TORSION POINTS ON CURVES OF THE FORM $y^n = x^d + 1$

VISHAL ARUL

Abstract. In this paper we study torsion points on curves of the form $y^n = x^d + 1$ where $n$ and $d$ are coprime. When $n + d \geq 8$, we show that the only torsion points on this curve are: (i) those whose $x$-coordinate is zero, (ii) those whose $y$-coordinate is zero, (iii) the point at infinity. When $n + d = 7$, there are more torsion points and we classify them all.

1. Introduction

Fix coprime integers $n, d \geq 1$, and assume neither is a power of 2. Let $C$ be the smooth projective model of the curve given by the equation

$$y^n = x^d + 1$$

in $\mathbb{A}^2_C$. Then $C$ has a unique point at infinity, denoted by $\infty$. The genus of $C$ is

$$g = \frac{1}{2}(n - 1)(d - 1).$$

Let $J$ be the Jacobian of $C$. Then $C$ naturally embeds into $J$ via the map $P \mapsto P - \infty$; that is, the point $P$ of $C$ goes to the divisor $P - \infty$.

A point $P$ of $C$ is called a torsion point if there exists an integer $k \geq 1$ such that $kP \sim k\infty$.

We seek to classify the torsion points on $C$.

Let $\zeta_n, \zeta_{2d} \in K$ be primitive $n$th and $2d$th roots of unity, respectively. For odd $0 \leq i \leq 2d - 1$ and any $0 \leq j \leq n - 1$, we have

$$\text{div}(x - \zeta_{2d}^i) = n(\zeta_{2d}^i, 0) - n\infty,$$

$$\text{div}(y - \zeta_n^j) = d(0, \zeta_n^j) - d\infty,$$

from which it follows that the points $(\zeta_{2d}^i, 0)$ and $(0, \zeta_n^j)$ are all torsion points of $C$. Of course, there is also the point $\infty$ of $C$, which also counts as a torsion point. We seek to show that these are the only torsion points on $C$. Indeed, this will be our main result as Theorem 4.20.

We restate it here as Theorem 1.1.

Theorem 1.1. Suppose $n, d$ are coprime integers with $n, d \geq 2$. The point at infinity of $C_{n,d}$, and points of $C_{n,d}$ whose $x$- or $y$- coordinate is zero are all torsion points. These are the only torsion points except in the following cases.

(1) $(n, d) \in \{(2, 3), (3, 2)\}$. Then $C_{n,d}$ is an elliptic curve, so it has infinitely many torsion points.

This research was supported in part by Simons Foundation grant #550033.
(2) \((n, d) \in \{(2, 5), (5, 2)\}\). The only other torsion points on \(C_{2,5}\) are \(\{(\zeta^i_2 \sqrt{4}, \pm \sqrt{5}) : 0 \leq i \leq 4\}\). The curves \(C_{2,5}\) and \(C_{5,2}\) are isomorphic via \((x, y) \in C_{2,5} \mapsto (\zeta_4 y, -x) \in C_{5,2}\), so torsion points on \(C_{5,2}\) are similar.

(3) \((n, d) \in \{(3, 4), (4, 3)\}\). The only other torsion points on \(C_{4,3}\) are \(\{(2\zeta_3^i, \pm \sqrt{3}) : 0 \leq i \leq 2\}\). As before, torsion points on the isomorphic curve \(C_{3,4}\) are similar.

First we reduce to the case when \(n, d\) are both primes. If \(p, q\) are primes satisfying \(p|n\) and \(q|d\) and \(C'\) is the projective normalization of the curve cut out by \(y^p = x^n + 1\) in the affine plane, then there is a dominant map \(C \to C'\) acting on points as \((x, y) \mapsto (x^{n/p}, x^{d/q})\) which must send torsion points to torsion points (as it also induces a map on their Jacobians). So if we can show that the torsion points on \(C'\) are precisely either (i) those whose \(x\)- or \(y\)-coordinates is zero or (ii) \(\infty\), it follows that the same must be true for \(C\).

So instead we now consider the curve \(C\) which is the projective normalization of the affine plane \(K\)-curve \(y^p = x^n + 1\), for distinct odd primes \(p, q \geq 3\).

Similar results are proven for the Fermat curve \(F_n\) given by the equation \(X^m + Y^m + Z^m = 0\) in [Col86]. A cusp is a point of \(F_n(K)\) such that one of its coordinates is zero. In this paper, Coleman shows that whenever \(P\) and \(Q\) are points of \(F_n(K)\) such that \(P - Q\) is zero and \(P\) is a cusp, then \(Q\) is also necessarily a cusp. Since our curve \(y^n = x^d + 1\) is a quotient of the Fermat curve \(F_{nd}\), we obtain a slightly stronger version of this result when \(m\) is of the form \(nd\).

In [Jęd14] and [Jęd16], Jędrzejak considers a slightly more general variant given by \(y^a = x^p + a\). Jędrzejak studies the rational torsion of the Jacobian. Letting the Jacobian of this curve be \(J_{q,p,a}\), Jędrzejak shows that the group \(J_{q,p,a}(\mathbb{Q})_{\text{tors}} \cong (\mathbb{Z}/2\mathbb{Z})^{e_2} \times (\mathbb{Z}/p\mathbb{Z})^{e_p} \times (\mathbb{Z}/q\mathbb{Z})^{e_q}\) where \(e_2, e_p, e_q \in \{0,1\}\). Jędrzejak also shows that when \(a\) is odd, that \(e_2 = 0\). It follows easily that \(J_{q,p,1}(\mathbb{Q})_{\text{tors}} \cong (\mathbb{Z}/p\mathbb{Z}) \times (\mathbb{Z}/q\mathbb{Z})\), generated by the points \((-1,0)\) and \((0,1)\). Moreover in the case \(a = 1\) we work out explicitly the torsion fields \(\mathbb{Q}(J_{q,p,1}[p], \mu_{pq})\) and \(\mathbb{Q}(J_{q,p,1}[q], \mu_{pq})\) in Theorem 3.6. The key ingredient is an understanding of the \(p\)-adic and \(q\)-adic valuation of certain Jacobi sums; this analysis is performed in [Aru19].

2. The Structure of \(T_{\ell}J\) as a \(\mathbb{Z}\)-Representation

Let \(J\) be the Jacobian of \(C\). For any prime \(\ell\), let \(T_{\ell} J\) be the \(\ell\)-adic Tate module of \(J\).

Now define \(Z\) to be the subgroup of \(\text{Aut}(C)\) generated by the automorphism sending \((x, y) \mapsto (\zeta y, \zeta^m y)\). Note that \(Z\) is naturally isomorphic to \(\mu_{pq}\). We will seek to understand \(T_{\ell} J\) as a representation of \(Z\).

For every positive integer \(m\), define \(H_{\infty,m}\) to be the following Galois group.

\[
H_{\infty,m} := \text{Gal}(\mathbb{Q}(\mu_{pq}, J[m^{\infty}])/\mathbb{Q}(\mu_{pq})).
\]

**Proposition 2.1.** Suppose \(K_{\ell}\) is any extension of \(\mathbb{Q}_{\ell}\) containing a primitive \(pq\)-th root of unity. Let \(O_{\ell} = O_{K_{\ell}}\). Let \(J_{\ell}\) be the group of characters (group homomorphisms) \(\chi: Z \to O_{\ell}^{\times}\) and \(T_{\chi} \subseteq T_{\ell} J \otimes_{\mathbb{Z}_{\ell}} O_{\ell}\) be the eigenspace corresponding to \(\chi\).

(1) We have that \(T_{\chi} \cong \begin{cases} O_{\ell} & \text{if } \chi \text{ is injective}. \\ 0 & \text{otherwise}. \end{cases}\)
(2) We have a decomposition

\[ T_{\ell}J \otimes \mathcal{O}_\ell \simeq \bigoplus_{\chi \in J_{\ell}} T_{\chi}. \]

that respects the \( H_{\infty, \ell} \) action; in particular, we get characters for \( H_{\infty, \ell} \) indexed by \( J_{\ell} \).

For injective \( \chi \), define

\[ \xi_{\chi} : H_{\infty, \ell} \to \text{Aut} T_{\chi} \simeq \mathcal{O}_\ell^\times \]

to be the action of \( H_{\infty, \ell} \) on \( T_{\chi} \).

(3) The Weil pairing extends to a nondegenerate symplectic pairing on \( T_{\ell}J \otimes \mathcal{O}_\ell \) taking values in \( T_{\ell} \mu_{\ell} \otimes \mathcal{O}_\ell \). Furthermore, for the Weil pairing we have

\[ \langle T_{\chi}, T_{\psi} \rangle = 0 \]

whenever \( \psi \neq \chi^{-1} \).

\[ \square \]

**Proof.**

(1) (a) (Case 1: \( \chi \) is not injective)

Either \( \chi^p = 1 \) or \( \chi^q = 1 \). The two cases are similar so we handle the former. Then \( \mathbb{Z}^p \) must act trivially on \( T_{\chi} \subseteq T_{\ell}J \otimes \mathcal{O}_\ell \), so in fact we must have the containment

\[ T_{\chi} \subseteq (T_{\ell}J \otimes \mathcal{O}_\ell)^{Z^p} = (T_{\ell}J)^{Z^p} \otimes \mathcal{O}_\ell. \]

Now note that under the quotient map \( C \to C/Z^p \) we get an induced map \( J \to \text{Jac}(C/Z^p) \) which induces \( T_{\ell}J \to (T_{\ell} \text{Jac}(C/Z^p)) \) and allows us to identify \( T_{\ell} \text{Jac}(C/Z^p) \) with \( (T_{\ell}J)^{Z^p} \). However \( C/Z^p \) is isomorphic to \( \mathbb{P}^1 \), so

\[ (T_{\ell}J)^{Z^p} \simeq T_{\ell} \text{Jac}(C/Z^p) \simeq T_{\ell} \mathbb{P}^1 = 0. \]

(b) (Case 2: \( \chi \) is injective)

Note that \( \text{Gal}(\mathbb{Q}(\mu_{pq})/\mathbb{Q}) \) acts on \( \mu_{pq} \), which is naturally isomorphic to \( \mathbb{Z} \). Then this Galois group must also act on the group of characters \( J_{\ell} \), and it acts on the injective characters transitively. Therefore \( \dim T_{\chi} \) is independent on \( \chi \) for the injective \( \chi \).

