Chin. Ann. Math. Ser. B
47(2), 2026, 389–412
DOI: 10.1007/s11401-026-0018-7

Chinese Annals of
Mathematics, Series B

New Conditions on the Nagell-Ljunggren Equation $\frac{x^p-1}{x-1}=y^q$

Han CHEN[1] Preda MIHĂILESCU[1]

Abstract In this paper, the authors consider the equation $\frac{x^p-1}{x-1}=p^e y^q$, for distinct odd prime exponents p,q, and show that, for $p>3$, it has no solutions under the condition that q does not divide h_p^-, the minus part of the class number of the p-th cyclotomic field.

Keywords Nagell-Ljunggren conjecture, Diophantine equations, Class field theory
2020 MR Subject Classification 11D61, 11R37

1 Introduction

The Diophantine equation of Nagell-Ljunggren

$$\frac{x^n-1}{x-1}=y^q \tag{1.1}$$

has six known solutions

$$(x,y,n,q)\in\mathcal{S}:=\{(3,\pm 11,5,2),(7,\pm 20,4,2),(18,7,3,3),(-19,7,3,3)\} \tag{1.2}$$

in integers x,y,q and n with $|x|,|y|,q>1$ and $n>2$. These solutions carry the exponents $q=2$ or $q=3$. The conjecture of Nagell and Ljunggren states that these are the only solutions in integers with exponents larger than one. This claim can be reduced (see [1]) to the following set of statements, relative to pairs of odd primed (p,q): The equations

$$\frac{x^p-1}{x-1}=p^e y^q,\quad e=\begin{cases}1, & \text{if } x\equiv 1 \bmod p,\\ 0, & \text{otherwise}\end{cases} \tag{1.3}$$

have no solutions outside $\mathcal{S}$. Here, the exponent of x is also prime, so one has to distinguish the cases in which $p\mid(x-1)$ and $p\nmid(x-1)$. Thus, recent work on the Nagell-Ljunggren equation advanced in two directions. On the one hand, one proved decreasing upper bounds in the number of factors of the exponent n in (1.1). The latest results in this direction are due to Bugeaud and the second author in [2], improved by Bennett and Levin in [3].

Of course, one can consider (1.3) as one independent of (1.1), and then the question remains, whether it has nontrivial solutions. Dupuy treated the case of solutions $x\equiv 1 \bmod p$ — thus $e=1$ — elegantly, under the additional assumption $q\nmid h_p^-$, in [4], proving the following theorem.

Manuscript received May 11, 2022. Revised October 14, 2023.
[1] Mathematisches Institut der Universität Göttingen, Bunsenstraße 3–5, 37073, Göttingen, Germany.
E-mail: han.chen@mathematik.uni-goettingen.de preda@uni-math.gwdg.de

Theorem 1.1 *If $q \nmid h_p^-$, then* (1.3) *has no solutions with $e = 1$ — or, equivalently, with $x \equiv 1 \bmod p$.*

In a certain sense, the main result of our paper can be seen as generalizing this result to arbitrary values of x. We thus prove the following theorem.

Theorem 1.2 *The equation* (1.3) *has no integer solutions outside $\mathcal{S}$, if $q \nmid h_p^-$ — except*[1] *possibly for some solutions with $p < 29$ and $x < 0$.*

Nagell and Ljunggren were the first having investigated this equation in [6–8] and they have eliminated some small exponents, showing that there are no other solutions for these values of the exponents p, q, so we may assume in the sequel that $p \geq 5$ is an odd prime. Later work of Cassels [9] showed that if the famous equation of Catalan $x^p - y^q = 1$ has nontrivial solutions $(x, y; p, q)$, then a fortiori (1.1) holds for these values, the index e equals one, and in addition, $p^{q-1} \mid (x - 1)$; $q^{p-1} \mid (y + 1)$. Since (1.1) appears as a partial condition for the Catalan equation, we see that the first is a harder condition. The challenge of solving the Nagell-Ljunggren conjecture was taken up by numerous number theorists, including Bennett [10], Bugeaud [5] with various coauthors, Le [11], Shorey [12], Shorey and Tijdeman [13], etc. We refer to [2–3] for detailed historical references.

In this paper, we shall study (1.3) under additional class number condition $q \nmid h_p^-$. We will have to analyse numerous subcases, applying a rich variety of approaches in order to eliminate all possible solutions that may arise under the given restriction.

We let $\mathbb{K} = \mathbb{Q}[\zeta_p]$ be the p-th cyclotomic field, with p an odd prime. The class group of a generic number field $\mathbf{K}$ will be denoted by $\mathcal{C}(\mathbf{K})$ and $h_{\mathbf{K}} = |\mathcal{C}(\mathbf{K})|$. In the case of the p-th cyclotomic field, we write $h_p = |h_{\mathbb{K}}|$ and $h_p^- = |\mathcal{C}(\mathbb{K})/\iota(\mathcal{C}(\mathbb{K}^+))|$, with $\iota : \mathcal{C}(\mathbb{K}^+) \to \mathcal{C}(\mathbb{K})$ being induced by the ideal lift map from the maximal totally real subfield $\mathbb{K}^+ \subset \mathbb{K}$ to $\mathbb{K}$.

In 2008, the second author [14] studied the diagonal case $p = q$. He gave general class number conditions which are the first known general algebraic necessary conditions for the equation to have solutions. Among others, these lead, based on computer results produced for the investigation of the Fermat Equation, to the conclusion that the equation has no solutions for $p < 12\ 000\ 000$. In another paper [15], he considered the case $q \neq p$. The following unconditional criterion can be found in that paper.

Theorem 1.3 *For $q \neq p$, two distinct odd primes,* (1.3) *has no solutions if $q > (p - 1)^2$.*

Further conditional results were provided in the paper, for the case when $q < (p - 1)^2$. In this paper we thus study the case $e = 0$ of (1.3), under the premise that $q \nmid h_p^-$ and $q \neq p$.

1.1 Plan and intermediate results of the paper

For primes $n \in \mathbb{N}$, we define the following two functions of n:

$$M(n) = \max\left(n, \frac{n(n-12)}{16}\right), \tag{1.4}$$

[1] Our criteria are independent of the sign of x, but we did use at places the bound $p \geq 29$ found in [5]: A second reading of the paper revealed that the authors used the premise $x > 0$, hence the possible exceptions.

$$M'(n) = C(n) n \log(n), \quad C(n) = \Big(\frac{\log(4)}{1 + \frac{1+\log\log(n)}{\log(n)}} - \frac{4\log(n)}{n-2} \Big)^{-1}.$$

The local approach is based on the following technically involved result on Fermat quotients.

Proposition 1.1 *Suppose that* (1.3) *has nontrivial solutions with primes* $p \neq q$, $q \nmid h_p^-$. *Then* $q \in [M(p), (p-1)^2) \cap \mathbb{N}$ *or* $q^2 \mid x - t$ *for some* $t \in \{0, \pm 1\}$.

This leads to the first major restriction of the set of the solutions, when $q < M(p)$.

Proposition 1.2 *Suppose that* (1.3) *has nontrivial solutions with primes* $p \neq q$, $q \nmid h_p^-$ *and* $q < M(p)$. *Then* $x \equiv \pm 1 \bmod q^2$.

Using a global bounding approach based on binomial series expansions, we then deduce the following proposition.

Proposition 1.3 *Suppose that* (1.3) *has nontrivial solutions with primes* $p \neq q$, $q \nmid h_p^-$. *Then* $q < M'(p)$.

Finally, we use the condition in Proposition 1.2 for local developments of the putative solutions, in order to obtain a contradiction to the upper bounds on $|y|$ established in [15]. We thus prove the following proposition.

Proposition 1.4 *The equation* (1.3) *has no solutions with primes* $p \neq q$, $q \nmid h_p^-$ *if* $q < M(p)$*— except possibly for some solutions with* $p < 29$ *and* $x < 0$.

And then conclude with the proof of Theorem 1.2.

1.2 Notations and general facts

We assume throughout this paper that $(x, y; p, q) \notin \mathcal{S}$, $q \neq p$ is an unknown solution to (1.3) and $q \nmid h_p^-$. In view of Dupuy's Theorem, we know that $e = 0$ and in view of Theorem 1.3, $q < (p-1)^2$. We let $\mathbb{K} = \mathbb{Q}[\zeta]$ be the p-th cyclotomic extension, with ζ a p-th primitive root of unity. We shall at places use also a primitive p-th root of unity ξ and let $\mathbb{K}' = \mathbb{Q}[\xi], \mathbb{L} = \mathbb{Q}[\zeta, \xi]$.

We let $\Phi_r(x)$ be the r-th cyclotomic polynomial and define $P = \{1, 2, \cdots, p-1\}$ and $\sigma_c \in G =$ Gal $(\mathbb{Q}(\zeta)/\mathbb{Q})$ be the automorphism of $\mathbb{K}$ with $\zeta \to \zeta^c$, for $c \in P$. Complex conjugation is denoted by $\jmath_p \in G$, or simply by $\jmath$ when the group is clear from the context; we may thus write $\overline{\alpha} = \sigma_{p-1}(\alpha) = \jmath_p(\alpha) = \alpha^\jmath$. We let $\lambda = (1 - \zeta)$ be an algebraic integer generating the unique ramified prime ideal $\wp$ above p in $\mathbb{K}$.

1.2.1 Cyclotomic properties of the equation

Recall that $x \not\equiv 1 \bmod p$ and $e = 0$. We define herewith the characteristic number of (1.3) by

$$\alpha = x - \zeta \quad \text{and} \quad \alpha_c = \sigma_c(\alpha).$$

The characteristic ideal is $\mathfrak{A} = (\alpha, y)$ and one verifies directly that $(\sigma_a(\alpha), \sigma_b(\alpha)) = 1$ for $a, b \in P$, two distinct integers. Indeed, $D(a, b) = (\sigma_a(\alpha), \sigma_b(\alpha))$ contains $\zeta^a - \zeta^b = \sigma_b(\alpha) - \sigma_a(\alpha)$, and thus $D(a,b) \mid \wp$. However, $x \not\equiv 1 \bmod p$ implies that $\alpha \notin \wp$, so $D(a,b) = 1$. As a consequence, we have

$$\mathfrak{A}^q = (\alpha), \quad \mathbf{N}(\alpha) = y^q; \quad \mathbf{N}(\mathfrak{A}) = (y). \tag{1.5}$$

We proved the following lemma.

