$ in (31), when $p > 53$.

For smaller values of p, one can check the double Wieferich criterion – Theorem I in Schoof's [10] – for $p < 59$ and all $q < (p-1)^2$. The PARI GP program in the Listing 1[5] of the footnote counts the number of pairs in this range, which verifies the double Wieferich criterion. The number is zero. This completes the proof. □

[5]The following PARI GP program does the verification of the double Wieferich condition:

4 Outlook

The result of this paper is not a new one. It is however valuable, in as much as it responds positively to a question that was frequently asked 20 years ago, about the original proof of the Catalan conjecture: *Can one provide a proof using only Stickelberger annihilation and avoiding the use of the class field theoretic more intricate Theorem of Thaine and its consequences for q-primary units?*. The fact that the question was not easy, even decades later, is confirmed by the fact that our first attempt underestimated the old known obstruction, thus containing an error. We apologize to the readers of the first published version for this mishap!

4.1 Using AI

The combinatorics required for the estimate of the number of distinct elements in $I(1)$ and $I(2)$ were quite involved, and they were first surprising before becoming simpler, step by step. Even programming in PARI GP was not evident, due to counter-intuitive stack rules in the call of nested functions. So we tried the use of AI, a new experience for us, about which we wish to share some of our experiences.

The first attempt consisted in a direct question: we defined Stickelberger elements as vectors of explicit coefficients over the group ring and asked for a count. The AI did provide a result, and it took some tricks to realize it was blatantly false! We then decided to sketch in detail the program we wished to write in PARI, with data description, input and output – and ask the AI to produce it. The jump in quality was unexpected. One of the AIs (we tested several systems available on line), produced the following perfect program. The program can be immediately run in PARI GP. It produces a list of the elements $\psi(m,a) := \sigma_a(\psi_m) \in I(1)$, expressed by their vectors of coefficients in the $\mathbb{Z}[G]$-expansion

$$\psi(m,a) = \sum_{c \in P^*} n_c(m,a)\sigma_c^{-1}.$$

The last four columns contain the pair (m,a) and a further pair which is $(0,0)$ if this is the first occurrence of the element in the list; if the second pair is not trivial, it then indicates that the present element collides with one previously appearing in the list; the pair then indicates the labeling pair of that element.

```
doubleWif( p, q ) = {
  rs=1; if( component( Mod(p, q^2 )^(q-1), 2 ) == 1, if( component(
      Mod(q, p^2 )^(p-1), 2 ) == 1, rs = 0)); rs
};
{ t=0;
  forprime( p = 47, 53, s = 0; forprime( q = nextprime( p+1), (p-1)
      ^2, s = s + (1-doubleWif( p, q ))); t = t+s); t}
```