Then as

\[ \dim_{\mathcal{O}_\ell} T_{\ell}J \otimes \mathcal{O}_\ell = \dim_{\mathbb{Z}_\ell} T_{\ell}J \]

\[ = 2g = (p - 1)(q - 1) \]

\[ = \#\{ \text{injective characters } \chi \}, \]

this shows that \( \dim T_{\chi} = 1 \) for injective \( \chi \).

(2) The previous argument shows that this decomposition exists. It respects the \( H_{\infty, \ell} \) action since the actions of \( H_{\infty, \ell} \) and \( \mathbb{Z} \) on \( C \) both commute with each other.

(3) The Weil pairing on \( T_{\ell}J \) is \( \mathbb{Z} \)-invariant (since \( \mathbb{Z} \) consists of automorphisms of the curve). It follows then that the dual of \( \langle T_{\chi}, T_{\psi} \rangle = 0 \) whenever \( \chi \psi \neq 1 \).

\[ \square \]

**Definition 2.2.** Since \( H_{\infty, \ell} \) acts on \( \mu_{l^\infty} \), it induces a map \( H_{\infty, \ell} \to \text{Aut}(\mu_{l^\infty}) = \mathbb{Z}_l^\times \). Let \( \lambda \) be the map

\[ \lambda : H_{\infty, \ell} \to \mathbb{Z}_l^\times \hookrightarrow \mathcal{O}_l^\times. \]

Since the Weil pairing is nondegenerate, we know that \( \mathbb{Q}(J[l^\infty]) \) contains \( \mathbb{Q}(\mu_{l^\infty}) \). Therefore,
Lemma 2.3. The image of $\lambda$ is the following.

$$\lambda(H_{\infty,\ell}) = \begin{cases} 
\mathbb{Z}_\ell^\times & \text{for } \ell \notin \{p, q\} \\
\ker(\mathbb{Z}_\ell^\times \to F_\ell^\times) & \text{for } \ell \in \{p, q\}.
\end{cases}$$

Lemma 2.4. For every $\chi \in J_\ell$ we have

$$\xi_\chi \xi_{\chi^{-1}} = \lambda.$$  

Proof. From Galois-equivariance of the Weil pairing we have that if $v \in T_\chi$, $w \in T_{\chi^{-1}}$, and $h \in H_{\infty,\ell}$ then

$$h(\langle v, w \rangle) = \langle h(v), h(w) \rangle = \langle \xi_\chi(h)v, \xi_{\chi^{-1}}(h)w \rangle = (\xi_\chi \xi_{\chi^{-1}})(h)(v, w).$$

If $\langle v, w \rangle$ is chosen to be a primitive element of $\mu_\infty$, the above shows that $\lambda(h) = \xi_\chi \xi_{\chi^{-1}}(h)$. □

Remark 2.5. By Proposition 2.1 (2), we have an embedding

$$H_{\infty,\ell} \hookrightarrow \prod_{\chi: \bar{\mathcal{O}}_\ell^\times \to \mathcal{O}_\ell^\times} \text{Aut} T_\chi \simeq (\mathcal{O}_\ell^\times)^{2g}.$$  

In particular, $H_{\infty,\ell}$ is abelian. Taking the direct sum over all $\ell$, we see that the torsion field $\mathbb{Q}(\mu_{pq}, J_{\text{tors}})$ is abelian over $\mathbb{Q}(\mu_{pq})$.

In particular, the group $\text{Gal}(\mathbb{Q}(\mu_{pq})/\mathbb{Q})$ acts on $H_{\infty,m}$ via conjugation in a well-defined way. This inspires the following definition

Definition 2.6. Let $\sigma \in \text{Gal}(\mathbb{Q}(\mu_{pq})/\mathbb{Q})$ be complex conjugation. For $h \in H_{\infty,m}$, define $\overline{h} = \sigma h \sigma^{-1} \in H_{\infty,m}$. (That is, lift $\sigma$ arbitrarily to $\tilde{\sigma}$ and then define $\overline{h} = \tilde{\sigma} h \tilde{\sigma}^{-1}$; this is well-defined since $H_{\infty,m}$ is abelian.)

Lemma 2.7. Let $h \in H_{\infty,\ell}$ and $\chi \in J_\ell$. The image of $\overline{h}$ under $\xi_\chi$ is

$$\xi_\chi(\overline{h}) = \xi_{\chi^{-1}}(h).$$

Proof. Pick $v \in T_\chi$, lift $\sigma \in \text{Gal}(\mathbb{Q}(\mu_{pq})/\mathbb{Q})$ arbitrarily to $\tilde{\sigma} \in \text{Gal}(\mathbb{Q}(\mu_{pq}, J[\ell^\infty])/\mathbb{Q})$ and consider the action of $\tilde{\sigma}$ on $T_\chi \otimes_{\mathbb{Z}_\ell} \mathcal{O}_\ell = \oplus_{\chi \in J_\ell} T_\chi$. We know that $\tilde{\sigma}$ will send $T_{\chi^{-1}}$ isomorphically to $T_\chi$. Since $\tilde{\sigma}$ is linear and $\tilde{\sigma} v \in T_{\chi^{-1}}$, we see that

$$\overline{h} v = \tilde{\sigma} h \tilde{\sigma}^{-1} v = \tilde{\sigma} (\xi_{\chi^{-1}}(h) \tilde{\sigma}^{-1} v) = \xi_{\chi^{-1}}(h) \tilde{\sigma}(\tilde{\sigma}^{-1} v) = \xi_{\chi^{-1}}(h) v.$$  

From this we conclude that $\xi_\chi(\overline{h}) = \xi_{\chi^{-1}}(h)$. □

Proposition 2.8. Let $h \in H_{\infty,\ell}$. The element $\overline{h}$ acts by $\lambda(h)$ on all of $T_\ell J \otimes_{\mathbb{Z}_\ell} \mathcal{O}_\ell$.

Proof. Combining Lemmas 2.4 and 2.7 gives that $\xi_\chi(\overline{h}) = \xi_\chi(h) \xi_{\chi^{-1}}(h) = \xi_{\chi^{-1}}(h)$ since $\lambda(h) = \lambda(h)$. Hence $\overline{h}$ acts by multiplication by $\lambda(h)$ on every $T_\chi$, and hence on $T_\ell J \otimes_{\mathbb{Z}_\ell} \mathcal{O}_\ell$. □

Lemma 2.9. Suppose $\ell \notin \{p, q\}$. Then the field $\mathbb{Q}(J[\ell])$ contains $\mathbb{Q}(\mu_{pq})$.

Proof. Define the homomorphism $\nu$ as the following.

$$\nu: \text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to \text{Gal}(\mathbb{Q}(\mu_{pq})/\mathbb{Q}) \simeq (\mathbb{Z}/pq\mathbb{Z})^\times.$$  

It suffices to show for every $\tau \in \text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ fixing $J[\ell]$ that we have $\nu(\tau) = 1$.

Let $D$ be any nonzero element of $J[\ell]$. Then for every $z \in \mathbb{Z}$, we have $\tau z D = z D$. Therefore,

$$D = z^{-1} z D = z^{-1} z D = z^{\nu(\tau)-1} z D = z^{\nu(\tau)-1} D.$$  

4
Therefore, \( z^{\nu(\tau) - 1} \) must fix every element of \( J[\ell] \). If it were the case that \( \nu(\tau) \neq 1 \), then either \( Z^p \) or \( Z^q \) must act trivially on \( J[\ell] \), which forces \( \ell \) to be either \( p \) or \( q \).

**Definition 2.10.** For \( \ell \in \{p, q\} \), let \( m_\ell \) be the maximal ideal of \( O_\ell \).

**Definition 2.11.** Let \( \zeta_p \) be the automorphism given by \( (x, y) \mapsto (x, \zeta_p y) \).

Let \( \zeta_q \) be the automorphism given by \( (x, y) \mapsto (\zeta_q x, y) \).

**Lemma 2.12.** Suppose that \( \alpha \) is an endomorphism of \( T_\ell J \) that commutes with \( Z \) and that \( k \) is a nonnegative integer.

1. Note that \( \alpha \) induces an endomorphism of \( J[\ell^k] \). Then \( \alpha \) acts as the identity on \( J[\ell^k] \) if and only if \( \alpha \) acts on each \( T_\chi \) by multiplication by some element of \( 1 + \ell^k O_\ell \) (i.e., if and only if \( \xi_\chi(\alpha) \in 1 + \ell^k O_\ell \) for every \( \chi \)).

2. Suppose \( \ell \in \{p, q\} \). Since \( \alpha \) commutes with \( \zeta_\ell \), it induces an endomorphism of \( J[(1 - \zeta_\ell)^k] \). Then \( \alpha \) acts as the identity on \( J[(1 - \zeta_\ell)^k] \) if and only if \( \alpha \) acts on each \( T_\chi \) by multiplication by some element of \( 1 + m_\ell^k \) (i.e., if and only if \( \xi_\chi(\alpha) \in 1 + m_\ell^k \) for every \( \chi \)).

**Proof.** By extending scalars, \( \alpha \) is also an endomorphism of \( T_\ell J \otimes_{\mathbb{Z}_\ell} O_\ell \).

1. As

\[
J[\ell^k] = T_\ell J/\ell^k,
\]

we see that \( \alpha \) acts trivially on \( J[\ell^k] \) if and only if it acts trivially on \( T_\ell J/\ell^k \), if and only if (after extending scalars), it acts trivially on each \( T_\chi/\ell^k \).