Lemma 1.1 *Assume that* (1.3) *has a nontrivial solution* $(x, y; p, q)$ *with* $x \not\equiv 1 \bmod p$ *and let* $\sigma_c(\alpha) = x - \zeta^c$. *Then the ideal* $\mathfrak{A} = (\alpha, y)$ *verifies* $\mathfrak{A}^q = (\alpha)$. *Moreover, for* $c \neq d \in P$, *we have*

$$(\sigma_c(\alpha), \sigma_d(\alpha)) = 1, \quad (\sigma_c(\mathfrak{A}), \sigma_d(\mathfrak{A})) = 1, \quad \mathbf{N}_{\mathbb{K}/\mathbb{Q}}(\mathfrak{A}) = (y).$$

1.2.2 Class number condition

We assume $q \nmid h_p^-$; then the image of the class $[\mathfrak{A}]$ in the quotient $(\mathcal{C}(\mathbb{K})/\iota(\mathcal{C}(\mathbb{K}^+)))$ is trivial. Since the conjugates of $\mathfrak{A}$ are pairwise coprime, it follows that $\mathfrak{A}$ is divisible by no real prime. It must consequently be a principal ideal, say $\mathfrak{A} = (\rho)$. It follows that $(\alpha) = (\rho^q)$ and by transforming the identity of ideals into one of algebraic numbers, we find the following lemma.

Lemma 1.2 *Under the premises above, there is a* $\rho \in \mathbb{Z}[\zeta]$ *and a unit* $\varepsilon \in (\mathcal{O}(\mathbb{K}^+))^\times$ *such that*

$$\alpha = \varepsilon \cdot \rho^q. \tag{1.6}$$

2 A General Result on Fermat Quotients

In this section, we prove a general result on the Fermat quotients of certain binary fractions.

Theorem 2.1 *Let* $p, q, \mathbb{K}$ *be like above, and suppose that* $x, y \in \mathbb{Z}$ *satisfy the condition that there exists a* $\beta \in \mathbb{K}$ *such that*

$$\frac{x + \zeta y}{x + \overline{\zeta} y} = \Big(\frac{\beta}{\overline{\beta}}\Big)^q.$$

If in addition $q < M(p)$ *with the function defined in* (1.4), *then there is an* $f \in \{-1, 0, 1\}$ *such that*

$$x + fy \equiv 0 \bmod q^2.$$

The proof of this theorem requires the rest of this section.

Lemma 2.1 *Let* p, q *be odd primes and let* x *and* y *be coprime integers with* $q \nmid x$ *and such that there is a* $\beta \in \mathbb{Q}(\zeta)$ *with*

$$\frac{x + \zeta^q \cdot y}{x + \overline{\zeta}^q \cdot y} = \pm\Big(\frac{\beta}{\overline{\beta}}\Big)^q. \tag{2.1}$$

Then

$$-(\zeta^q - \overline{\zeta}^q)\varphi(t) \equiv \sum_{k=1}^{q-1} \frac{t^k - t^{2-k}}{k} \cdot (\zeta^k - \overline{\zeta}^k) \bmod q, \tag{2.2}$$

where $\varphi(a) \equiv \frac{a^q - a}{q} \bmod q$ *for* $a \in \mathbb{Z}/(q^2\mathbb{Z})$ *is the Fermat quotient function and* $t :\equiv -\frac{y}{x} \bmod q^2$.

Proof We have
$$\frac{x+\zeta^q\cdot y}{x+\overline{\zeta}^q\cdot y}\equiv\Big(\frac{x+\zeta\cdot y}{x+\overline{\zeta}\cdot y}\Big)^q \bmod q\mathbb{Z}[\zeta],$$
and thus, from (2.1)
$$\pm\frac{\beta}{\overline{\beta}}=\frac{x+\zeta\cdot y}{x+\overline{\zeta}\cdot y}+q\cdot\mu$$
with $\mu\in\mathbb{Q}(\zeta)$ being a q-integer. By raising to the power q, it follows again from (2.1), that
$$\frac{x+\zeta^q\cdot y}{x+\overline{\zeta}^q\cdot y}\equiv\pm\Big(\frac{x+\zeta\cdot y}{x+\overline{\zeta}\cdot y}\Big)^q \bmod q^2\mathbb{Z}[\zeta]. \tag{2.3}$$

Note that $\varphi(a)\equiv\frac{a^q-a}{q} \bmod q$, for $(a,q)=1$ and $t\equiv-\frac{y}{x} \bmod q^2$, so $-\big(\frac{y}{x}\big)^q\equiv t+q\varphi(t) \bmod q^2$. Now
$$\begin{aligned}(x+\zeta\cdot y)^q&\equiv x^q\cdot(1-t\cdot\zeta)^q\equiv(x+q\varphi(x))\cdot(1-t\zeta)^q\\&\equiv(x+q\varphi(x))\cdot(1-t\zeta^q+qf(\zeta)) \bmod q^2,\end{aligned}$$
where
$$f(\zeta)=-\zeta^q\cdot\varphi(t)+\frac{1}{q}\cdot\sum_{k=1}^{q-1}\binom{q}{k}(-t\zeta)^k\equiv-\Big(\zeta^q\cdot\varphi(t)+\sum_{k=1}^{q-1}\frac{t^k\zeta^k}{k}\Big) \bmod q.$$

Writing $\alpha=1+\frac{y}{x}\zeta^q=1-t\zeta^q+q^2z$ for some $z\in\mathbb{Z}[\zeta]$ and eliminating denominators in (2.3), we find that
$$\alpha\cdot(x+q\varphi(x))(\overline{\alpha}+q\cdot f(\overline{\zeta}))\equiv\overline{\alpha}\cdot(x+q\varphi(x))\cdot(\alpha+q\cdot f(\zeta)) \bmod q^2$$
and
$$\alpha\cdot f(\overline{\zeta})\equiv\overline{\alpha}\cdot f(\zeta) \bmod q.$$
We let $S=\sum\limits_{k=1}^{q-1}\frac{t^k\zeta^k}{k}$. Regrouping the terms, we find
$$(1-t\overline{\zeta}^q)\cdot(\varphi(t)\cdot\zeta^q+S)\equiv(1-t\zeta^q)\cdot(\varphi(t)\cdot\overline{\zeta}^q+\overline{S}) \bmod q,$$
hence
$$-(\zeta^q-\overline{\zeta}^q)\varphi(t)\equiv(1-t\overline{\zeta}^q)S-(1-t\zeta^q)\overline{S} \bmod q$$
and
$$-(\zeta^q-\overline{\zeta}^q)\varphi(t)\equiv\sum_{k=1}^{q-1}\frac{t^k}{k}(\zeta^k-\overline{\zeta}^k)-\sum_{k=1}^{q-1}\frac{t^{k+1}}{k}(\zeta^{k-q}-\overline{\zeta}^{k-q}) \bmod q.$$
We regroup the powers of ζ using $q-k\equiv-k \bmod q$, thus $\frac{\zeta^{k-q}}{k}\equiv-\frac{\overline{\zeta}^{q-k}}{q-k}$, which can be applied in the above for $k=1,2,\cdots,q-1$,
$$-(\zeta^q-\overline{\zeta}^q)\varphi(t)\equiv\sum_{k=1}^{q-1}\frac{t^k-t^{2-k}}{k}\cdot(\zeta^k-\overline{\zeta}^k) \bmod q,$$

which is the statement of (2.2).

Lemma 2.1 essentially yields a system of equations modulo q in the unknown t. It turns out that under some additional conditions on p and q, there are only three possible values for t (one of which is $t = 0$). This reflects the main ideas which will subsequently lead, by a more in depth study of the system (2.2), to a sharper inequality between p and q. The light result is the following proposition.

Proposition 2.1 *Let x, y be coprime integers, assume that $p > q$ are odd primes and there is a $\beta \in \mathbb{Q}(\zeta)$ such that* (2.1) *holds. Then*

$$x + f \cdot y \equiv 0 \bmod q^2 \tag{2.4}$$

for some $f \in \{-1, 0, 1\}$.

Proof Assume first that $x \equiv 0 \bmod q$ and $x = qu$ with $(u, q) = 1$. Since $(x, y) = 1$ and $p \neq q$, it follows that $(x + \zeta^a y, q) = 1$, so the right-hand side of (2.1) is a q-integer. The equation is Galois-invariant, so we can replace ζ by ζ^q. Thus (2.1) becomes

$$\frac{y + q\overline{\zeta}^q u}{y + q\zeta^q u} = \gamma^q$$

with $\gamma = \pm\zeta^{-2} \cdot \frac{\beta}{\overline{\beta}}$. From the definition, we see that $\gamma^q \equiv 1 \bmod q\mathbb{Z}[\zeta]$. Let $\mathfrak{Q} \subset \mathbb{Z}[\zeta]$ be a prime above q, and f be its height. We have a fortiori $\gamma^q \equiv 1 \bmod \mathfrak{Q}$ and raising the identity to the power q^{f-1}, we obtain

$$\gamma \equiv \gamma^{q \cdot q^{f-1}} \equiv 1 \bmod \mathfrak{Q}.$$

This holds for all primes of $\mathbb{Z}[\zeta]$ above q, so $\gamma \equiv 1 \bmod q$. Consequently, $\gamma = 1 + qw$ for some $w \in \mathbb{Z}[\zeta]$, and thus $\gamma^q = (1 + qw)^q \equiv 1 \bmod q^2$, and $y + qu\zeta^q \equiv y + qu\overline{\zeta}^q \bmod q^2$. Thus $u \cdot (\zeta^q - \overline{\zeta}^q) \equiv 0 \bmod q$. Since $p \neq q$, this is only possible if $u \equiv 0 \bmod q$ and thus $x \equiv 0 \bmod q^2$. We can interchange x and y, so this proves that if x or y is divisible by q, then it is divisible by q^2, which takes care of $f = 0$ in this case.

We may now assume that $q \nmid x, y$ and use the previous lemma, which implies that (2.2) holds under the given premises. Since the set $\{\zeta, \zeta^2, \cdots, \zeta^{p-1}\}$ builds a basis of the algebra $\mathbb{Z}[\zeta]/(q \cdot \mathbb{Z}[\zeta])$, the coefficients of the single powers in the above identity must all vanish and $p > q + 1$ implies that the coefficient of ζ is $a_1 = t(1 - t^{-4})$ and thus

$$t^4 \equiv 1 \bmod q$$

must hold. Furthermore, if $q + 2 < p$, then the coefficient of ζ^2 is

$$2 \cdot a_2 = (t^2 - t^{-4}) \equiv 0,$$

hence

$$t^6 - 1 \equiv 0 \bmod q.$$

The last two congruences in t have the only common solution $t^2 = 1 \bmod q$. One easily verifies that if this holds, then the right-hand side in (2.2) vanishes and thus $\varphi(t) \equiv 0 \bmod q$. This leads to the possible solution $x \pm y \equiv 0 \bmod q^2$. Inserting the value back shows that this is indeed a solution of (2.1). If $p = q + 2$, then we still have $a_1 = t^{-3}(t^4 - 1)$ so $t^4 \equiv 1 \bmod q$. If $t^2 - 1 \equiv 0 \bmod q$, we find the previous solution. So let us assume that $t^2 \equiv -1 \bmod q$ and consider the second coefficient. But $\varphi(t)\overline{\zeta}^q = \varphi(t)\zeta^2$ has in this case a contribution to a_2. We estimate this coefficient by using $t^2 \equiv -1 \bmod q$,

$$\begin{aligned} 2 \cdot a_2 &\equiv t^2 - t^{-4} - 2\varphi(t) \equiv -t^{-4}(t^6 - t^2 + t^2 - 1 + 2t^4\varphi(t)) \\ &\equiv t^2 - 1 + 2\varphi(t) \equiv 2(\varphi(t) - 1) \bmod q, \end{aligned}$$

a congruence which is satisfied by $\varphi(t) \equiv 1 \bmod q$. We have to consider also

$$\begin{aligned} 3 \cdot a_3 = (t^3 - t^{-5}) - (t^{q-1} - t^{-q-1}) \equiv 0 \bmod q \\ \Leftrightarrow 0 \equiv t^{-5}(t^8 - 1) - (1 - t^{-2}) \bmod q. \end{aligned}$$

If $t^2 \equiv -1 \bmod q$, then the first term vanishes while the second is $-2 \not\equiv 0 \bmod q$, so $t^2 \equiv -1 \bmod q$ is not possible. This takes care of the case $p = q + 2$ as well, thus completing the proof of the proposition.