Listing 1: PARI/GP: Double Wieferich Check

```
jacobi_data_with_selective_counts(p) = {
  my(half = (p-1)\2, j1vecs = vector(half), fp1vecs = vector(half),
      fps_map, jDat, v, fp_v, found, a, c, r, m, fps);

  \\ Precompute base vectors j(1, m) and their fingerprints
  for(m = 1, half,
  v = vector(p-1, c,
  my(yc = (m * c) % p);
  if((c + yc) >= p, 1, 0)
  );
  j1vecs[m] = v;
  fp1vecs[m] = sum(c = 1, p-1, v[c] * 2^(c-1));
  );

  fps_map = Map(); \\ Map: fingerprint -> index m
  jDat = matrix(half, 4); \\ Result matrix

  \\ Store all unique vectors from first-occurrence orbits with their
      source information
  my(all_vectors_info = List()); \\ List of [vector_str, n, a]
  my(orbit_unique_vectors = vector(half));

  for(n = 1, half,
  my(j1n = j1vecs[n]); \\ Vector for j(1, n)
  fps = vector(p-1); \\ Fingerprints for the orbit j(a, a*n)
  found = 0; \\ Collision flag

  \\ Store unique vectors in current orbit with their source information
  my(orbit_vecs_set = Set());
  my(orbit_vecs_info = List()); \\ Store [vector_str, n, a] for current
      orbit

  for(a = 1, p-1,
  \\ Build vector j(a, a*n) by permuting j1n
  v = vector(p-1, c,
  r = lift(Mod(a*c, p)); \\ Residue in [1, p-1]
  j1n[r] \\ Permuted component
  );
  fp_v = sum(c = 1, p-1, v[c] * 2^(c-1)); \\ Fingerprint
  fps[a] = fp_v;

  \\ Convert vector to string for storage
  my(vec_str = concat(vector(p-1, c, Str(v[c]))));

  \\ Check if this vector is new in the orbit
  if(!setsearch(orbit_vecs_set, vec_str),
  orbit_vecs_set = setunion(orbit_vecs_set, [vec_str]);
  listput(orbit_vecs_info, [vec_str, n, a]);
  );
```

```
\\ Check collision with j(1, m) for m < n
if(!found && n > 1,
if(mapisdefined(fps_map, fp_v, &m),
jDat[n, 3] = m;
jDat[n, 4] = a;
found = 1;
);
);
);

orbit_unique_vectors[n] = Vec(orbit_vecs_info);

jDat[n, 1] = n; \\ Store n
jDat[n, 2] = #orbit_vecs_set; \\ Orbit length (number of unique
    vectors)

if(!found, \\ No collision
jDat[n, 3] = 0;
jDat[n, 4] = 0;
\\ If it's a first-occurrence orbit, add all its unique vectors to the
     total list
for(i = 1, #orbit_unique_vectors[n],
listput(all_vectors_info, orbit_unique_vectors[n][i]);
);
);

mapput(fps_map, fp1vecs[n], n); \\ Add j(1, n) to map
);

\\ Calculate the sum of the second column for rows where the third
    column is 0
my(total_unique_vectors = 0);
for(n = 1, half,
if(jDat[n, 3] == 0,
total_unique_vectors += jDat[n, 2];
);
);

print("Number of unique vectors in first-occurrence orbits: ",
    total_unique_vectors);

\\ Output all unique vectors from first-occurrence orbits with source
    information
print("\nAll unique vectors from first-occurrence orbits:");
for(i = 1, #all_vectors_info,
my(info = all_vectors_info[i]);
print("Vector ", i, ": ", info[1], " (", info[2], ",", info[3], ")");
);

\\ Compute pairwise sums of all unique vectors from first-occurrence
    orbits
```

```
\\ Use lists to store sum vectors, their counts, and the pairs that
    generate them
my(sum_vectors = List());
my(sum_counts = List());
my(sum_pairs = List());

for(i = 1, #all_vectors_info,
for(j = i, #all_vectors_info,
\\ Get vector information
my(info_i = all_vectors_info[i]);
my(vec_i_str = info_i[1]);
my(n_i = info_i[2]);
my(a_i = info_i[3]);

my(info_j = all_vectors_info[j]);
my(vec_j_str = info_j[1]);
my(n_j = info_j[2]);
my(a_j = info_j[3]);

\\ Compute the sum of two vectors (component-wise addition)
my(sum_vec = vector(p-1, c,
digit_i = eval(Strchr(Vecsmall(vec_i_str)[c]));
digit_j = eval(Strchr(Vecsmall(vec_j_str)[c]));
digit_i + digit_j
));

\\ Convert the sum vector to a string for comparison
my(sum_str = concat(vector(p-1, c, Str(sum_vec[c]))));

\\ Create a description of the pair that generated this sum using (n,a
    ) format
my(pair_desc = Str("((", n_i, ",", a_i, ")+(", n_j, ",", a_j, "))"));

\\ Check if this sum vector is already in our list
my(found_sum = 0);
for(k = 1, #sum_vectors,
if(sum_vectors[k] == sum_str,
sum_counts[k] = sum_counts[k] + 1;
listput(sum_pairs[k], pair_desc);
found_sum = 1;
break;
);
);

if(!found_sum,
listput(sum_vectors, sum_str);
listput(sum_counts, 1);
my(new_pair_list = List());
listput(new_pair_list, pair_desc);
listput(sum_pairs, new_pair_list);
);
```

```
  );
  );

  print("\nNumber of distinct sums of two Jacobi sums: ", #sum_vectors);

  \\ Output all distinct sum vectors with their counts and generating
      pairs
  print("\nAll distinct sum vectors:");
  for(i = 1, #sum_vectors,
  if(sum_counts[i] > 1,
  print("Sum vector ", i, ": ", sum_vectors[i], " (count: ", sum_counts[
      i], ")");
  print(" Generated by: ", Vec(sum_pairs[i]));
  ,
  print("Sum vector ", i, ": ", sum_vectors[i]);
  print(" Generated by: ", Vec(sum_pairs[i]));
  );
  );

  \\ Automatically print the jDat matrix at the end
  print("\njDat matrix:");
  print(jDat);

  return(jDat);
}

\\ Example usage with automatic execution:
p = 5; \\ Replace with your prime
print("Starting computation for p = ", p);
jDat = jacobi_data_with_selective_counts(p);
```

Listing 2: PARI/GP: jacobi_data_with_selective_counts

This immediately confirmed our results for $I(1)$ for small primes p, and allowed deducing the formula (24). We even received some valuable indications for a possible proof, which could be worked out to a correct validation of the formula found. Only after completing this proof, we found the simpler proof provided above. Also for $I(2)$, asking the AI to list the collisions occurring for fixed p, we received indicative numerical confirmation of the analysis that then led to the final proof of the formula (31). In conclusion, our experience suggests that AI is a new step beyond programming of symbolic computation software, as support for mathematical research. Unlike programming, where the result only depends on writing a correct software, with AI, the programming is spared at the expense of some trial and error, verification of results, and improvement of questions. But in the case of success, one obtains a powerful aid that in the end avoids the work of improving and extending the software: Once the AI has produced the first working program, asking more complex questions along the same lines becomes easier, and the extended program is ready in seconds!

References

[1] Albu, T., *Cogalois theory*, Taylor & Francis, 2002.

[2] Bilu, Y., Bugeaud, Y. and Mignotte, M., *The problem of Catalan*, Springer, 2014.

[3] Cassels, J. W. S., *On the equation* $a^x - b^y = 1$, II, Proc. Cambridge Philos. Soc. **56** (1960), 97–103.

[4] Fueter, R., *Kummer's Kriterien zum letzten Theorem von Fermat*, Math. Ann. **85** (1922), 11-20.

[5] Iwasawa, K., *A note on Jacobi sums*, Proc. Sympos. Pure Math. **15** (1975), 447-459.

[6] Mihăilescu, P., *Class number conditions for the diagonal case of the equation of Nagell and Ljunggren*, Festschrift to the 70-th Birthday of Wolfgang Schmidt (Eds. Schlickewei, H. P. et al.), Berlin, Springer, 243-274, 2008.

[7] Mihăilescu, P., *New bounds and conditions for the equation of Nagell-Ljunggren*, J. Number Theory **124** (2007), no. 2, 380-395.

[8] Mihăilescu, P., *Primary cyclotomic units and a proof of Catalan's conjecture*, J. Reine Angew. Math. **572** (2004), 167-195.

[9] Ribenboim, P., *Catalan's conjecture*, Academic Press, 1994.

[10] Schoof, R., *Catalan's conjecture*, Universitext, Springer, 2008.

[11] Washington, L., *Introduction to cyclotomic fields*, 2nd ed., Graduate Texts in Mathematics, Springer, 1997.