Since \( \alpha \) commutes with \( Z \), it acts on each \( T_\chi \) by multiplication by an element of \( O_\ell \). As \( T_\chi \cong O_\ell \) (as \( O_\ell \)-modules) induces the isomorphism \( T_\chi/\ell^k \cong O_\ell/\ell^k O_\ell \), the conclusion of the lemma follows.

2. This proof is very similar to the previous part. Replace “\( \ell \)” with “\( 1 - \zeta_\ell \)” and \( \ell^k O_\ell \) with \( m_\ell^k \).

Using results of [Kat81], we can get an expression for \( \xi_\chi(g) \) when \( g \) is a Frobenius element, in terms of Jacobi sums. To do so, first select a prime \( r \notin \{p, q, \ell\} \) and a prime \( \tau \) of \( \mathbb{Q}(\mu_{pq}) \) lying over \( r \) whose residue field is \( F_\tau \). Since \( r \notin \{p, q\} \), all the automorphisms in \( Z \) can be reduced to automorphisms over \( F_\tau \). Moreover, if we let \( Z_\tau \) be the collection of these automorphisms defined over \( F_\tau \), there is a natural isomorphism \( Z \cong Z_\tau \) and also a natural isomorphism \( Z_\tau \cong \mu_{pq}(F_\tau) \).

The following lemma is essentially a reformulation of some of the results in the first three sections of [Kat81].

**Lemma 2.13.** Select a prime \( r \notin \{p, q, \ell\} \), a prime \( \tau \) of \( \mathbb{Q}(\mu_{pq}) \) lying over \( r \) with residue field \( F_\tau \), and a Frobenius element \( \text{Frob}_\tau \in H_{\infty, \ell} \) (note this is well-defined since \( \mathbb{Q}(\mu_{pq}, J[\ell^{\infty}]) \) is unramified over \( r \) by the criterion of Néron-Ogg-Shafarevich). Suppose that the size of \( F_\tau \) is \( R \).

Suppose \( \chi : Z \to O_\ell^{\times} \) is a character. Define \( \overline{\chi} : F_\tau^{\times} \to O_\ell^{\times} \) as the composite of the “exponentiation by \( (R - 1)/(pq) \) map” \( F_\tau^{\times} \to \mu_{pq}(F_\tau) \) and the natural isomorphisms \( \mu_{pq}(F_\tau) \cong Z_\tau \), \( Z_\tau \cong Z \), and the character \( \chi : Z \to O_\ell^{\times} \).

Then

\[
\xi_\chi(\text{Frob}_\tau) = - \sum_{\alpha \in F_\tau \setminus \{0, 1\}} \overline{\chi}^p(\alpha) \overline{\chi}^q(1 - \alpha).
\]

\[\square\]
Proof. From [Kat81] Lemma 1.1, we know that Frob\(_r\) operates on the \(\chi\)-isotypical part \(T\chi\) of the Tate module \(T_\ell J \cong H^1_{et}(C, \mathbb{Q}_\ell)\) via multiplication by

\[
-S(C/F_r, \chi, 1) := -\frac{1}{|Z|} \sum_{z \in Z} \chi(z) \# \text{Fix(Frob}_r z^{-1})
\]

In this last expression, the quantity \(\# \text{Fix(Frob}_r z^{-1})\) is the number of points of \(C(F_r)\) fixed by \(\text{Frob}_r z^{-1}\).

Now choose some \(z \in Z\). Write \(z = z_p z_q\) where \(z_p\) and \(z_q\) have order \(p\) and \(q\), respectively. Since \(Z \cong Z_r\), we identify \(z, z_p, z_q\) with automorphisms of \(C\) defined over \(F_r\). Let \(\zeta_p^a\) and \(\zeta_q^a\) denote elements of \(F_r^\times\) such that (i) \(\zeta_p^a\) is the scalar by which \(z_p\) acts on the \(y\)-coordinate by multiplication, (ii) \(\zeta_q^a\) is the scalar by which \(z_q\) acts on the \(x\)-coordinate by multiplication.

Recall that \(R\) is the size of \(F_r\). Note that \((x, y)\) is fixed by \(\text{Frob}_r z^{-1}\) if and only if we have the following:

\[
\begin{align*}
x^R &= \zeta_q^a x \\
y^R &= \zeta_p^a y.
\end{align*}
\]

From these equations we see that \(x^q\) and \(y^p\) are both fixed by \(\text{Frob}_r\), so \(x^q, y^p \in F_r\). We also have \(y^p = x^q + 1\). Setting \(\alpha = -x^q\) then, we have that \(x^q = -\alpha, y^p = 1 - \alpha\) and that \(\alpha \in F_r\).

Suppose \(x, y \neq 0\). Then from \(\alpha\) we can recover \(\zeta_p^a\) and \(\zeta_q^a\) by \(\zeta_q^a = x^{R-1} = (-\alpha)^{(R-1)/q}\) and \(\zeta_p^a = y^{R-1} = (1 - \alpha)^{(R-1)/p}\). In particular, from our definition of \(\widetilde{\chi}\) we know that

\[
\begin{align*}
\chi(z_q) &= \widetilde{\chi}^p(-\alpha) \\
\chi(z_p) &= \widetilde{\chi}^q(1 - \alpha)
\end{align*}
\]

When \(R\) is odd, we know that \((R - 1)/q\) will be even, so \(\widetilde{\chi}^p(-\alpha) = \widetilde{\chi}^p(-1)\widetilde{\chi}^p(\alpha) = \widetilde{\chi}^p(\alpha)\) as \((-1)^{(R-1)/q} = 1\). When \(R\) is even, we know that \(\alpha = -\alpha\) so in any case we can remove the minus sign to get

\[
\begin{align*}
\chi(z_q) &= \widetilde{\chi}^p(\alpha) \\
\chi(z_p) &= \widetilde{\chi}^q(1 - \alpha).
\end{align*}
\]

Multiplying these two equations gives

\[
\chi(z) = \widetilde{\chi}^p(\alpha) \widetilde{\chi}^q(1 - \alpha).
\]

Since we made the assumption that \(x, y \neq 0\), let \(C(F_q)^*\) be the subset of \(C(F_q)\) where neither the \(x\)- nor the \(y\)-coordinate is zero.
Going back to our sum, we then have
\[
\sum_{z \in Z} \chi(z) \# \text{Fix}(\text{Frob}_\chi z^{-1}) = \sum_{z \in Z} \chi(z) \sum_{(x,y) \in (\mathbb{F}_r)^*} \chi(z) + \sum_{z \in Z} \chi(z) \sum_{(x,y) \in (\mathbb{F}_r)^*} \chi(z)
\]
\[
= \sum_{(x,y) \in (\mathbb{F}_r)^*} \sum_{\text{Frob}_r(x,y) = z(x,y)} \chi(z) + \sum_{z \in Z} \chi(z) \sum_{(x,y) \in (\mathbb{F}_r)^*} \chi(z) \sum_{(x,y) \in (\mathbb{F}_r)^*} \chi(z)
\]
\[
= \sum_{\alpha \in \mathbb{F}_l \setminus \{0,1\}} (x,y,z) \in (\mathbb{F}_r)^2 \times Z \sum_{\text{Frob}_r(x,y) = z(x,y)} \chi(z) \sum_{x^q = -\alpha, y^p = 1 - \alpha} \chi^p(\alpha) \chi^q(1 - \alpha)
\]
\[
+ \sum_{z \in Z} \chi(z) \sum_{y \in \mathbb{F}_r} \sum_{y^p = 1} \chi(z) = \sum_{z \in Z} \chi(z) = 0 = 0.
\]

Therefore, we have
\[
\sum_{z \in Z} \chi(z) \# \text{Fix}(\text{Frob}_\chi z^{-1}) = \sum_{\alpha \in \mathbb{F}_l \setminus \{0,1\}} (x,y,z) \in (\mathbb{F}_r)^2 \times Z \sum_{\text{Frob}_r(x,y) = z(x,y)} \chi(z) \sum_{x^q = -\alpha, y^p = 1 - \alpha} \chi^p(\alpha) \chi^q(1 - \alpha).
\]

In this final inner sum, we know that \( z \) is determined by \( \alpha \): the equations \( x^q = -\alpha \) and \( x^p = 1 - \alpha \) force \( x^R = (-\alpha)^{(R-1)/q} \) and \( y^R = (1 - \alpha)^{(R-1)/p} \), so that means that \( z \) must scale the \( x \)-coordinate by \( (-\alpha)^{(R-1)/q} \) and the \( y \)-coordinate by \( (1 - \alpha)^{(R-1)/p} \). So we may rewrite this as
\[
\sum_{z \in Z} \chi(z) \# \text{Fix}(\text{Frob}_\chi z^{-1}) = \sum_{\alpha \in \mathbb{F}_l \setminus \{0,1\}} (x,y) \in (\mathbb{F}_r)^2 \sum_{x^q = -\alpha, y^p = 1 - \alpha} \chi^p(\alpha) \chi^q(1 - \alpha)
\]

Since \( \alpha \notin \{0,1\} \), there are exactly \( pq = |Z| \) such pairs \( (x,y) \) satisfying \( x^q = -\alpha, y^p = 1 - \alpha \). So this sum simplifies to
\[
\sum_{z \in Z} \chi(z) \# \text{Fix}(\text{Frob}_\chi z^{-1}) = |Z| \sum_{\alpha \in \mathbb{F}_l \setminus \{0,1\}} \chi^p(\alpha) \chi^q(1 - \alpha).
\]

Dividing both sides by \(-|Z|\) and recalling that \( \xi_\chi(\text{Frob}_\chi) = -\frac{1}{|Z|} \sum_{z \in Z} \chi(z) \# \text{Fix}(\text{Frob}_\chi z^{-1}) \), we are done. \( \square \)
Definition 2.14. For two characters $\psi, \psi' : F^\times_\ell \to \mathcal{O}^\times_\ell$, define the Jacobi sum $J(\psi, \psi')$ to be

$$J(\psi, \psi') := \sum_{\alpha \in F^*} \psi(\alpha)\psi'(1 - \alpha).$$

That is, $\xi_\chi(\text{Frob}_\ell) = -J(\bar{\chi}_p, \bar{\chi}_q)$. Applying Lemma 2.12 to our situation with $\alpha = \text{Frob}_\ell$, we obtain the following corollary.