It follows from Lemma 2.1 that the following corollary holds.

Corollary 2.1 *If $p > q > 3$ are odd primes for which* (1.3) *has nontrivial solutions and such that $q \nmid h_p^-$, then* (2.4) *holds.*

Proof The premises of Lemma (2.1) are fulfilled and thus (2.2) holds. By setting $\beta = \rho_1$ in this equation, we find that the hypotheses of Proposition 2.1 also hold, and by its proof it follows that (2.4) must be true.

2.1 Sharpening

In order to gain more information from (2.2), also in cases when $q > p$, we need to first introduce some operations on sequences. Let $\mathbf{k}$ be a field and $\mathcal{T}$ be the space of two-sided sequences on $\mathbf{k}(t)$. We define the following operators on $\mathcal{T}$: For $a = (a_n)_{n\in\mathbb{N}} \in \mathcal{T}$, the maps $\theta_+, \theta_-, \Theta : \mathcal{T} \to \mathcal{T}$ produce the following sequences:

$$\begin{aligned} \theta_+(a)_n &= a_n - t \cdot a_{n-1}, \\ \theta_-(a)_n &= t \cdot a_n - a_{n-1}, \\ \Theta(a)_n &= \theta_+(\theta_-(a))_n. \end{aligned} \tag{2.5}$$

Furthermore, we let Δ be the classical forward difference operator

$$\Delta a_n \;=\; a_n - a_{n-1}$$

and

$$n^{\underline{k}} = n \cdot (n-1) \cdots (n-k+1)$$

be the k-th falling power of n, so $\Delta n^{\underline{k}} = k \cdot (n-1)^{\underline{k-1}}$. The main properties of the operators in (2.5) are given by the following lemma.

Lemma 2.2 *The operators θ_+ and θ_- are linear, and they commute*

$$\Theta = \theta_+ \circ \theta_- = \theta_- \circ \theta_+.$$

Furthermore,

$$\begin{cases} \theta_+(t^n) = 0, & \theta_+(t^{-n}) = (1-t^2)t^{-n}, \\ \theta_-(t^{-n}) = 0, & \theta_-(t^n) = -(1-t^2)t^{n-1} \end{cases} \tag{2.6}$$

and

$$\begin{aligned} \theta_+^l(n^{\underline{k}} \cdot t^n) &= \frac{k!}{l!} \cdot (n-l)^{\underline{k-l}} \cdot t^n, \\ \theta_-^l(n^{\underline{k}} \cdot t^{-n}) &= \frac{k!}{l!} \cdot (n-l)^{\underline{k-l}} \cdot t^{-(n-l)}, \end{aligned} \tag{2.7}$$

where we set $a^{\underline{k-l}} = 0$ if $k < l$. In particular, we have

$$\begin{aligned} \theta_+^k(n^{\underline{k}} \cdot t^n) &= k! \cdot t^n, \\ \theta_-^k(n^{\underline{k}} \cdot t^{-n}) &= k! \cdot t^{-(n-k)}, \\ \Theta^k(n^{\underline{k}} \cdot t^n) &= k! \cdot (t^2-1)^k \cdot t^{n-k}, \\ \Theta^k(n^{\underline{k}} \cdot t^{-n}) &= k! \cdot (-1)^k \cdot (t^2-1)^k \cdot t^{-(n-k)}. \end{aligned} \tag{2.8}$$

Proof Commutativity follows from a straightforward computation from

$$\theta_+ \circ \theta_-(a_n) = \theta_- \circ \theta_+(a_n) = t \cdot (a_n + a_{n-2}) - (t^2+1)a_{n-1}.$$

The rules (2.6) are also easily verified, and they yield (2.7) by induction on k. Finally, the first two actions in (2.8) are obtained by setting $l = k$ in (2.7), while the action of Θ is obtained due to commutativity, by setting $\Theta^k = \theta_-^k \circ \theta_+^k$ or $\Theta^k = \theta_+^k \circ \theta_-^k$, depending on whether the operand is t^n or t^{-n}.

Remark 2.1 Note that $k+1$ consecutive values of a_n are necessary for applying $\theta_\pm^k$, while Θ^k requires $2k+1$ consecutive values.

We shall call the set $\{-1, 0, 1\}$ the admissible solutions. The task we pursue is to improve our estimates on pairs p, q for which the system (2.2) has no other solutions except (2.4). In particular, we are concerned with $p < q$, since Proposition 2.1 deals already with $p > q$. We shall use the fact on which the proof of Proposition 2.1 relies: $(\zeta^k)_{k=1}^{p-1}$ forms a basis of the algebra $\mathbb{Z}[\zeta]/(q\mathbb{Z}[\zeta])$ and this allows us to consider (2.2) as a linear system modulo q. Concretely, the coefficients of $\zeta^k - \overline{\zeta}^k$ in that equation must vanish, for $k = 1, 2, \cdots, \frac{p-1}{2}$. Let $0 < \nu < \frac{p-1}{2}$ be the value for which $\nu \equiv q \bmod p$ or $\nu \equiv -q \bmod p$. Then, with δ_{ij} the Kronecker δ, the previous observation yields the equations

$$-\delta_{\nu,k} \cdot \varphi(t) \equiv \sum_{j \geq 0; jp+k<q} \frac{t^{k+pj} - t^{2-(k+pj)}}{pj+k}$$

$$- \sum_{j\geq 0; jp+(p-k)<q} \frac{t^{p-k+pj} - t^{2-(p-k+pj)}}{p-k+jp} \bmod q. \tag{2.9}$$

The index value ν plays a singular role in the equations above: First, it is the only index for which the equations are not homogeneous. Second, the number of terms in the sums in the right-hand side changes between $0 < k < \nu$ and $\frac{p}{2} > k > \nu$. In these two intervals, (2.9) yields homogeneous equations which manifest itself in the vanishing of polynomials of fixed degree in k. This suggests the use of the difference operators defined above. Let $5 \leq p < q$ be primes. We shall take the approach of choosing either the interval $0 < k < \nu$ or $\nu < k < \frac{p}{2}$, whichever has more elements: In that interval, (2.9) translates into polynomial equations of the type $f_q(k;t) = 0$. Having a contiguous interval on which this equation holds, one can use the iteration of Θ in order to reduce the degree in k of the polynomial f_q. We have thus to distinguish the cases $\nu < \frac{p}{4}$ and $\nu > \frac{p}{4}$.[2]

Proposition 2.2 *Let $5 \leq p < q$ be primes such that* (2.2) *holds and ν be defined above. Suppose that $\nu > \frac{p}{4}$. If in addition, $q < \frac{p(p-12)}{16}$, then* (2.4) *holds.*

Proof Let $n = \lfloor \frac{q}{p} \rfloor$. The equation (2.9) yields on the interval $0 < k < \nu$:

$$\sum_{0\leq j\leq n} \frac{t^{k+pj} - t^{2-(k+pj)}}{pj+k} \equiv \sum_{0\leq j<n} \frac{t^{p-k+pj} - t^{2-(p-k+pj)}}{p-k+jp} \bmod q. \tag{2.10}$$

After eliminating denominators, this yields a polynomial equation

$$\begin{aligned} &(-1)^n k^{2n} \cdot \sum_{0\leq j\leq n} (t^{k+pj} - t^{2-(k+pj)}) + O(k^{2n-1}) \\ \equiv\ &(-1)^{n-1} k^{2n} \cdot \sum_{0\leq j<n} (t^{p-k+pj} - t^{2-(p-k+pj)}) + O(k^{2n-1}) \bmod q. \end{aligned}$$

In order to eliminate the lower order terms in k, we may take Θ^{2n} on both sides of the congruence. This requires at least $2(2n) + 1$ contiguous points, so $1 \leq k - 2n < k + 2n < \frac{p}{4}$, which means $2(2n) + 1 < \frac{p}{4}$. If this is provided, the equation reduces, after simplifying by $(-1)^n \cdot (2n)! \cdot (1 - t^2)^{2n}$, to

$$\begin{aligned} &\sum_{0\leq j\leq n} (t^{k+pj-2n} - t^{2-(k+pj-2n)}) \\ &+ \sum_{0\leq j<n} (t^{p-k+2n+pj} - t^{2-(p+2n-k+pj)}) \equiv 0 \bmod q. \end{aligned} \tag{2.11}$$

If $t \notin \{-1, 0, 1\}$, then we can apply θ_+ and θ_- independently to the above congruence. This yields

$$0 \equiv \sum_{0\leq j\leq n} t^{2-(k+pj-2n)} - \sum_{0\leq j<n} t^{p-k+2n+pj} \bmod q$$

[2] One may also take the approach of considering the whole interval $0 < k < \frac{p}{2}$. In this case, the polynomials $f_q(k;t)$ change the degree and shape when k passes the "singular" value $k = \nu$. The computations become more intricate, for a gain of a factor at most 2. We choose to analyse here the simpler approach.

and

$$0 \equiv \sum_{0\le j\le n} t^{k+pj-2n} - \sum_{0\le j<n} t^{2-(p+2n-k+pj)} \bmod q.$$

Upon multiplication by the lowest power of t, we obtain

$$0 \equiv \sum_{0\le j\le n} t^{pj} - \sum_{0\le j<n} t^{p(n+1)-2+pj} \bmod q, \quad 0 \equiv \sum_{0\le j\le n} t^{pn+pj-2} - \sum_{0\le j<n} t^{pj} \bmod q. \qquad (2.12)$$

Adding the two congruences, we obtain $t^{pn} \equiv -t^{pn-2} \bmod q$ with the solutions $t \equiv 0$ and $t^2 \equiv -1 \bmod q$. The first solution is admissible. We reinsert $t^2 \equiv -1 \bmod q$ in (2.11), using the fact that $t^m \equiv (-1)^m t^{-m} \bmod q$ for all m. This yields, after some computations,

$$(t^k + t^{-k}) \cdot \Big(1 + \sum_{j=1}^{n} t^{pj}(1+(-1)^j)\Big) \equiv 0 \bmod q.$$

The inner sum is

$$1 + 2\sum_{0<2l\le n} (-1)^l = 1 + 2(-1+1-1\cdots+(-1)^{[\frac{n}{2}]}) = \begin{cases} 1, & \text{if } \left[\dfrac{n}{2}\right] \equiv 0 \bmod 2, \\ -1, & \text{otherwise,} \end{cases}$$

and the previous condition thus becomes $\pm(t^k + t^{-k}) \equiv 0 \bmod q$, and since $t \not\equiv 0$, it follows that $(-1)^k + 1 \equiv 0 \bmod q$. It suffices to take k even, to obtain $t \equiv 0 \bmod q$, again, an admissible solution.