Corollary 2.15. Suppose that $\ell$ is a prime and $k$ is a nonnegative integer.

1. Then $\text{Frob}_\ell$ acts as the identity on $J[\ell^k]$ if and only if for every character $\bar{\chi} : F^\times_\ell \to \mathcal{O}^\times_\ell$, we have $$1 + J(\bar{\chi}_p, \bar{\chi}_q) \in \ell^k \mathcal{O}_\ell.$$  

2. Suppose $\ell \in \{p, q\}$. Then $\text{Frob}_\ell$ acts as the identity on $J[(1 - \zeta_\ell)^k]$ if and only if for every character $\bar{\chi} : F^\times_\ell \to \mathcal{O}^\times_\ell$, we have $$1 + J(\bar{\chi}_p, \bar{\chi}_q) \in \mathfrak{m}_\ell^k.$$  

Lemma 2.16. Fix a prime $\ell$ and a nonnegative integer $k$. Then

1. Suppose $\ell \not\equiv \{p, q\}$ and $D$ is a divisor of exact order $\ell^{k+1}$; that is, $D \in J[\ell^{k+1}] \setminus J[\ell^k]$. Then $Q(D, \mu_{pq}) = Q(J[\ell^{k+1}], \mu_{pq})$.

2. Suppose $\ell \in \{p, q\}$ and $D$ is a divisor of exact order $(1 - \zeta_\ell)^{k+1}$; that is, $D \in J[(1 - \zeta_\ell)^k+1] \setminus J[(1 - \zeta_\ell)^k]$. Then $Q(D, \mu_{pq}) = Q(J[(1 - \zeta_\ell)^{k+1}], \mu_{pq})$.

Proof. The inclusions $Q(D, \mu_{pq}) \subseteq Q(J[\ell^{k+1}], \mu_{pq})$ and $Q(D, \mu_{pq}) \subseteq Q(J[(1 - \zeta_\ell)^{k+1}], \mu_{pq})$ are immediate.

1. By Galois theory, it suffices to show that any $h \in \text{Gal}(Q(J[\ell^{k+1}], \mu_{pq})/Q(\mu_{pq}))$ fixing $D$ must be the identity. Suppose $h$ is such an element. By the Chebotarev Density Theorem, we can assume $h = \text{Frob}_\ell$. By Corollary 2.15 (1), we need to show that $1 + J(\bar{\chi}_p, \bar{\chi}_q) \in \ell^{k+1} \mathcal{O}_\ell$ for every $\bar{\chi}$. We see that $J(\bar{\chi}_p, \bar{\chi}_q)$ is actually an element of $Z[\zeta_{pq}]$, so we just need to show that $1 + J(\bar{\chi}_p, \bar{\chi}_q) \in \ell^{k+1} \mathcal{O}_\ell$ for some $\bar{\chi}$ (since the others are just Galois conjugates of our favorite one).

Consider the map

$$T_\ell J \hookrightarrow T_\ell T_\chi \otimes_{\mathbf{Z}_\ell} \mathcal{O}_\ell \cong \oplus_\chi T_\chi.$$

Taking a quotient by $\ell^{k+1}$, we get a map

$$J[\ell^{k+1}] \hookrightarrow \oplus_\chi T_\chi/\ell^{k+1}T_\chi.$$

Note that the image of $J[\ell^k]$ will be $\oplus_\chi (T_\chi/\ell^{k+1}T_\chi)$. Since $D \in J[\ell^{k+1}] \setminus J[\ell^k]$, there will be some $\chi$ such that the image of $D$ in the projection to $T_\chi/\ell^{k+1}T_\chi$ will land in

$$(T_\chi/\ell^{k+1}T_\chi) \setminus (\ell T_\chi/\ell^{k+1}T_\chi).$$

For convenience, let $D_\chi$ be the image of $D$ in $T_\chi/\ell^{k+1}T_\chi$. Since $h$ fixes $D$, we know that $\xi_\chi(h)$ fixes $D_\chi$. We also have

$$D_\chi \in (T_\chi/\ell^{k+1}T_\chi) \setminus (\ell T_\chi/\ell^{k+1}T_\chi) \simeq (\mathcal{O}_\ell/\ell^{k+1}\mathcal{O}_\ell) \setminus (\ell\mathcal{O}_\ell/\ell^{k+1}\mathcal{O}_\ell).$$

Let $R_\ell$ be the local ring $\mathcal{O}_\ell/\ell^{k+1}\mathcal{O}_\ell$ with maximal ideal $\mathfrak{m}_{R_\ell} = \ell R_\ell$. Then $D_\chi$ is a unit of $R_\ell$, so its annihilator must be zero. Hence the image of $\xi_\chi(h) - 1$ is zero in $R_\ell$. In other words, we know that $\xi_\chi(h) - 1 \not\in \ell^{k+1}\mathcal{O}_\ell$. Hence by Lemma 2.13 we have $1 + J(\bar{\chi}_p, \bar{\chi}_q) = 1 - \xi_\chi(h) \in \ell^{k+1}\mathcal{O}_\ell$, which completes the proof.
(2) The proof is very similar to the previous part. Replace “ℓ” with $1 - \zeta_\ell$. □

3. Computation of Some Torsion Fields

In this section, we use results of [Aru19] to compute some torsion fields.

**Definition 3.1.** Let $\zeta_p$ be the automorphism given by $(x, y) \mapsto (x, \zeta_p y)$.

Let $\zeta_q$ be the automorphism given by $(x, y) \mapsto (\zeta_q x, y)$.

For nonnegative $i, j$ define

$$L_{i,j} := \mathbb{Q}(J[(1 - \zeta_p)^i(1 - \zeta_q)^j])$$

Note that $J[(1 - \zeta_p)^{p-1}(1 - \zeta_q)^{q-1}] = J[pq]$, so $L_{p-1,q-1} = \mathbb{Q}(J[pq])$.

**Lemma 3.2.** We have the following facts about the fields $L_{1,1}$, $L_{1,2}$, $L_{2,1}$:

1. The field $L_{1,1}$ is $\mathbb{Q}(\mu_{pq})$.
2. The field extension $L_{2,1}/L_{1,1}$ is generated by the $p$th roots of the numbers $1 - \zeta_p^i$.
3. The field extension $L_{1,2}/L_{1,1}$ is generated by the $q$th roots of the numbers $1 - \zeta_q^j$.
4. The field extensions $L_{2,1}/L_{1,1}$ and $L_{1,2}/L_{1,1}$ are nontrivial.

**Proof.**

1. The field $L_{1,1}$ is generated by the points whose $x$-coordinates are zero and the points whose $y$-coordinates are zero, so it is exactly $\mathbb{Q}(\mu_{pq})$.
2. Let $L$ be a number field containing $L_{1,1} = \mathbb{Q}(\mu_{pq})$ and $A = L[T]/(T^q + 1)$. Then we know that we have the “$x - T$” map

$$J(L)/(1 - \zeta_p)J(L) \hookrightarrow \ker(A^\times/(A^\times)^p \xrightarrow{N} L^\times/(L^\times)^p)$$

which is essentially the Galois cohomology coboundary map

$$J(L)/(1 - \zeta_p)J(L) \hookrightarrow H^1(L, J[1 - \zeta_p])$$

arising from the short exact sequence

$$0 \rightarrow J[1 - \zeta_p] \rightarrow J \xrightarrow{1 - \zeta_p} J \rightarrow 0.$$

The key point is that $H^1(L, J[1 - \zeta_p])$ is isomorphic to $\ker(A^\times/(A^\times)^p \xrightarrow{N} L^\times/(L^\times)^p)$ and that the image of a point $(a, b) \in C(L) \subseteq J(L)$ is $a - T$.

See [Poo06] Section 6.3 for more details of the “$x - T$” map; Poonen treats the hyperelliptic situation for the “multiplication by 2” isogeny, but this generalizes in a straightforward fashion in the superelliptic case and the “multiplication by $(1 - \zeta_p)$” isogeny.

In order for $L$ to contain the $(1 - \zeta_p)^2(1 - \zeta_q)$-torsion of $J$, it only needs to contain the $(1 - \zeta_p)^2$-torsion (since it already has the $(1 - \zeta_p)(1 - \zeta_q)$-torsion as it contains $\mu_{pq}$). Therefore it is sufficient for it to contain all divisors $D$ such that $(1 - \zeta_p)D \sim (-\zeta_q^i, 0) - \infty$ for $0 \leq i \leq q - 1$. That is, the points $(-\zeta_q^i, 0) - \infty$ need to be mapped to the identity in the above “$x - T$” map.

So we need $-\zeta_q^i - T$ to lie in $(A^\times)^p$. The Chinese remainder theorem gives that this is equivalent to $\zeta_q^i - \zeta_q^k$ being a $p$th-power for all $i, j$. Since $L$ already contains $\mu_{pq}$, this is equivalent to $1 - \zeta_q^k$ being a $p$th power for all $k$. 

9
(3) Similar to the proof of part 2.

(4) We see that $L_{2, i}/L_{1, i}$ is nontrivial; the latter is similar. Consider the ramification of $L_{2, i}$ and $L_{1, i}$ above the prime $q$. Note that as $L_{2, i}$ contains $(1 - \zeta_q)^{1/p}$, we see that $e_q(L_{2, i}/\mathbb{Q}) \geq p(q - 1)$. However the degree of the extension $L_{1, i}/\mathbb{Q}$ is only $(p - 1)(q - 1)$, so $L_{2, i}$ has to strictly contain $L_{1, i}$.

\[ \Box \]

**Lemma 3.3.** We have the following facts about the fields $L_{p - 1, i}, L_{1, q - 1}, L_{p - 1, q - 1}$.

1. The extension $L_{p - 1, i}/L_{1, i}$ is a $p$-Kummer extension; it is generated by $p$-th roots of elements of $L_{1, i} = \mathbb{Q}(\mu_{p q})$.