We now verify the conditions necessary for our derivation. For the final application of $\theta_\pm$ and the condition that k be even, we need

$$2n+1 \le k \le \frac{p}{4} - (2n+1),$$

which is satisfied by the even value $k = 2(n+1)$, provided that $4n+3 < \frac{p}{4}$. On the other hand, we find from the definition of ν and the fact that $\nu > \frac{p}{4}$, that $p(4n+3) > 4q$, and thus

$$\frac{p^2}{4} > p(4n+3) > 4q$$

as claimed. On the other hand, we find from the definition of n that if $q < \frac{p(p-12)}{16}$, then

$$4n+3 \le 4\left\lfloor \frac{q}{p} \right\rfloor + 3 < 4\frac{p-12}{16} + 3 = \frac{p}{4}$$

as claimed.

Proposition 2.3 *Let $5 \le p < q$ be primes such that* (2.2) *holds and ν be defined above. Suppose that $\nu < \frac{p}{4}$ and $q < \frac{p(p-12)}{16}$. Then* (2.4) *holds.*

Proof The proof of this proposition follows the same line as the previous one, but encounters a few particular obstructions. We shall let

$$n = \begin{cases} \left\lfloor \dfrac{q}{p} \right\rfloor - 1, & \text{if } (q \bmod p) < \dfrac{p}{4}, \\ \left\lfloor \dfrac{q}{p} \right\rfloor, & \text{if } (q \bmod p) > \dfrac{3p}{4}. \end{cases}$$

The equation (2.9) yields now on the interval $\nu < k < \frac{p}{4}$:

$$\sum_{0\leq j\leq n} \frac{t^{k+pj} - t^{2-(k+pj)}}{pj+k} \equiv \sum_{0\leq j\leq n} \frac{t^{p-k+pj} - t^{2-(p-k+pj)}}{p-k+jp} \bmod q. \tag{2.13}$$

Note that there are equally many terms in the sums of both sides of the above congruence, unlike the case of the previous proposition. If $t \notin \{-1, 0, 1\}$, this perpetuates down to the analogue of (2.12), in which the two congruences become identical:

$$0 \equiv \sum_{0\leq j\leq n} t^{pj} + \sum_{0\leq j\leq n} t^{p(n+1)+pj-2} \bmod q. \tag{2.14}$$

If $t^p \equiv 1 \bmod q$, we get $(n+1)(t^2+1) \equiv 0 \bmod q$ and $t^2 \equiv -1 \bmod q$, hence $1 \equiv t^p \equiv (-1)^{\frac{p-1}{2}} t \bmod q$, showing that t is admissible; so we can assume $t^p \not\equiv 1 \bmod q$. Then

$$t^{p(n+1)} \equiv 1 \bmod q \quad \text{or} \quad t^{p(n+1)} \equiv -t^2 \bmod q. \tag{2.15}$$

Note that this condition is equivalent to applying any of $\theta_+\Theta^{2n+1}$ or $\theta_-\Theta^{2n+1}$ to the original system (2.9).

In order to arrive at a contradiction, we shall have to consider lower order terms in k. Let

$$\sigma_j = t^{k+pj} - t^{2-(k+pj)} \quad \text{and} \quad \tau_j = t^{p-k+pj} - t^{2-(p-k+pj)}.$$

With some additional work, the first congruence yields, after elimination of denominators:

$$\begin{aligned} &\sum_{0\leq j\leq n} \sigma_j \cdot (k^{\underline{2n+1}} - ((n+j+1)p - (2n+1)n) \cdot k^{\underline{2n}}) \\ &+ \sum_{0\leq j\leq n} \tau_j \cdot (k^{\underline{2n+1}} - ((n-j)p - (2n+1)n) \cdot k^{\underline{2n}}) + O(k^{2n-1}) \equiv 0 \bmod q. \end{aligned} \tag{2.16}$$

We assume that the first congruence of (2.15) is satisfied. Then

$$\sum_{0\leq j\leq n} t^j = \frac{1 - t^{p(n+1)}}{1 - t^p} \equiv 0 \bmod q$$

and

$$\sum_{0\leq j\leq n} t^{-j} = \frac{1 - t^{-p(n+1)}}{1 - t^{-p}} \equiv 0 \bmod q.$$

Herewith, (2.16) reduces to

$$\begin{aligned} &\sum_{0\leq j\leq n} (t^{k+pj} - t^{2-(k+pj)}) \cdot (-jp) \cdot k^{\underline{2n}} \\ &+ \sum_{0\leq j\leq n} (t^{p-k+pj} - t^{2-p+k-pj}) \cdot (jp) \cdot k^{\underline{2n}} + O(k^{2n-1}) \equiv 0 \bmod q. \end{aligned}$$

Applying Θ^{2n} to the above and after simplifying by $(2n)! \cdot (1-t^2)^{2n}$, we get

$$\sum_{0\leq j\leq n} (t^{k-2n+pj} - t^{2-k+2n-pj}) \cdot (-jp)$$

$$+\sum_{0\le j\le n}(t^{p-k+2n+pj}-t^{2-p+k-2n-pj})\cdot(jp)\equiv 0 \bmod q.$$

Then we have

$$(t^{k-2n}-t^{p-k+2n})\cdot\sum_{0\le j\le n}t^{pj}j\equiv(t^{2-k+2n}-t^{2-p+k-2n})\cdot\sum_{0\le j\le n}t^{-pj}j \bmod q,$$

hence

$$(t^{k-2n}-t^{p-k+2n})\cdot\frac{1+n}{t^p-1}\equiv(t^{k-2n}-t^{p-k+2n})\cdot(-t^{2-p})\cdot\frac{1+n}{t^{-p}-1} \bmod q,$$
$$t^{k-2n}-t^{p-(k-2n)}\equiv t^2(t^{k-2n}-t^{p-k+2n}) \bmod q.$$

Since $t^2\equiv 1 \bmod q$ leads to admissible solutions, it remains that $t^p\equiv t^{2(k-2n)} \bmod q$, which must hold for instance for two successive values of k. Hence, by dividing the corresponding congruences, we get $t^2\equiv 1 \bmod q$, which has only admissible solutions.

Now we claim that

$$t^{p(n+1)}\not\equiv -t^2 \bmod q. \tag{2.17}$$

If not, we have

$$\sum_{0\le j\le n}t^{pj}\equiv\frac{1-t^{p(n+1)}}{1-t^p}\equiv\frac{1+t^2}{1-t^p} \bmod q$$

and

$$\sum_{0\le j\le n}t^{-pj}=\frac{1-t^{-p(n+1)}}{1-t^{-p}}\equiv\frac{1+t^{-2}}{1-t^{-p}} \bmod q.$$

This time, (2.16) is reduced to

$$\Big(t^k\cdot\sum_{0\le j\le n}t^{pj}-t^{2-k}\sum_{0\le j\le n}t^{-pj}\Big)\cdot k^{\underline{2n+1}}$$
$$+\Big(t^{p-k}\cdot\sum_{0\le j\le n}t^{pj}-t^{2-p+k}\sum_{0\le j\le n}t^{-pj}\Big)\cdot k^{\underline{2n+1}}+O(k^{2n})\equiv 0 \bmod q.$$

Then we have

$$\Big(t^k\cdot\frac{1+t^2}{1-t^p}-t^{2-k}\frac{1+t^{-2}}{1-t^{-p}}\Big)\cdot k^{\underline{2n+1}}$$
$$+\Big(t^{p-k}\cdot\frac{1+t^2}{1-t^p}-t^{2-p+k}\frac{1+t^{-2}}{1-t^{-p}}\Big)\cdot k^{\underline{2n+1}}+O(k^{2n})\equiv 0 \bmod q.$$

It follows that

$$\Big(\frac{t^2+1}{1-t^p}\Big)\cdot(t^k+t^{p-k})\cdot 2\cdot k^{\underline{2n+1}}+O(k^{2n})\equiv 0 \bmod q.$$

We apply Θ^{2n+1} to the above and get

$$\Big(\frac{t^2+1}{1-t^p}\Big)\cdot(t^{k-2n-1}+t^{p-k+2n+1})\cdot 2\cdot(2n+1)!\cdot(t^2-1)^{2n+1}\equiv 0 \bmod q.$$

We have excluded $t^2 \equiv -1 \bmod q$ and the vanishing of the second factor leads to $t^p \equiv t^{2(k-2n-1)}$, which implies like before, that $t^2 - 1 \equiv 0 \bmod q$, thus only admissible solutions are possible. This confirms the claim (2.17).

We finally have to derive the inequality between p and q, for which the proof above holds. The condition is that the interval $\left(\frac{p}{4}, \frac{p}{2}\right)$ contains sufficient contiguous points for applying both $\theta_{\pm}\Theta^{2n+1}$ and $\theta_{+}^2\Theta^{2n}$. This requires at least $2(2n+1)+1$ contiguous points, so we need $\frac{p}{4} < k - 2(n+1) < k + 2(n+1) + 1 < \frac{p}{2}$, which means $4n + 3 < \frac{p}{4}$. Note that by definition of n, if $q < \frac{p(p-12)}{16}$, then

$$4n + 3 \leq 4\left\lfloor \frac{q}{p} \right\rfloor + 3 < \frac{p}{4}.$$

This completes the proof of Proposition 1.1.

3 The Local Approach

We consider the cases when a solution to (1.3) has $x \equiv f \bmod q^2$, with $f \in \{-1, 0, 1\}$. Let $\mu := \frac{\overline{\rho}}{\rho}$ and note that $\delta := \overline{\alpha} \cdot \left(\frac{\mu-1}{1-\zeta^2}\right)^q \in \mathcal{O}(\mathbb{K})^{\times}$, as follows by the following computations:

$$\delta = \overline{\alpha} \cdot \left(\frac{\mu - 1}{1 - \zeta^2}\right)^q = -\zeta^q \overline{\alpha}\left(\frac{\rho - \overline{\rho}}{(\zeta - \overline{\zeta})\overline{\rho}}\right)^q = -\zeta^q \frac{\overline{\alpha}}{\overline{\rho}^q} \cdot \left(\frac{\rho - \overline{\rho}}{\zeta - \overline{\zeta}}\right)^q = -\zeta^q \overline{\varepsilon} \cdot \left(\frac{\rho - \overline{\rho}}{\zeta - \overline{\zeta}}\right)^q.$$

We note that $A - \overline{A} \equiv 0 \bmod (\zeta - \overline{\zeta})$ for arbitrary $A \in \mathbb{Z}[\zeta]$, so consequently $\frac{\rho - \overline{\rho}}{\zeta - \overline{\zeta}} \in \mathbb{Z}[\zeta]$. In view of Lemma 1.2, we have

$$\alpha - \overline{\alpha} = \zeta - \overline{\zeta} = \varepsilon \cdot (\rho - \overline{\rho}) \cdot \left(\frac{\rho^q - \overline{\rho}^q}{\rho - \overline{\rho}}\right),$$

hence

$$\varepsilon^{-1} = \left(\frac{\rho - \overline{\rho}}{\zeta - \overline{\zeta}}\right) \cdot \left(\frac{\rho^q - \overline{\rho}^q}{\rho - \overline{\rho}}\right);$$

the two factors on the right-hand side are integral, and their product is a unit, so both must be units, individually[3]. We note also that $\overline{\alpha} \cdot \left(\frac{\mu-1}{\pi}\right)^q \in \mathcal{O}(\mathbb{K})^{\times}$ for every $\pi \in \wp$ that generates the prime above p. We shall adapt various values for π to the different values of f. We may write $\delta(1 - \zeta^2)$ in the above case, to indicate that the unit is defined with respect to the choice $\pi = 1 - \zeta^2$.