2. The extension $L_{1, q - 1}/L_{1, i}$ is a $q$-Kummer extension; it is generated by $q$-th roots of elements of $L_{1, i} = \mathbb{Q}(\mu_{p q})$.

**Proof.** Both parts are similar so we prove the first. From Remark 2.3 we know that the extension is abelian. Since $L_{1, i}$ already contains the $p$th roots of unity, it suffices to check that the exponent of $\text{Gal}(L_{p - 1, i}/L_{1, i})$ divides $p$. To do so, we need to check that for every $h \in \text{Gal}(L_{p - 1, i}/L_{1, i})$ that $h^p = 1$. From Lemma 2.12 we know that $\xi_\chi(h) \in 1 + m_p$ for all $\chi$. In particular,

\[ \xi_\chi(h^p) = \xi_\chi(h)^p \in (1 + m_p)^p \in 1 + p\mathcal{O}_p, \]

so again by Lemma 2.12 we see that $h^p$ acts trivially on $J[p]$; hence, $h^p = 1$ as desired. \[ \Box \]

Using the main result in [Aru19], we can actually determine $L_{p - 1, i}$ and $L_{1, q - 1}$. Theorem 1.5 in [Aru19] states the following.

**Theorem 3.4.** Fix a prime $\tau$ of $\mathbb{Q}(\mu_{p q})$ lying over a prime $r$ of $\mathbb{Q}$ such that $r \notin \{p, q\}$. Let $\zeta_p$ and $\zeta_q$ denote primitive $p$th and $q$th roots of unity in $\mathbf{F}_r$. Take any Jacobi sum $J = J(\chi^p, \tilde{\chi}^q)$ and an integer $k$ in the range $2 \leq k \leq p - 1$. Then

1. $J + 1$ always lies in $m_p$.
2. $J + 1$ lies in $m_p$ if and only if for each $i$ in the range $0 \leq i \leq k - 2$ and $j$ in the range $1 \leq j \leq q - 1$, we have

\[ \prod_{r=0}^{p-1} (1 - \zeta_q^{j r})^{(i)} \in (\mathbf{F}_\tau^\times)^p. \]

Combining Theorem 3.4 and Corollary 2.15 we conclude that

**Corollary 3.5.** Fix a prime $\tau$ of $\mathbb{Q}(\mu_{p q})$ lying over a prime $r$ of $\mathbb{Q}$ such that $r \notin \{p, q\}$. Let $k$ be an integer in the range $2 \leq k \leq p - 1$. Then

1. $\text{Frob}_\tau$ always acts as the identity on $J[(1 - \zeta_p)]$.
2. $\text{Frob}_\tau$ acts as the identity on $J[(1 - \zeta_p)^k]$ if and only if for each $i$ in the range $0 \leq i \leq k - 2$ and $j$ in the range $1 \leq j \leq q - 1$, we have

\[ \prod_{r=0}^{p-1} (1 - \zeta_q^{j r})^{(i)} \in (\mathbf{F}_\tau^\times)^p. \]

Now applying the Chebotarev density theorem, Corollary 3.3 allows us to understand the field extension $L_{k, i}/\mathbb{Q}(\mu_{p q})$ for every $k$ in the range $1 \leq k \leq p - 1$. (Similarly, we also understand $L_{1, k}$ for $k$ in the range $1 \leq k \leq q - 1$.) We get the following result.

**Theorem 3.6.** Let $k$ be an integer in the range $2 \leq k \leq p - 1$. Then
\( (1) \quad L_{1,1} = \mathbb{Q}(\mu_{pq}). \)

\( (2) \quad L_{k,1} = L_{k-1,1} \left( \sqrt[p]{\prod_{r=0}^{p-1} \left( 1 - \zeta_r^j \zeta_p^q \right)^{(r)} : 1 \leq j \leq q-1 \right). \)

Let us investigate the case \( k = 4 \) a bit more closely.

**Lemma 3.7.** Suppose \( p \geq 5 \) and that \( q^2 \not\equiv 1 \mod p \). Then the field \( L_{4,1} \) contains

\[ \mathbb{Q}(\mu_{pq}, \sqrt[p]{q}) \quad \text{and} \quad \mathbb{Q} \left( \mu_{pq}, \sqrt[p]{\prod_{s=0}^{p-1} \left( 1 - \zeta_s^q \right)^{s^2}} \right) \]

*The intersection of these subfields is \( \mathbb{Q}(\mu_{pq}) \).*

**Proof.** We already get \( \sqrt[p]{q} \) in \( L_{2,1} \) because setting \( i = 0 \) and taking a product over the \( j \) gives

\[ \prod_{j=1}^{q-1} \prod_{s=0}^{p-1} \sqrt[p]{\left( 1 - \zeta_q^j \zeta_p^s \right)^{(r)}} = \prod_{j=1}^{q-1} \sqrt[p]{1 - \zeta_q^{p^j}} = \sqrt[p]{q}. \]

Now do the same with \( i = 2 \) (we suppress the \( p \)-th root symbol for now):

\[ \prod_{j=1}^{q-1} \prod_{s=0}^{p-1} \left( 1 - \zeta_q^j \zeta_p^s \right)^{(r)} = \prod_{s=0}^{p-1} \left( \frac{1 - \zeta_q^{qs}}{1 - \zeta_p^s} \right)^{(r)} \]

Up to an element of \( (L_{1,1}^\times)^p \), we can simplify this expression further using the fact that \( \binom{a}{2} \equiv \binom{b}{2} \mod p \) whenever \( a \equiv b \mod p \). This gives

\[ \prod_{j=1}^{q-1} \prod_{s=0}^{p-1} \left( 1 - \zeta_q^j \zeta_p^s \right)^{(r)} = \prod_{s=0}^{p-1} \left( 1 - \zeta_p^s \right)^{(s^2)} \]

\[ \equiv \prod_{s=1}^{p-1} \left( 1 - \zeta_p^s \right)^{-(\binom{s}{2}) + \binom{p}{2}} \quad (\text{mod } (L_{1,1}^\times)^p) \]

\[ \equiv \prod_{s=1}^{p-1} \left( 1 - \zeta_p^s \right)^{\frac{s^2-1}{2}} \cdot \prod_{s=1}^{p-1} \left( 1 - \zeta_p^s \right)^{\binom{p-1}{2}} \quad (\text{mod } (L_{1,1}^\times)^p) \]

\[ \equiv \left( \prod_{s=1}^{p-1} \left( 1 - \zeta_p^s \right)^{s^2} \right)^{\frac{q^2-1}{2}} \cdot \left( \prod_{s=1}^{p-1} \left( 1 - \zeta_p^s \right)^{s} \right)^{\frac{1-q}{2}} \quad (\text{mod } (L_{1,1}^\times)^p) \]
The first term is \((-1)^{(p^2-1)/8}\), the second term is \(\zeta_p^{p(p^2-1)/24}\). Since \(p \geq 5\), the second term is just 1. In any case, this means that the entire expression is a \(p\)th power. So

\[
\prod_{j=1}^{q-1} \eta_{2,j} \equiv \left( \prod_{s=1}^{p-1} (1 - \zeta_p^s)^{s^2} \right)^{\frac{q^2-1}{2}} \pmod{(L_1^\chi)^p}.
\]

Since by assumption \(p\) is odd and \(q^2 \not\equiv 1 \mod p\), it follows that \(\prod_{s=1}^{p-1} (1 - \zeta_p^s)^{s^2}\) has a \(p\)th root if and only if \(\prod_{j=1}^{q-1} \eta_{2,j}\) does. Since each \(\eta_{2,j}\) is a \(p\)th power in \(L_{4,1}\), it follows that so is \(\prod_{s=1}^{p-1} (1 - \zeta_p^s)^{s^2}\).

For the last part of the lemma, we will show that the intersection of \(\mathbb{Q}(\mu_{pq}, \sqrt[p]{q})\) and \(\mathbb{Q} \left( \mu_{pq}, \sqrt[p]{\prod_{s=0}^{p-1} (1 - \zeta_p^s)^{s^2}} \right)\) is exactly \(\mathbb{Q}(\mu_{pq})\). Note that both of these are Kummer extensions of \(\mathbb{Q}(\mu_{pq})\) and their degrees divide \(p\), so we just need to show that they do not equal each other. Note that the extension \(\mathbb{Q}(\mu_{pq}, \sqrt[p]{q})/\mathbb{Q}(\mu_{pq})\) is totally ramified at \(q\), but the latter extension \(\mathbb{Q} \left( \mu_{pq}, \sqrt[p]{\prod_{s=0}^{p-1} (1 - \zeta_p^s)^{s^2}} \right) / \mathbb{Q}(\mu_{pq})\) is unramified at \(q\). So the two fields are not equal, and hence their intersection must be \(\mathbb{Q}(\mu_{pq})\). \(\square\)

**Proposition 3.8.** Under the assumptions \(p \geq 5\) and \(q^2 \not\equiv 1 \mod p\), we have \([L_{4,1} : L_{1,1}] \geq p^2\).

**Proof.** By Lemma 3.7 we need only check that the subextensions \(\mathbb{Q}(\mu_{pq}, \sqrt[p]{q})/\mathbb{Q}(\mu_{pq})\) and \(\mathbb{Q} \left( \mu_{pq}, \sqrt[p]{\prod_{s=0}^{p-1} (1 - \zeta_p^s)^{s^2}} \right) / \mathbb{Q}(\mu_{pq})\) are nontrivial. The former is nontrivial since it ramifies at \(q\).

So we need only check that the latter is nontrivial. To do so, we need some notation for unit groups of cyclotomic fields. We follow [Was97] for this part. Let \(E\) be the group of units of \(\mathbb{Q}(\zeta_p)^+\), the totally real subfield of \(\mathbb{Q}(\zeta_p)\). Let \(C\) be the subgroup of cyclotomic units. The \(p\)-adic characters of \(\text{Gal}(\mathbb{Q}(\zeta_p)/\mathbb{Q})\) are of the form \(\omega^i\) for \(0 \leq i \leq p - 2\), where \(\omega\) is a Teichmüller character. The \(p\)-adic characters of \(\text{Gal}(\mathbb{Q}(\zeta_p)/\mathbb{Q})\) are of the form \(\omega^i\) where \(i\) is even and in the range \(0 \leq i \leq p - 2\). Let \(\varepsilon_{\omega^i}(E/C)\) be the \(\omega^i\)-isotypic component of \(E/C\). Let \(A\) be the ideal class group of \(\mathbb{Q}(\zeta_p)^+\) and \(\varepsilon_{\omega^i}(A)\) be the \(\omega^i\)-isotypic component of...
A (again, where \( i \) is even). From [Was97] Theorem 15.7 we have
\[
|\varepsilon_{\omega^i}A| = |\varepsilon_{\omega^i}(E/C)_p|
\]
That is, the equality \(|A| = |(E/C)_p|\) (that is, the class number of \( Q(\zeta_p)^+ \) equals the index of the cyclotomic units inside the full unit group) holds component by component. Moreover, the \( \omega^i \)-isotypic component of \((E/C)_p\) is nontrivial if and only if the unit \( E_i \) as defined in [Was97] (Section 8.3, page 155) is a \( p \)th power in \( E \). From a calculation done on the same page, \( E_i \) is a \( p \)th power if and only if the following is:
\[
\prod_{a=1}^{p-1} \left( \zeta_p^{a(1-g)/2} \frac{1 - \zeta_p^{ag}}{1 - \zeta_p^a} \right)^{a^{p-1-i}}.
\]
From a similar argument as in the proof of Lemma 3.7, this expression is a \( p \)th power in \( Q(\zeta_p)\) if and only if
\[
\prod_{a=1}^{p-1} (1 - \zeta_p^a)^{a^{p-1-i}}
\]
is a \( p \)th power in \( Q(\zeta_p) \). (The only difference between the unit group of \( Q(\zeta_p)^+ \) and the unit group of \( Q(\zeta_p) \) is the torsion; i.e, the roots of unity.) So we conclude with \( i = p - 3 \) that the condition that \( \prod_{s=0}^{p-1}(1 - \zeta_p^s)^{s^2} \) not being a \( p \)th power in \( Q(\zeta_p) \) is equivalent to \(|\varepsilon_{\omega^{p-3}}A| = 1\). From [Kur92] Corollary 3.8 we know that \(|\varepsilon_{\omega^{p-3}}A| = 1\), so indeed \( \prod_{s=0}^{p-1}(1 - \zeta_p^s)^{s^2} \) is not a \( p \)th power in \( Q(\zeta_p) \).

A bit of Galois theory wraps up the rest: we have seen so far that the extension
\[
Q \left( \zeta_p, p \sqrt{\prod_{s=0}^{p-1}(1 - \zeta_p^s)^{s^2}} \right)/Q(\zeta_p)
\]
is nontrivial. Moreover the extension \( Q(\mu_{pq})/Q(\zeta_p) \) is disjoint from this extension because the latter is totally ramified at \( q \), while the former is unramified at \( q \). So finally the extension
\[
Q \left( \mu_{pq}, p \sqrt{\prod_{s=0}^{p-1}(1 - \zeta_p^s)^{s^2}} \right)/Q(\mu_{pq})
\]
is nontrivial.

\[\Box\]

4. Application to Torsion Points on \( C \)

Corollary 4.1. Let \( m \) be a positive integer. For each prime \( \ell \) dividing \( m \), choose an element \( \lambda_\ell \) such that
\[
\lambda_\ell \in \begin{cases} Z^*_\ell & \text{if } \ell \not\in \{p, q\} \\ 1 + \ell Z_\ell & \text{if } \ell \in \{p, q\} \end{cases}
\]
Then there exists an element \( \tau \) of \( H_{\infty,m} \) such that for each \( \ell \) dividing \( m \), this element \( \tau \) acts on \( J[\ell^\infty] \) acts by multiplication as \( \lambda_\ell \).
Proof. Let $h$ be any element of $H_{\infty,m}$ such that for each $\ell$ dividing $m$, the restriction of $h$ to $\text{Gal}(\mathbb{Q}(\mu_{\infty})/\mathbb{Q}(\mu_{p^\ell}))$ is the one which raises each element of $\mu_{\infty}$ to the power of $\lambda_\ell$. Choosing $\tau = hh$ then satisfies the desired property by Proposition 2.8.

We will also need a corollary of the Castelnuovo-Severi inequality. We state the Castelnuovo-Severi inequality here as it appears in [Poo07].

**Proposition 4.2** (Castelnuovo-Severi inequality). Let $F, F_1, F_2$ be function fields of curves over $k$, of genera $g, g_1, g_2$, respectively. Suppose that $F_i \subseteq F$ for $i = 1, 2$ and the compositum of $F_1$ and $F_2$ in $F$ equals $F$. Let $d_i = [F : F_i]$ for $i = 1, 2$. Then

$$g \leq d_1g_1 + d_2g_2 + (d_1 - 1)(d_2 - 1).$$

**Corollary 4.3.** Suppose that we have two maps $\mathcal{C} \to \mathbb{P}^1$ of degrees $d_1$ and $d_2$. If $d_1$ and $d_2$ are coprime, then

$$\text{genus}(\mathcal{C}) \leq (d_1 - 1)(d_2 - 1).$$

**Proof.** Let $F$ be the function field of $\mathcal{C}$. Each map gives an embedding of the function field of $\mathbb{P}^1$ into $F$; let their images be $F_1$ and $F_2$. Since $[F : F_i] = d_i$ and the $d_i$ are coprime, it follows that the compositum $F_1F_2$ equals $F$. We apply the Castelnuovo-Severi inequality in this situation with $g_1 = g_2 = 0$ to obtain the result.

4.1. **The non-hyperelliptic case.** The results up till now did not depend on $p, q \geq 3$, but we will impose $p, q \geq 3$ in this section.

**Lemma 4.4.** Let $P_i, Q_i$ be points of $\mathcal{C}$ for $i \in \{1, 2, 3\}$.

1. If $P_1 + P_2 \sim Q_1 + Q_2$, then $\{P_1, P_2\} = \{Q_1, Q_2\}$.
2. Suppose $p, q \geq 5$. If $P_1 + P_2 + P_3 \sim Q_1 + Q_2 + Q_3$, then $\{P_1, P_2, P_3\} = \{Q_1, Q_2, Q_3\}$.

**Proof.** Let $f$ be a rational map $f : \mathcal{C} \to \mathbb{P}^1$ such that in case (1) we have $\text{div}(f) = P_1 + P_2 - Q_1 - Q_2$ and in case (2) we have $\text{div}(f) = P_1 + P_2 + P_3 - Q_1 - Q_2 - Q_3$. (1) Suppose $\{P_1, P_2\} \neq \{Q_1, Q_2\}$. Let $f$ be a rational map $f : \mathcal{C} \to \mathbb{P}^1$ such

$$\text{div}(f) = P_1 + P_2 - Q_1 - Q_2$$

Then $f$ is either a degree 1 map or a degree 2 map to $\mathbb{P}^1$. We also have degree $p$ and $q$ maps to $\mathbb{P}^1$ via the $x$-map and $y$-map, respectively. By Corollary 4.3 this means that

$$g \leq (2 - 1)(p - 1) \text{ and } g \leq (2 - 1)(q - 1).$$

Since $g = (p - 1)(q - 1)/2$, this would mean that $q \leq 3$ and $p \leq 3$, a contradiction since $p, q$ are distinct odd primes.

(2) In the same way as the proof of the previous part, we assume $\{P_1, P_2, P_3\} \neq \{Q_1, Q_2, Q_3\}$ to obtain a rational map $f : \mathcal{C} \to \mathbb{P}^1$ such

$$\text{div}(f) = P_1 + P_2 + P_3 - Q_1 - Q_2 - Q_3.$$

Then the degree of $f$ is at most 3, and since $p, q \geq 5$ we know that Corollary 4.3 gives

$$g \leq (3 - 1)(p - 1) \text{ and } g \leq (3 - 1)(q - 1)$$

from which we get that $p, q \leq 5$ which contradicts our assumption that $p, q$ are distinct primes that are at least 5.

□
Then by construction, we see that

\[ \text{Lemma 4.4 (1) it follows that} \]

Proof. Suppose \( m \) is a positive integer such that \( m(P - \infty) \sim 0 \). Without loss of generality, suppose \( m \) is divisible by \( 2pq \). Let \( \{r_1, \ldots, r_M\} \) be the set of primes that divide the prime-to-\( 2pq \)-part of \( m \).

Using Corollary 4.1, choose \( \tau_1, \tau_2, \tau_3 \in \text{Gal}(\mathbb{Q}(J[m^\infty])/\mathbb{Q}(\mu_{pq})) \) such that

\[
\begin{align*}
\tau_1 \text{ acts on} & \quad J[2^\infty] \text{ as multiplication by } 1 + 2 \\
& \quad J[p^\infty] \text{ as multiplication by } 1 + p \\
& \quad J[q^\infty] \text{ as multiplication by } 1 + q \\
& \quad J[r_i^\infty] \text{ as multiplication by } 2 \\
\tau_2 \text{ acts on} & \quad J[2^\infty] \text{ as multiplication by } 1 - 2 \\
& \quad J[p^\infty] \text{ as multiplication by } 1 - p \\
& \quad J[q^\infty] \text{ as multiplication by } 1 - q \\
& \quad J[r_i^\infty] \text{ as multiplication by } -2 \\
\tau_3 \text{ acts on} & \quad J[2^\infty] \text{ as multiplication by } 1 \\
& \quad J[p^\infty] \text{ as multiplication by } 1 \\
& \quad J[q^\infty] \text{ as multiplication by } 1 \\
& \quad J[r_i^\infty] \text{ as multiplication by } -1
\end{align*}
\]

Then by construction, we see that \((\tau_1 + \tau_2) - (\tau_3 + 1)\) acts as the identity on \( J[m^\infty] \). In particular,

\[ \tau_1 P + \tau_2 P \sim \tau_3 P + \tau_1 P \sim \tau_3 P + P. \]

Then by Lemma 4.4 (1) it follows that \( P \) is either \( \tau_1 P \) or \( \tau_2 P \).