We shall compute a q-adic development of δ and its norm, and compare this to the value of one which should be the result, since we have seen that δ is a unit.

3.1 The case $f = 0$

Let $x = q^l z$ with $z \in \mathbb{Z}$, $p \nmid z$ and $l \geq 2$; we have

$$\mu^q = \zeta^{2q} \frac{1 - \overline{\zeta}^q x}{1 - \zeta^q x},$$

[3] We mention for later use, that the decomposition so far is independent of the value of x.

so

$$\mu = \zeta^2 \cdot \frac{1 - \overline{\zeta}^q q^{l-1} z}{1 - \zeta^q q^{l-1} z} + O(q^{2(l-1)}) = \zeta^2 \cdot (1 + q^{l-1} z \cdot (\zeta^q - \overline{\zeta}^q)) + O(q^{2(l-1)}),$$

hence

$$\frac{\mu - 1}{\zeta^2 - 1} = 1 + \frac{x}{q} \cdot \frac{\zeta^q (1 - \overline{\zeta}^{2q})}{1 - \overline{\zeta}^2} + O(q^{2(l-1)}),$$

$$\delta = 1 + x \cdot \Big(\frac{\zeta^q (1 - \overline{\zeta}^{2q})}{1 - \overline{\zeta}^2} + \overline{\zeta}^q \Big) + O(q^{2(l-1)}).$$

By defining $B = \frac{\zeta^q (1 - \overline{\zeta}^{2q})}{1 - \overline{\zeta}^2}$, and taking the norm of $\delta(\zeta^2 - 1)\overline{\zeta}$, we see that $\mathbf{N}(\delta) = 1$ implies $\mathbf{Tr}(B - \overline{\zeta}^q) \equiv 0 \bmod q$ and so $\mathbf{Tr}(B) \equiv -1 \bmod q$. Let $q \equiv r \bmod p$, where $1 \le r \le p - 1$; since

$$\overline{B} = \overline{\zeta}^r (1 + \zeta^2 + \cdots + \zeta^{2(r-1)}) = \zeta^{-r} + \zeta^{2-r} + \cdots + \zeta^{r-2},$$

we see that for odd r, none of the terms in this sum carries the exponent zero. There are r terms in the sum B, so

$$\mathbf{Tr}(B - \zeta^{-q}) = \begin{cases} -r + 1, & \text{if } r \text{ is odd,} \\ -r + 1 + p, & \text{otherwise.} \end{cases} \tag{3.1}$$

If r is even, which only happens when $q > p$, then $0 < -r + 1 + p < q$ and hence $q \nmid (-r + 1 + p)$. If r is odd, the vanishing condition requires $r = 1$ so $q \equiv 1 \bmod p$.

In this case, we consider some higher order terms:

$$\mu = \zeta^2 (1 - q^l z \overline{\zeta}^q)^{\frac{1}{q}} (1 - q^l z \zeta^q)^{-\frac{1}{q}}. \tag{3.2}$$

By expanding (3.2) under the condition $\zeta^q = \zeta$, we obtain

$$\begin{aligned} \mu &= \zeta^2 \Big(1 - q^{l-1} z \overline{\zeta} + q^{2l-2} z^2 \frac{1 - q}{2} \overline{\zeta}^2 \Big) \cdot \Big(1 + q^{l-1} z \zeta + q^{2l-2} z^2 \frac{1 + q}{2} \zeta^2 \Big) + O(q^{3l-3}) \\ &= \zeta^2 \Big(1 + q^{l-1} z (\zeta - \overline{\zeta}) + q^{2l-2} z^2 \frac{\overline{\zeta}^2 (1 - q) - 2 + \zeta^2 (1 + q)}{2} \Big) + O(q^{3l-3}), \\ \frac{\mu - 1}{1 - \overline{\zeta}^2} &= -\Big(1 + q^{l-1} z \frac{\zeta - \overline{\zeta}}{1 - \overline{\zeta}^2} + q^{2l-2} z^2 \frac{\overline{\zeta}^2 \frac{1-q}{2} - 1 + \frac{q+1}{2} \zeta^2}{1 - \overline{\zeta}^2} \Big) + O(q^{3l-3}), \end{aligned}$$

implying that

$$-\Big(\frac{\mu - 1}{1 - \zeta^2} \Big)^q = 1 + q^l z \frac{\zeta - \overline{\zeta}}{1 - \overline{\zeta}^2} + q^{2l-1} z^2 \Big(\frac{\overline{\zeta}^2 \frac{1-q}{2} - 1 + \frac{q+1}{2} \zeta^2}{1 - \overline{\zeta}^2} + \Big(\frac{\zeta - \overline{\zeta}}{1 - \overline{\zeta}^2} \Big)^2 \frac{q - 1}{2} \Big) + O(q^{2l}).$$

Hence

$$\delta = 1 + q^{2l-1} z^2 \Big(\frac{\overline{\zeta}^2 \frac{1-q}{2} - 1 + \frac{q+1}{2} \zeta^2}{1 - \overline{\zeta}^2} + \zeta^2 \frac{q - 1}{2} \Big) + O(q^{2l}).$$

Taking the norm of the last equality, we obtain

$$\mathbf{N}(\delta) = 1 + q^{2l-1} z^2 \cdot \mathbf{Tr} \Big(\frac{\overline{\zeta}^2 \frac{1-q}{2} - 1 + \frac{q+1}{2} \zeta^2}{1 - \overline{\zeta}^2} + \zeta^2 \frac{q - 1}{2} \Big) + O(q^{2l})$$

$$= 1 + q^{2l-1} z^2 \Big(- \frac{p-1}{2} \cdot \frac{1-q}{2} - \frac{p-1}{2} + \frac{q+1}{2} \cdot \frac{p-3}{2} + \frac{1-q}{2} \Big) + O(q^{2l})$$
$$= 1 + q^{2l-1} z^2 \frac{1-p}{2} + O(q^{2l}) \not\equiv \pm 1 \bmod q^{2l}.$$

We assumed that $q \equiv 1 \bmod p$, so the factor $\frac{1-p}{2} \not\equiv 0 \bmod q$ and consequently $\mathbf{N}(\delta) = 1 + Cq^{2l-1} + O(q^{2l})$, for an integer constant $C = z^2 \frac{1-p}{2} \not\equiv 0 \bmod q$. This is inconsistent with the fact that δ is a unit, which completes the proof of case $f = 0$ in Proposition 1.2.

4 Diophantine Approximation

Let $\Theta = \sum_{c \in P} m_c \sigma_c \in \mathbb{Z}_{\geq 0}[G]$ have the weight $w(\Theta) = \sum_{c \in P} n_c$. We define the following formal binomial series:

$$F_{n\sigma_c}(T) = (1 + \zeta^c T)^{\frac{n}{q}} = 1 + \sum_{k=1}^{\infty} \binom{\frac{n}{q}}{k} (\zeta^c T)^k \tag{4.1}$$

and

$$F_\Theta(T) = \prod_{c=1}^{p-1} F_{m_c \sigma_c}(T) := 1 + \sum_{k=1}^{\infty} a_k(\Theta) T^k \in \mathbb{K}[[T]],$$

where the coefficients $a_k(\Theta)$ are obtained by multiplying out the elementary series and rearranging in ascending order of the powers of the indeterminate T.

Alternatively, we may fix $g \in P$ as a generator of $\mathbb{F}_p^\times$ and fix $\sigma = \sigma_g \in G$, which is then a generator of the cyclic Galois group. We then write our generic group ring element as

$$\Theta = \sum_{j \in P} n_j \sigma^j = \sum_{j \in P} n_j \sigma_{g^j},$$

and the formal power series $F_\Theta(T) := \prod_{j \in P} F_{n_j \sigma_{g^j}}(T)$.

The series $F_\Theta\big(-\frac{1}{x}\big)$ are absolutely convergent in $\mathbb{C}$, for $|x| > 1$, and in particular for integers $x \in \mathbb{Z} \setminus \{-1, 0, 1\}$ and

$$\Big(F_{(1-\jmath)\sigma^d}\Big(-\frac{1}{x}\Big)\Big)^q = \mu^q,$$

so there exist exponents $a(d), a(d, \Theta) \in \mathbb{Z} \cap \big[-\frac{q-1}{2}, \frac{q-1}{2}\big]$ such that

$$\sigma^d\Big(\frac{\rho}{\overline{\rho}}\Big) = \xi^{a(d)} \cdot F_{(1-\jmath)\sigma^d}\Big(-\frac{1}{x}\Big) \quad \text{and} \quad \sigma^d\Big(\frac{\rho}{\overline{\rho}}\Big)^\Theta = \xi^{a(d,\Theta)} F_{\sigma^d(1-\jmath)\Theta}\Big(-\frac{1}{x}\Big);$$

we may also write $a(\Theta) = a(1, \Theta)$. We may estimate the power series by appealing to [16, Lemma 7], which leads to

$$W(\sigma^d \Theta) := \Big| F_{(1-\jmath)\sigma^d \Theta}\Big(-\frac{1}{x}\Big) - 1 \Big| < 3 \frac{w(\Theta)}{q|x|}. \tag{4.2}$$

By considering the action of Θ on $\sigma^d\big(\frac{\rho}{\overline{\rho}}\big)$, one verifies

$$a(d, \Theta) \equiv \sum_{i=1}^{p-1} n_i a(d+i) \bmod p. \tag{4.3}$$

Recall that $\sigma^{\frac{p-1}{2}}$ is the complex conjugation, hence $a(i+\frac{p-1}{2},\Theta) = -a(i,\Theta)$ for $1 \leq i \leq \frac{p-1}{2}$. Let $\nu_i = n_i - n_{i+\frac{p-1}{2}}$; by inserting this in (4.3), we obtain

$$a(d,\Theta) \equiv \sum_{i=1}^{\frac{p-1}{2}} \nu_i a(d+i) \bmod p, \quad d = 1,2,\cdots,p-1. \tag{4.4}$$

We call the system (4.4) over $\mathbb{F}_p$, the associated linear system of $\Theta \in \mathbb{Z}_{\geq 0}[G]$. The above considerations show that we may interpret the vector $\overrightarrow{A} := (a(d))_{d=1}^{\frac{p-1}{2}}$ as given and consider the ν_i as unknowns for a given constant vector $\overrightarrow{X}$. We shall impose certain conditions on the vector $\overrightarrow{X}$, which will determine n_i and herewith, a $\Theta \in \mathbb{Z}[G]$. For some $\tau \in \mathbb{N}$ with $1 \leq \tau \leq \frac{p-1}{2}$, we let the entries $X_j = 0$ for $j \leq \tau$ in the right-hand side vector of the system (4.4), meaning that we wish to find Θ, such that the exponents $a(\sigma^d\Theta) = 0$ for $d \leq \tau$. The remaining entries in $\overrightarrow{X}$ are free. We shall maximize τ subject to the condition that the homogeneous system built from the first τ equations in (4.4) has a nontrivial integer solution $\overrightarrow{\nu}$ with $\|\overrightarrow{\nu}\|_1 < 2$. The values X_d for $d > \tau$ will be determined by this solution.

Focusing herewith on the first τ equations, we consider the linear map with matrix $M = (a(d,i))_{d,i=1}^{\tau,\frac{p-1}{2}}$ and its action on the vectors in $V_2 = \{0,1\}^{\frac{p-1}{2}}$. For $v \in V_2$, we consider the image $w = Mv \in (\mathbb{Z}/(q\cdot\mathbb{Z}))^{\tau}$. By an application of the pigeonhole principle, we see that as soon as

$$2^{\frac{p-1}{2}} > q^{\tau} \Leftrightarrow \tau < \frac{p-1}{2\log_2(q)}, \tag{4.5}$$

there are two different vectors $v_1, v_2 \in V_2$ with identical image, so letting $v = v_1 - v_2$, we obtain an integer vector with entries in $\{-1,0,1\}$ and such that $Mv \equiv 0 \bmod q$. We may thus choose the value

$$\tau = \left[\frac{p-1}{2\log_2(q)}\right],$$

and the previous reasoning implies there is a vector $\overrightarrow{\nu}$ with entries in $\{-1,0,1\}$ which annihilates modulo q the first τ equations in (4.4). This vector defines a group ring element $\theta = \sum_{i=1}^{\frac{p-1}{2}} \nu_i\sigma^i \in \mathbb{Z}[G]$. It can be transformed into a positive element $\Theta \in \mathbb{Z}_{\geq 0}[G]$ as follows: We observe that in θ, the coefficients $\nu_i = 0$ for $i > \frac{p}{2}$. We derive the element $\Theta = \sum_{c\in P} n_c\sigma^c$ as follows: Set all $n_j = 0$ and then, for $c = 1,2,\cdots,\frac{p-1}{2}$, let $n_j = \nu_j$ if $\nu_j \geq 0$ and $n_{p-j} = -\nu_j$ otherwise. This defines Θ uniquely, and by the above, we know that $a(\sigma_d\Theta) = a(\sigma_d\theta) = 0$ for all $d \leq \tau$. The weight verifies $w(\Theta) = \sum_{i=1}^{\frac{p-1}{2}} |\nu_i| \leq \frac{p-1}{2}$.

4.1 The norm of $\left(\frac{\rho}{\overline{\rho}}\right)^{\Theta} - 1$

We move on to compute the norm of $\left(\frac{\rho}{\overline{\rho}}\right)^{\Theta} - 1$:

$$\mathbf{N}\Big(\Big(\frac{\rho}{\overline{\rho}}\Big)^{\Theta} - 1\Big) = \prod_{c\in P}\Big(\Big(\sigma^c\Big(\frac{\rho}{\overline{\rho}}\Big)\Big)^{\Theta} - 1\Big) = \prod_{c\in P}\Big(\xi^{a(c,\Theta)} \cdot F_{\sigma^d(1-\jmath)\Theta}\Big(-\frac{1}{x}\Big) - 1\Big) =: P_1 \cdot P_2,$$

where P_1 denotes the product of the factors with $a(c,\Theta) \neq 0$ and P_2 denotes the one with $a(c,\Theta) = 0$. Let the number of factors appearing in P_2 be $t(\Theta)$, so by construction, $2\tau \leq t(\Theta)$.

Then by (4.2), for each factor $\phi_c := \big(\xi^{a(c,\Theta)} \cdot F_{\sigma^d(1-\jmath)\Theta}\big(-\frac{1}{x}\big) - 1\big)$ appearing in P_1, we have

$$\begin{aligned}\Big|\xi^{a(c,\Theta)} \cdot F_{\sigma^d(1-\jmath)\Theta}\Big(-\frac{1}{x}\Big) - 1\Big| &\leq \Big|F_{\sigma^d(1-\jmath)\Theta}\Big(-\frac{1}{x}\Big) - 1\Big| + 1 \\ &\leq 3\frac{w(\Theta)}{q|x|} + 2 \leq 2\Big(1 + \frac{3(p-1)}{4q|x|}\Big);\end{aligned} \tag{4.6}$$

similarly, for each factor ϕ_c occurring in P_2, we have

$$\Big|F_{\sigma^d(1-\jmath)\Theta}\Big(-\frac{1}{x}\Big) - 1\Big| \leq 3\frac{w(\Theta)}{q|x|} \leq \frac{3(p-1)}{2q|x|}. \tag{4.7}$$

Let $q = cp\log(p)$ for some $c > 1$; we search for a minimal value for c, such that the existence of a solution to (1.3) to which we apply Θ as derived here, leads to a contradiction. If $p > 29$, then $\frac{3(p-1)}{2q} < \frac{1}{2}$, hence by (4.6)–(4.7), we obtain

$$\Big|\mathbf{N}\Big(\Big(\frac{\rho}{\overline{\rho}}\Big)^{\Theta} - 1\Big)\Big| = |P_1 \cdot P_2| < 3^{p-1-t(\Theta)} \cdot \Big(\frac{1}{|x|}\Big)^{t(\Theta)}.$$

Therefore

$$|\mathbf{N}((\rho - \overline{\rho})^{\Theta})| = |\mathbf{N}(\overline{\rho}^{\Theta})| \cdot \Big|\mathbf{N}\Big(\Big(\frac{\rho}{\overline{\rho}}\Big)^{\Theta} - 1\Big)\Big| < |y|^{w(\Theta)} \cdot 3^{p-1-t(\Theta)} \cdot \Big(\frac{1}{|x|}\Big)^{t(\Theta)}; \tag{4.8}$$

note that $(\rho - \overline{\rho})$ is an associate of $(\zeta - \overline{\zeta})$, hence

$$\mathbf{N}((\rho - \overline{\rho})^{\Theta}) = p^{w(\Theta)},$$

and (4.8) is equivalent to

$$1 < \Big(\frac{|y|}{p}\Big)^{w(\Theta)} \cdot 3^{p-1-t(\Theta)} \cdot \Big(\frac{1}{|x|}\Big)^{t(\Theta)}.$$

Recall that we have the known result $|y| \geq 2p+1$, $t(\Theta) \geq 2\tau$ and $w(\Theta) \leq \frac{p-1}{2}$, hence

$$1 < \Big(\frac{|y|}{p}\Big)^{w(\Theta)} \cdot 3^{p-1-t(\Theta)} \cdot \Big(\frac{1}{|x|}\Big)^{t(\Theta)} \leq \Big(\frac{|y|}{p}\Big)^{\frac{p-1}{2}} \cdot 3^{p-1-2\tau} \cdot \Big(\frac{1}{|x|}\Big)^{2\tau}. \tag{4.9}$$

Note that

$$y^q = \frac{x^p - 1}{x - 1} < \frac{4}{3}|x|^{p-1};$$

if $p > 9$, we obtain from (4.9)

$$1 < \Big(\frac{4}{3}\Big)^{\frac{p-1}{2q}} \cdot \Big(\frac{9}{p}\Big)^{\frac{p-1}{2}} \cdot |x|^{\frac{(p-1)^2}{2q} - 2\tau} \leq (|x|^{\frac{p-1}{2q} - \frac{2\tau}{p-1}})^{p-1}, \tag{4.10}$$

which is impossible if

$$\tau \geq \frac{(p-1)^2}{4q}. \tag{4.11}$$

We see that if there is an integer τ with upper bound given by (4.5) and lower bound given by (4.11), then (4.9) leads to a contradiction, derived from the assumption that (1.3) has a solution. In other words, provided that q is sufficiently large for the interval

$$\frac{(p-1)\log 2}{2\log(q)} > \tau \geq \frac{(p-1)^2}{4q} \tag{4.12}$$

to contain an integer, we may conclude that there are no solutions of (1.3) for such p, q, if $q \nmid h_p^-$. If the simple inequality

$$\frac{(p-1)\log(2)}{2\log(q)} > \frac{(p-1)^2}{4q} + 1 \tag{4.13}$$

holds, then the given interval necessarily contains an integer. Multiplying both sides by $\frac{4q}{(p-1)^2}$, we get the equivalent inequality

$$\frac{q\log(4)}{(p-1)\log(q)} > 1 + \frac{4q}{(p-1)^2}. \tag{4.14}$$

For $c < 2.7$, we use $\log(q) < 1 + \log(p) + \log\log(p)$ and $\frac{4q}{(p-1)^2} \leq \frac{c\log(p)}{p-2}$, so

$$\frac{q\log(4)}{(p-1)\log(q)} > \frac{c\log(4)}{1 + \frac{1+\log\log(p)}{\log(p)}},$$

hence the condition is fulfilled if

$$c \cdot \left(\frac{\log(4)}{1 + \frac{1+\log\log(p)}{\log(p)}} - \frac{4\log(p)}{p-2} \right) > 1. \tag{4.15}$$

Thus, defining $C(p)$ to be the inverse of the cofactor of c and $M'(p) = C(p)p\log(p)$, we recover the definitions in (1.4), in which $C(p)$ is a function with asymptotic value $\lim\limits_{p\to\infty} C(p) = \frac{1}{\log(4)} < 0.75$. Note that for $p = 29$ we have $M(p) < M'(p)$, while asymptotically we obviously have $\frac{M(p)}{M'(p)} \to \infty$. With some elementary analysis, one also verifies that the difference of the two functions only has one zero on $x > 29$. This is determined numerically to be $x \in (113, 127)$, which confirms the last statement of Proposition 1.3.

Finally, for $29 \leq p \leq 113$, we need to check if there is any prime $M(p) < q < M'(p)$, such that the pair (p, q) satisfies the previously derived conditions. Note that there always is a loss, when deriving a general condition that excludes all pairs verifying it. For concrete numbers, one may still show that no solution exists, by some concrete verification. We have done this using PARI and the following three conditions:

(1) Existence of an integer τ satisfying the condition of (4.12).

(2) Using the conditions on ν in Proposition 2.2 and Proposition 2.3 that provide inequalities between ν and $4n + 3$.

(3) Showing that the only possible values of t in (2.10) $\big($if $\nu > \frac{p}{4}\big)$ and (2.13) $\big($if $\nu < \frac{p}{4}\big)$ belong to $\{0, \pm 1\}$.