If \( P = \tau_1 P \), then writing \( P - \infty = D_2 + D_p + D_q + \sum D_{\ell} \) for divisors \( D_\ell \in J[\ell^\infty] \), we see that \( \tau_1 D_\ell = D_\ell \) for each \( \ell \). So in particular, \( 2D_2 \sim 0, pD_p \sim 0, qD_q \sim 0, \) and \( D_{r_i} \sim 0 \) for each \( i \). Hence \( 2pq(P - \infty) \sim 0 \).

If \( P = \tau_2 P \), then a similar analysis shows that either (i) \( 2pq(P - \infty) \sim 0 \) or (ii) \( p, q \geq 5 \) and \( 6pq(P - \infty) \sim 0 \).

So the last case to consider is \( p, q \geq 5 \) and \( 6pq(P - \infty) \sim 0 \). In that case, find \( \tau_4 \in \text{Gal}(\mathbb{Q}(J[m^\infty])/\mathbb{Q}(\mu_{pq})) \) such that \( \tau_4 \) acts on \( J[3^\infty] \) as multiplication by \(-1\) and \( \tau_4 \) acts on \( J[(2pq)^\infty] \) as the identity. Then \( 3P \sim 3\tau_4 P \). Then by Lemma 4.4 (2) it follows that \( P = \tau_4 P \), so a similar analysis as before shows again that \( 2pq(P - \infty) \sim 0 \).

Next, we would like to remove the “2” in the statement of Proposition 4.5. To do so, we need to study ramification in torsion fields.

Lemma 4.6. We have the following.

1. The torsion field \( \mathbb{Q}(J[2]) \) is ramified at 2.
2. Suppose \( D \) is a nonzero element of \( J[2] \). Then the field \( \mathbb{Q}(D, \mu_{pq}) \) is ramified at 2.
3. The torsion field \( \mathbb{Q}(J[pq]) \) is unramified at 2.

Proof.
(1) From \texttt{Jęd16} applied with \(a = 1\) we know that the reduction of the jacobian \(J\) at 2 is not ordinary. Applying Lemma 1.4 of \texttt{GR78} now tells us that \(\mathbb{Q}(J[2])\) is ramified at 2.

(2) From Lemma \texttt{2.16} we know that \(\mathbb{Q}(D, \mu_{pq}) = \mathbb{Q}(J[2])\). So we are done by the previous part and Lemma \texttt{2.9}.

(3) This follows from the criterion of Néron-Ogg-Shafarevich. \(\Box\)

**Proposition 4.7.** If \(P\) is a torsion point on \(C\), then \(pq(P - \infty) \sim 0\).

**Proof.** From Proposition \texttt{4.5} we know that \(P - \infty = D_2 + D_p + D_q\), where \(2D_2, pD_p, qD_q \sim 0\). Suppose \(D_2 \neq 0\).

From Lemma \texttt{4.6} it follows that \(\mathbb{Q}(J[pq])\) cannot contain \(\mathbb{Q}(D_2, \mu_{pq})\). Hence we can find a \(\tau \in \text{Gal}(\mathbb{Q}(J[2pq])/\mathbb{Q}(J[pq]))\) which acts nontrivially on \(\mathbb{Q}(D_2, \mu_{pq})\). Since \(\mathbb{Q}(\mu_{pq}) \subseteq \mathbb{Q}(J[pq])\) (due to the Weil pairing) it follows that \(\tau\) must act nontrivially on \(D_2\).

Hence \(D_2 \neq \tau D_2\) which implies \(P \neq \tau P\) and yet \(2(P - \tau P) = 2(D_2 - \tau D_2) \sim 0\), which violates Lemma \texttt{4.4}. This contradiction implies \(D_2 = 0\), as desired. \(\Box\)

**Definition 4.8.** Choose \(a, b\) minimal such that 
\[
(1 - \zeta_p)^a (1 - \zeta_q)^b P \sim 0.
\]

Define \(D_p, D_q\) such that \(P - \infty \sim D_p + D_q, pD_p \sim 0\), and \(qD_q \sim 0\).

In order to get a contradiction whenever \(a\) and \(b\) are large, we will use an argument with inflectionary weights of Weierstrass points. The following definitions can be found in an introductory book on Riemann surfaces, e.g. \texttt{FK92}.

**Definition 4.9.** Given a point \(R\) on a nonsingular algebraic curve \(X\) of genus \(g\), an integer \(k\) is a gap of \(R\) if there is no rational function on \(X\) with a pole at \(R\) of exact order \(k\). By Riemann-Roch, there will be exactly \(g\) gaps and they will lie in the range \([1, 2g - 1]\). The set of non-gaps forms a monoid, denoted by \(\text{WM}(R)\), the Weierstrass monoid of \(R\). If the gaps of \(R\) are \(k_1 < k_2 < \cdots < k_g\), then the inflectionary weight of \(R\) is 
\[
\text{wt}(R) = \sum_{i=1}^{g} (k_i - i).
\]
The point \(R\) is called a Weierstrass point of \(X\) if \(\text{wt}(R) > 0\).

We now use a basic result about Weierstrass points on a Riemann surface, found in \texttt{Mir95} as Corollary 4.17.

**Theorem 4.10.** The sum of the inflectionary weights of all the Weierstrass points on a Riemann surface \(X\) of genus \(g\) is \(g^3 - g\).

**Lemma 4.11.** Define 
\[
S_P = \{hzP : h \in \text{Gal}(\mathbb{Q}(J[pq])/\mathbb{Q}(\mu_{pq})), z \in \mathbb{Z}\}
\]

(1) If \(a \geq 2\) and \(b \geq 1\), then \(S_P\) has size at least \(pq[L_{a,1} : L_{1,1}]\) and for each \(Q \in S_P\) we have \(p - 1, p \in \text{WM}(Q)\).

(2) If \(a \geq 1\) and \(b \geq 2\), then \(S_P\) has size at least \(pq[L_{1,b} : L_{1,1}]\) and for each \(Q \in S_P\) we have \(q - 1, q \in \text{WM}(Q)\).
Proof. Both parts are similar so we show the first.

Define
\[ E = D_p + (1 - \zeta_q)^{b-1}D_q \]
Since \( a \geq 2, b \geq 1 \), we know that \( E \) is a divisor of exact order \( (1 - \zeta_p)^a(1 - \zeta_q) \). Therefore \( E \) is defined over \( L_{a,1} \).

To show that \( |S_P| \geq pq[L_{a,1} : L_{1,1}] \), we instead show the stronger statement that
\[ S_E = \{hzE : h \in \text{Gal}(Q(D_p,\mu_{pq})/Q(\mu_{pq})), z \in \mathbb{Z} \} \]
already has size exactly equal to
\[ [Q(D_p,\mu_{pq}) : Q(\mu_{pq})] : |Z|. \]
(2)

By Lemma 2.16 we know that \( Q(D_p,\mu_{pq}) = L_{a,1} \) so this latter number is exactly equal to \([L_{a,1} : L_{1,1}]pq\). To do so, we need to check that all the elements \( hzE \) are distinct. Since \( \text{Gal}(Q(D_p)/Q(\mu_{pq})) \) is abelian and commutes with \( Z \), it suffices to check that if \( hE = zE \), then \( h = 1 \) and \( z = 1 \).

So assume now that \( hE = zE \). Since \( \zeta_p \) commutes with \( h, z \) we have that
\[ h(1 - \zeta_p)^{a-1}E = z(1 - \zeta_q)^{a-1}E. \]

But \((1 - \zeta_p)^{a-1}E\) is a \((1 - \zeta_p)(1 - \zeta_q)\)-torsion divisor, and is hence defined over \( Q(\mu_{pq}) \), so \( h \) is forced to fix it. Hence
\[ (1 - \zeta_p)^{a-1}E = z(1 - \zeta_q)^{a-1}E. \]

But as \((1 - \zeta_p)^{a-1}E\) has exact order \((1 - \zeta_p)(1 - \zeta_q)\), the only element of \( Z \) that can fix it is \( 1 \); hence, \( z = 1 \).

As \( z = 1 \), we now assume \( hE = zE = E \). In particular, \( h \) also fixes \( D_p \). Hence \( h = 1 \) as well. We have now shown that
\[ |S_P| \geq |S_E| = [Q(D_p,\mu_{pq}) : Q(\mu_{pq})] : |Z| = pq[Q(D_p,\mu_{pq}) : Q(\mu_{pq})] \]
which suffices to check that \( p - 1, p \in \text{WM}(P) \). For this, let \( h \in \text{Gal}(J[pq]/Q(\mu_{pq})) \) be such that its restriction to \( \text{Gal}(Q(D_p,\mu_{pq})/Q(\mu_{pq})) \) is nontrivial and its restriction to \( \text{Gal}(Q(J[q])/Q(\mu_{pq})) \) is trivial. (This can be done since \( a \geq 2 \).) Then
\[ h^iP \neq P \text{ for } 0 \leq i \leq p - 1. \]

(We know that \( h^p = 1 \) since the entire extension is \( p \)-Kummer.)

Since \( h \) fixes the \( q \)-torsion, we know that \( h(pP) \sim pP \). Therefore, \( pP \sim p(hP) \) implies that \( p \in \text{WM}(P) \). Moreover, we also see that \( p \in \text{WM}(Q) \) for all \( Q \in S_P \).