Indeed, the combination of the three criteria helped eliminate all remaining cases. It turns out that (3) is the most powerful condition. We herewith know that the only possible cases remaining verify $x \equiv \pm 1 \bmod q^2$. In order to complete the proof of Theorem 1.2 we shall eliminate these cases by using a local power series development method.

4.2 Local approximation in the cases $x \equiv \pm 1 \bmod q^2$

In this section, we use local approximation for eliminating these two remaining cases. Using local power series expansions, we prove the following proposition.

Proposition 4.1 *The equation* (1.3) *has no solutions with* $x \equiv f \bmod q^{l+1}$ *for* $p > 23$ *and* $f \in \{-1, 1\}$ *and an integer* $l \geq 1$.

The proof of the proposition covers the rest of this section. We give first a brief description of our approach, which starts from the assumption that $(x, y; p, q)$ is a solution with odd primes p, q and $x \equiv \pm 1 \bmod q^2$.

4.2.1 The μ-map

Let

$$D_G = \Big\{ t \in \mathbb{Z}[G]^\times : t = \sum_{c \in P} n_c \sigma_c \text{ with } n_c \in \{0,1\};\ n_c + n_{p-c} \leq 1;\ n_c \cdot n_{p-c} = 0 \Big\}. \tag{4.16}$$

Note that the set D_G is G-stable, since for $t \in D_G$, the conjugates σt will also fulfill the defining conditions of D_G.

Then, for $t \in D_G$, the product $Z(t) := y \cdot \left(\frac{\rho}{\bar{\rho}}\right)^t \in \mathbb{Z}[\zeta]$, as follows from the definition of D_G together with the fact that the conjugates of ρ are pairwise coprime and have norm y. We note that the map

$$\Delta : D_G \hookrightarrow \mathcal{O}^\times(\mathbb{K}), \quad t \mapsto Z(t) \tag{4.17}$$

is indeed injective. This follows, for instance, by induction on the weight of t, from the coprimality of the conjugates of ρ.

Moreover, in the cases of interest, when $x \equiv \pm 1 \bmod q^{l+1}$, there is a convergent q-adic binomial series

$$\mu(t) = \left(\frac{\rho}{\bar{\rho}}\right)^t = 1 + \sum_{n=1}^{\infty} a_n(t) q^{ln}, \quad a_n(t) \in \mathbb{Z}[\zeta] \tag{4.18}$$

and

$$\sigma(a_n(t)) = a_n(\sigma t), \quad \forall \sigma \in G. \tag{4.19}$$

Here we assume that the coefficients $a_n(t)$ are elements of the minimal set of representatives $W = \big\{ \sum_{c \in P} w_c \zeta^c : -\frac{q^l-1}{2} \leq w_c \leq \frac{q^l-1}{2} \big\}$ for $\mathbb{Z}[\zeta]/q^l\mathbb{Z}[\zeta]$; thus, the binomial power series has been reordered in order for the coefficients to match this condition. Thus, $Z(t) = y \cdot \mu(t)$ and the coefficients of the power series are Galois covariant, by (4.19). Note also that $\mu(t) \notin \mathbb{Z}[\zeta]$ and herewith, the power series (4.18) has infinitely many nonvanishing coefficients.

We fix a $\theta \in D_G$ such that $\sigma_a\theta \neq \sigma_b\theta$ for $a \neq b$, so $|G\theta| = p - 1$. Let $Q = q^N$ for some large N, such that

$$Q > |py|^3.$$

The series (4.18) can be regrouped in terms of powers of Q, with some coefficients $b_m = b_m(\theta) \in W_Q$, where

$$W_Q = \Big\{ \sum_{c\in P} w_c \zeta^c : -\frac{Q-1}{2} \leq w_c \leq \frac{Q-1}{2} \Big\}$$

is a set of representatives for $\mathbb{Z}[\zeta]/Q\mathbb{Z}[\zeta]$:

$$\mu(\theta) = \Big(\frac{\rho}{\overline{\rho}}\Big)^{\theta} = 1 + \sum_{n=1}^{\infty} b_n Q^n, \quad b_n \in W_Q. \tag{4.20}$$

The coefficients are also Galois covariant, so $b_n(\sigma\theta) = \sigma(b_n(\theta))$. Moreover, $\|b_n\|_1 \leq \frac{Q-1}{2}$.

4.2.2 Scalar products and various representations of field elements

Let $\sigma = \sigma_g \in G$ be a generator of this cyclic group. We endow that number field $\mathbb{K} = \mathbb{Q}[\zeta]$ with the basis $\mathcal{Z} = \{e_c = \sigma^c(\zeta) : c \in P\}$, as a $\mathbb{Q}$-vector space; this is at the same time the power normal basis of the integers in $\mathbb{Z}[\zeta]$ and a fortiori, integral numbers are represented by vectors of rational integer coefficients with respect to this basis. We let $V = \mathbb{Q}^{p-1}$ and denote the coefficient map of linear algebra by

$$\kappa : \mathbb{K} \to V, \quad \alpha = \sum_{c=1}^{p-1} a_c \sigma^c(\zeta) \mapsto \overrightarrow{a} = (a_1, a_2, \cdots, a_{p-1}) \in V.$$

It will also be convenient to introduce a notation for the vectors of conjugate elements of $\mathbb{K}$, so let

$$\mathbb{K}_G = \{(\sigma^c(x))_{c\in P} \in \mathbb{K}^{p-1} : x \in \mathbb{K}\} \subset \mathbb{K}^{p-1},$$

and let $\nu : x \mapsto (\sigma^c(x))_{c\in P}$ be the associated embedding of $\mathbb{K}$ in $\mathbb{K}_G$.

Let $x = \sum_{i=1}^{p-1} x_c \zeta^c, y = \sum_{i=1}^{p-1} y_c \zeta^c \in \mathbb{K}$. Then

$$\mathbf{Tr}(x \cdot y) = \mathbf{Tr}\Big(\sum_{j=1}^{p-1} x_j y_{p-j} + \sum_{m=1}^{p-1} \zeta^m \sum_{j+k \equiv m \bmod p-1} x_j y_k \Big) = p \cdot \sum_{j=1}^{p-1} x_j y_{p-j} - \mathbf{Tr}(x) \cdot \mathbf{Tr}(y).$$

We observe that the trace has a particularly simple form, if for instance $\mathbf{Tr}(y) = 0$. Thus we have the following lemma.

Lemma 4.1 *Notations being like above, let $x, y \in \mathbb{K}$ and assume that $\mathbf{Tr}(y) = 0$. Then*

$$\mathbf{Tr}(x \cdot \overline{y}) = p \cdot \langle \kappa(x), \kappa(y) \rangle,$$

where $\langle \cdot, \cdot \rangle$ is the standard scalar product on $\mathbb{Q}^{p-1}$. The right-hand side does not depend upon simultaneous permutations of the coefficients, so the coefficient map may also be κ.

We note that in the above trace we had to take the complex conjugate of y, in order to obtain the standard scalar product on the right-hand side. At the same time, the left-hand side becomes a non-degenerate Hermitian bilinear form.

Let θ be fixed like above and $\mathcal{T} = G\theta$ be its orbit. We map the elements of $\mathcal{T}$ in the following way: A D-vector $\mathfrak{W}$ is a triple of maps

$$\begin{aligned} &\mathfrak{W}_c = \mathcal{T} \to \mathbb{K}_G, \quad \mathfrak{W}_s = \mathcal{T} \to V, \quad \mathfrak{W}_f = \mathcal{T} \to \mathbb{K}, \\ &\nu^{-1} : \mathbb{K}_G \to \mathbb{K}, \quad \kappa : \mathbb{K} \to V, \quad \Phi = \kappa \circ \nu^{-1} : \mathbb{K}_G \to V, \\ &\quad (Gx) \mapsto x \mapsto (\kappa(x)), \end{aligned} \tag{4.21}$$

such that the diagram commutes, as illustrated in the diagram below

$$\begin{tikzcd} & \mathcal{T} \arrow[ld, "\mathfrak{W}_c"'] \arrow[d, "\mathfrak{W}_f"] \arrow[rd, "\mathfrak{W}_s"] & \\ \mathbb{K}_G \arrow[r, "\nu^{-1}"] \arrow[rr, bend right, "\Phi"'] & \mathbb{K} \arrow[r, "\kappa"] & V \end{tikzcd}$$

Thus the vector can be given by any of its three representations, and the other two follow. We give some examples of D-vectors that we shall intensively use the following example.

Example 4.1 (a) The coefficients $b_n(t)$ in (4.20) give rise to a D-vector $\mathfrak{W}(b_n)$ presented in $\mathbb{K}$ by the map $\mathfrak{W}_f : t \in \mathcal{T} \mapsto b_n(t)$. We denote the image $(\mathfrak{W}(b_n)) =: \mathbf{b}_n \in V$.

(b) We present here the standard basis of V as a set of D-vectors

$$\Delta = \{\mathfrak{D}(i) : i = 1, 2, \cdots, p-1\}, \quad \text{given by } \mathfrak{D}(i)_f :\mapsto \sigma^i(\zeta).$$

The standard basis of V arises also as

$$\mathcal{E} = \{e_i : i \in P\} = (\mathfrak{D}(i)_c(\mathcal{T}))_{i=1}^{p-1} \subset V.$$

(c) Let $\ell \in \mathbb{Z}[\zeta]$ be an indeterminate. It will be useful to impose the condition $\mathbf{Tr}(\ell) = 0$. For this we define the D-vector $\mathfrak{U}$ induced by $\mathfrak{W}_s(\mathfrak{U}) = U := (1, 1, \cdots, 1)$. The scalar product $\langle U, \kappa(\ell)\rangle = \mathbf{Tr}(\ell)$ and thus $\mathbf{Tr}(\ell) = 0$ if and only if $\kappa(\ell) \perp U$.

For $\vec{v} \in V$ we define the norm to be the 1-norm $\|\vec{v}\| = \|\vec{v}\|_1 = \max_{c \in P}(|v_c|)$ and for $w \in \mathbb{K}_G$ we define $\|w\| = \|\Phi(w)\|_1$.

4.2.3 Strategy of proof

The principle of our proof is the following: For the unknown ℓ we impose the conditions $\ell \perp U$ and $\ell \perp b_1$ and define

$$\mathfrak{d} := y \cdot \mathbf{Tr}(\mu(\theta) \cdot \overline{\ell}) \in \mathbb{Z}. \tag{4.22}$$

The choice of ℓ should assure that $\mathfrak{d} \neq 0$. Since the complex absolute value $|\mu(\sigma_c\theta)| = 1$ for all $c \in P$, assuming that $\|\ell\| \leq L$ for some $L \in \mathbb{R}_{>0}$, we have the upper bound

$$|\mathfrak{d}| < (p-1)|y|L < Q \quad \text{if } L \leq Q^{\frac{1}{2}}. \tag{4.23}$$

It will suffice to let $L \leq Q^{\frac{1}{2}}$, in order to reach a contradiction. For this, we use the following principle.

4.2.4 The Siegel box principle

This is a simple estimate for short nonvanishing solutions of homogeneous integer linear systems. The question being related to the one of successive minima in lattices, it has been known since Siegel's original use — about one hundred years ago — numerous developments in various heights and different number fields, the one of Bombieri and Vaaler [17] being the most frequently used. Due to the rational scalar product introduced above, we shall be able to apply here the original version of Siegel, which is related to the pigeonhole principle application we used in (4.5). It claims the following lemma.

Lemma 4.2 *Let* $A = (a_{i,j})_{i,j=1}^{r,s}$ *be an integer matrix, with* $r < s$ *and entries bounded by* $B = \|A\|_1$*. Then there is a solution* $X = (X_1, X_2, \cdots, X_s) \in \mathbb{Z}^s \setminus \{0\}$*, with norm*

$$\|X\|_1 \le (sB)^{\frac{r}{s-r}}. \tag{4.24}$$

Under the same condition, Bombieri and Vaaler also prove

$$\|X\|_1 \le \left(\sqrt{\det(AA^T)}\right)^{\frac{1}{s-r}}.$$

4.2.5 Finding ℓ

Let $\Lambda = \mathbb{Z}^{p-1} \subset V$ and $\mathbf{B} = \{x \in \Lambda \ : \ \|x\| \le Q^{\frac{1}{2}}\}$. In view of the result of Bugeaud, Hanrot and Mignotte [5], we may assume that $p \ge 29$. If A is an $r \times (p-1)$ matrix with $\|A\|_1 < Q$, and $r \le 8$, then (4.24) implies that the system $AX = 0$ has at least one nontrivial solution in Λ, with

$$\|X\| < ((p-1)Q)^{\frac{8}{20}} < Q^{\frac{1}{2}}.$$

Here is how we use these degrees of freedom. First we impose the conditions $X \perp \mathcal{V}_0 := [U, \mathbf{b}_1(\theta)]_{\mathbb{Q}}$. The coefficients are clearly dominated by Q, so we let $X_1 \in \mathbf{B}$ be a nontrivial solution. With this we let $\mathcal{V}_1 = \mathcal{V}_0 \oplus \mathbb{Q}X_1$, and find a further solution $X_2 \in \mathbf{B}^*$, which is in addition perpendicular to X_1. We may in this way find at least six vectors $X_i \in \mathbf{B}^*$ which are mutually orthogonal and all orthogonal to $\mathcal{V}_0$. We shall use these degrees of freedom in order to find $\ell \in \mathbf{B}$, such that $\mathfrak{d} \neq 0$ in (4.22).

We note the following substitution, which leaves the sum invariant. For $\nu \in \Lambda$, we define

$$T_\nu(b_n, b_{n+1}) = (b_n + Q\nu, b_{n+1} - \nu). \tag{4.25}$$

The substitution replaces a pair of successive terms in the sequence of coefficients of the series, by leaving the sum in (4.20) unchanged. This follows immediately by considering the contribution of these terms:

$$Q^n \cdot (b_n + Qb_{n+1}) = Q^n(b_n + Q\nu + Q(b_{n+1} - \nu)).$$

In practice, $\nu \in \mathcal{E}$, the standard basis of V. Then, the modified coefficient $T_\nu(b_n)$ still verifies $|T_\nu(b_n)|_1 < \frac{3}{2Q}$, while $|T_\nu(b_{n+1})|_1 < Q$.

The choice of ν uses the following lemma.

Lemma 4.3 *Let $\mathcal{E} = \{e_j : j = 1, \cdots, p-1\}$ be the standard basis of $V = \mathbb{Q}^{p-1}$ and let $x = (x_1, x_2, \cdots, x_{p-1}) \in V$ have trace $\tau = \sum\limits_{i=1}^{p-1} x_i$. Let $t \in \mathbb{Z} \setminus \{-\tau, 0\}$ and*

$$\mathcal{F}(t) := x + t\mathcal{E} := \{x + te_i : i = 1, 2, \cdots, p-1\} \subset V.$$

Then, $\operatorname{span}\{F\} = [\mathcal{F}]_{\mathbb{Q}}$ *of $\mathcal{F}$ has dimension* $\dim(F) = p-1$.

Proof Since the $p-2$ linearly independent vectors $t(e_1 - e_i) \in F$, $i = 2, 3, \cdots, p-1$, it follows that $\dim(F) \geq p-2$. Suppose that there is a vanishing linear combination $\sum\limits_{i=1}^{p-1} \lambda_i(x + te_i) = 0$ and let $L := \sum\limits_{i=1}^{p-1} \lambda_i$ be the "trace" of $\overset{\rightarrow}{\lambda} := \sum\limits_{i \in P^*} \lambda_i e_i$. Unfolding the vanishing condition, we have

$$t\overset{\rightarrow}{\lambda} + Lx = \overset{\rightarrow}{0}.$$

If $L = 0$, then $t = 0$, which was excluded, or $\overset{\rightarrow}{\lambda} = \overset{\rightarrow}{0}$, so the linear combination was trivial to start with. If $L \neq 0$, we take traces again in the previous identity, and get $L(t + \tau) = 0$. Since $t \neq -\tau$, we obtain a contradiction, showing that the $\dim(F) = p-1$ indeed.

Let $S_0 = \mathbf{B}^* \cap \mathcal{V}_0^{\perp}$. We show that we may modify $\mathbf{b}_2$ by a substitution (4.25) in such a way, that $\mathbf{b}_2 \notin \mathcal{V}_0$ and there is at least one vector $z_0 \in S_0$ such that $z_0 \not\perp \mathbf{b}_2$. In view of Lemma 4.3 and since $\dim(S_0) > 4$, there is at least one translation of $\mathbf{b}_2$ by some base element $\nu \in Q\mathcal{E}$, which is not perpendicular to S_0. We assume thus that the condition is fulfilled, and use no new notation for the possibly modified coefficients $\mathbf{b}_2, \mathbf{b}_3$.

Let $\tau = \mathbf{Tr}(\mathbf{b}_2)$ and $t = 1$ if $\tau \leq 0$ and $t = -1$ otherwise. We define $w_j = \mathbf{b}_2 + te_j$ and $S^{(j)} = S_0 \cap w_j^{\perp}$. By Lemma 4.3, w_j span$\{V\}$, and it follows that

$$\mathbb{Q} \cdot \Big(\sum_{j \in P^*} S^{(j)} \Big) \supset \mathbb{Q}S_0.$$

There is a set

$$\emptyset \neq J = \{j \in P^* : \mathbb{Q}(w_j^{\perp} \cap S_0) \not\subset \mathbb{Q}(\mathbf{b}_2^{\perp} \cap S_0)\}.$$

A fortiori, there is a $j \in P^*$ such that $w_j^{\perp} \cap S_0 \not\subset (\mathbf{b}_2^{\perp} \cap S_0)$. For such j, we may choose $w \in S_0$ with $w \perp w_j$ but $w \not\perp \mathbf{b}_2$. We claim that $\ell = w$ satisfies our needs. Indeed, we have

$$\langle \mathbf{b}_2, \ell \rangle = \langle w_j, \ell \rangle - t\langle e_j, \ell \rangle = -t\ell_j \neq 0.$$

Consequently,

$$\mathfrak{d} \equiv Q^2(-pty\ell_j + O(Q)).$$

But $\|pty\ell_j\| \leq |y|Q^{\frac{1}{2}} < Q$, and since the choice of ℓ_j ascertains that $\ell_j \neq 0$, it follows that $\mathfrak{d} \not\equiv 0 \bmod Q^3$ and a fortiori, $\mathfrak{d} \neq 0$. But $\mathfrak{d} \equiv 0 \bmod Q^2$ implies $|\mathfrak{d}| \geq Q^2$, which contradicts the upper bound (4.23). The contradiction confirms Proposition 4.1 and completes the proof.

Declarations

Conflicts of interest The authors declare no conflicts of interest.

References

[1] Ribenboim, P., Catalan's Conjecture, *Séminaire de Philosophie et Mathématiques*, **6**, 1994, 1–11.

[2] Bugeaud, Y. and Mihăilescu, P., On the Nagell-Ljunggren equation $\frac{x^n-1}{x-1} = y^q$, *Mathematica Scandinavica*, **101**(2), 2007, 177–183.

[3] Bennett, M. A. and Levin, A., The Nagell-Ljunggren equation via Runge's method, *Monatshefte für Mathematik*, **177**, 2015, 15–31.

[4] Dupuy, B., A class number criterion for the equation $\frac{x^p-1}{x-1} = py^q$, *Acta Arithmetica*, **127**, 2007, 391–401.

[5] Bugeaud, Y., Hanrot, G. and Mignotte, M., Sur L'équation Diophantienne $\frac{x^n-1}{x-1} = y^q$ III, *Proceedings of the London Mathematical Society*, **84**(1), 2002, 59–78.

[6] Ljunggren, W., Noen Setninger om ubestemte likninger av formen $\frac{x^n-1}{x-1} = y^q$, *Norsk Matematisk Tidsskrift*, **25**(1), 1943, 17–20.

[7] Nagell, T., Note sur l'équation indéterminée $\frac{x^n-1}{x-1} = y^q$, *Norsk Matematisk Tidsskrift*, **2**, 1920, 75–78.

[8] Nagell, T., Des équations indétermines $x^2 + x + 1 = y^n$ et $x^2 + x + 1 = 3y^n$, *Norsk Matematisk Forenings Skrifter I*, **2**, 1921, 14.

[9] Cassels, J. W. S., On the equation $a^x - b^y = 1$, II, *Mathematical Proceedings of the Cambridge Philosophical Society*, **56**(2), 1960, 97–103.

[10] Bennett, M. A., Rational approximation to algebraic numbers of small height: The Diophantine equation $|ax^n - by^n| = 1$, *Journal für die Reine und Angewandte Mathematik*, **535**, 2001, 1–49.

[11] Le, M. H., A note on the Diophantine equation $\frac{x^m-1}{x-1} = y^n + 1$, *Mathematical Proceedings of the Cambridge Philosophical Society*, **116**(3), 1994, 385–389.

[12] Shorey, T. N., On the equation $z^q = \frac{x^n-1}{x-1}$, *Indagationes Mathematicae* (*Proceedings*), **89**(3), 1986, 345–351.

[13] Shorey, T. N. and Tijdeman, R., Exponential Diophantine Equations, Cambridge University Press, Cambridge, 1986.

[14] Mihăilescu, P., Class number conditions for the diagonal case of the equation of Nagell and Ljunggren, Diophantine Approximation, 2008, Springer-Verlag, Vienna, 245–273.

[15] Mihăilescu, P., New bounds and conditions for the equation of Nagell-Ljunggren, *Journal of Number Theory*, **124**(2), 2007, 380–395.

[16] Mihăilescu, P., On the class groups of cyclotomic extensions in presence of a solution to Catalan's equation, *Journal of Number Theory*, **118**(1), 2006, 123–144.

[17] Bombieri, E. and Vaaler, J., On Siegel's lemma, *Inventiones Mathematicae*, **73**, 1983, 11–32.