Moreover, note that \( 1 + h + h^2 + \cdots + h^{p-1} \) is an endomorphism of \( J[pq] \). From Lemma 2.12 we know that for all \( \chi, \xi_\chi(h) \in 1 + \mathfrak{m}_p \). Therefore,
\[ 1 + \xi_\chi(h) + \xi_\chi(h)^2 + \cdots + \xi_\chi(h)^{p-1} \in \mathfrak{m}_p^{p-1} = p\mathcal{O}_p. \]

So again by Lemma 2.12 we know that \( 1 + h + h^2 + \cdots + h^{p-1} \) acts trivially on \( J[p] \). Since \( h \) acts trivially on \( J[q] \), we conclude that
\[ (1 + h + h^2 + \cdots + h^{p-1}) - p \text{ acts trivially on } J[pq] \]
Therefore
\[ hP + h^2P + \cdots + h^{p-1}P \sim (p - 1)P \]
and as \( P \neq h^iP \) (by equation (2)) we see that \( p - 1 \in \text{WM}(P) \) as well. Since this argument only used the fact that \( P \in J[pq] \), we see it also applies to all \( Q \in S_P \). \( \square \)
Proposition 4.12.

(1) If $a \geq 2$ and $b \geq 1$, then we must have $q = 3$ and $a \in \{2, 3\}$.

(2) If $a \geq 1$ and $b \geq 2$, then we must have $p = 3$ and $b \in \{2, 3\}$.

Proof. Both parts are similar so we prove the first. By Lemma 4.11, there are at least $pq[L_{a,1} : L_{1,1}]$ points $P$ such that $p - 1, p \in \text{WM}(P)$.

We first obtain a lower bound on $\text{wt}(P)$ for such $P$. Since $p - 1, p \in \text{WM}(P)$, we know that $u(p - 1) + vp \in \text{WM}(P)$ for any $u, v \geq 0$. In particular, we know that

$$\{p - 1, p, 2p - 2, 2p - 1, 2p, 3p - 3, 3p - 2, 3p - 1, 3p, \cdots \} \subseteq \text{WM}(P).$$

Therefore, a lower bound on the weight of $P$ is

$$\begin{align*}
\text{wt}(P) &= \sum_{i=1}^{g} (k_i - i) \\
&\geq (1 - 1) + (2 - 2) + ((p - 2) - (p - 2)) \\
&\quad + ((p + 1) - (p - 1)) + ((p + 2) - (p)) + \ldots + ((2p - 3) - (2p - 5)) \\
&\quad + ((2p + 1) - (2p - 4)) + ((2p + 2) - (2p - 3)) + \ldots + ((3p - 4) - (3p - 9)) \\
&\quad + \ldots \\
&= 0 + \ldots + 0 + 2 + \ldots + 2 + 5 + \ldots + 5 + 9 + \ldots + 9 + \ldots \\
&\quad + 0 + \ldots + 0 + 2 + \ldots + 2 + 4 + \ldots + 4 + 6 + \ldots + 6 + \ldots \\
&= \left((p - 1) \left(0 + 2 + 4 + 6 + \ldots + 2 \left(\frac{q - 1}{2} - 1\right)\right)\right) \\
&= \left((p - 1)(q - 3)(q - 1)\right) \\
&= g \left(\frac{q - 3}{2}\right).
\end{align*}$$

By Lemma 3.3 (4) we know that $[L_{a,1} : L_{1,1}] \geq [L_{2,1} : L_{1,1}] \geq p$, so we have at least $p^2q$ of these points. Hence the total weight of all points on $C$ is at least

$$g \left(\frac{q - 3}{2}\right) p^2q.$$

If $q \geq 5$, then we know that $q(q - 3) \geq \frac{5}{8}(q - 1)^2$ which means the total weight is at least

$$g \left(\frac{q - 3}{2}\right) p^2q \geq g \left(\frac{5}{16}(q - 1)^2\right) p^2 \geq \frac{5}{16}g((p - 1)(q - 1))^2 \geq \frac{5}{4}q^3.$$

This contradicts Theorem 4.10, which states that the total weight of all points on $C$ is $g^3 - g$.

Hence $q = 3$. If $a \geq 4$, then we know from Proposition 3.8 that $[L_{a,1} : L_{1,1}] \geq [L_{4,1} : L_{1,1}] \geq p^2$, so we have at least $p^2q$ of these points of weight at least

$$\text{wt}(P) \geq (1 - 1) + (2 - 2) + \cdots + ((p - 2) - (p - 2)) + ((p + 1) - (p - 1)) = 2.$$
which means that the total weight is at least \(2p^3q = 6p^3\). Since \(g = (p-1)(q-1)/2 = p-1\) and Theorem 4.10 states that the total weight of all points on \(C\) is \(g^3 - g = (p-1)^3 - (p-1) < p^3\), we have yet again a contradiction.

So the only remaining possibility is \(q = 3\) and \(a \in \{2, 3\}\). \(\square\)

**Proposition 4.13.**

1. It is impossible for \(a = b = 1\).
2. If \(a = 0\), then \(b \leq 1\).
3. If \(b = 0\), then \(a \leq 1\).

**Proof.**

(1) Suppose \(a = b = 1\). Then

\[
(1 - \zeta_p)(1 - \zeta_q)P \sim 0
\]

which we rearrange to get

\[
P + \zeta_p\zeta_qP \sim \zeta_pP + \zeta_qP.
\]

From Lemma 4.4 it follows that either \(P = \zeta_pP\) or \(P = \zeta_qP\), meaning that either \(a\) or \(b\) is 0.

(2) Suppose \(a = 0\) and \(b \geq 1\). We seek to show that \(b = 1\).

Then \(qP \sim q\infty\). Let \(f\) be a function such that

\[
\text{div}(f) \sim qP - q\infty.
\]

Since \(f\) only has poles at \(\infty\), it follows that \(f\) is a polynomial in \(x\) and \(y\). Since the pole order is \(q\), it follows that \(f(x, y) = y - g(x)\) where \(\text{deg}(g) < q/p\). Let \(x_P\) be the \(x\)-coordinate of \(P\). From this it follows that

\[
\text{div} \left( \prod_{i=0}^{p-1} (\zeta_p^i y - g(x)) \right) = q \left( \sum_{i=0}^{p-1} \zeta_p^i P \right) - pq\infty
\]

Moreover we also have

\[
\text{div} ((x - x_P)^q) = q \left( \sum_{i=0}^{p-1} \zeta_p^i P \right) - pq\infty,
\]

so it follows that \(\prod_{i=0}^{p-1} (\zeta_p^i y - g(x))\) and \((x - x_P)^q\) are the same up to a scalar. Simplifying the former expression, we see that

\[
\prod_{i=0}^{p-1} (\zeta_p^i y - g(x)) = y^p - g(x)^p = x^q + 1 - g(x)^p,
\]

so we conclude

\[
(*) \quad x^q + 1 - g(x)^p = (x - x_P)^q.
\]

Rewrite this as

\[
x^q + 1 - (x - x_P)^q = g(x)^p.
\]

If \(g\) is nonconstant, then the right hand side has at least one root of order at least 3. However, this is not true of the left hand side: to see this, let \(L(x) = x^q + 1 - (x - x_P)^q\) and note

\[
L(x) - \frac{1}{q}(x - x_P)L'(x) = x_p x^{q-1} + 1
\]
With the help of a computer \cite{Dev19}, it does not take long to find such an \( R \) with the smallest possible such \( \eta \) then shows that \( x_P = 0 \). Hence \( P \) is a \((1 - \zeta_3)\) torsion point, forcing \( b = 1 \).

(3) Similar to the proof of the previous part.

Combining Propositions \ref{prop:4.12} and \ref{prop:4.13} the only cases we have left to consider are

(1) \( q = 3, a \in \{2, 3\}, b = 1 \)
(2) \( p = 3, a = 1, b \in \{2, 3\} \)

Both cases are similar so we handle the first. Hence from now on we suppose that \( q = 3, b = 1, a \in \{2, 3\} \), and \( p \geq 5 \). We have that

\[(1 - \zeta_3(1 - \zeta_3)P \sim 0, \]

which we rewrite as

\[(\zeta_3^3 - \zeta_3^3 - 3\zeta_3^2 \zeta_3 + 3\zeta_3^2 + 3\zeta_3 - 3\zeta_3 + 1)P \sim 0 \]

which we can rewrite as

\[\zeta_3^3 \zeta_3 P + 3\zeta_3^2 P + 3\zeta_3 P + P \sim \zeta_3^3 P + 3\zeta_3^2 \zeta_3 P + 3\zeta_3 P + \zeta_3 P \]

Since \( p \geq 5 \) and \( q = 3 \), the only way for any of these points \( \{\zeta_3^2 \zeta_3 P\} \) to equal another is for \( P \) to be in either \( J[1 - \zeta_3] \) or \( J[1 - \zeta_3] \). Therefore from this we get a degree 8 map to \( \mathbb{P}^1 \).

Since we also have a degree 3 map to \( \mathbb{P}^1 \), we know from Corollary \ref{cor:4.3} that

\[g \leq (3 - 1)(8 - 1).\]

Since \( g = (3 - 1)(p - 1)/2, \) this means that

\[\frac{p - 1}{2} \leq 8 - 1,\]

so \( p \leq 15 \). Therefore we need only check that at the primes \( p \in \{5, 7, 11, 13\} \) that there are no points \( P \in J[(1 - \zeta_3(1 - \zeta_3)] \) in order to finish.

For the remaining three curves, the first step will be to compute explicitly the Galois action on \( T_p J \) to find that \( L_{3.1}/L_{2.1}/L_{1.1} \) is a tower where each successive step is a nontrivial \( p \)-extension. The bottom extension \( L_{2.1}/L_{1.1} \) is known to be nontrivial by Lemma \ref{lem:3.2} (4). So we need to show that the top extension is nontrivial.

The strategy will be to find primes \( r \) such that for some prime \( \mu \) of \( \mathbb{Q}(\mu_{3p}) \) lying above \( r \), we have \( \xi_\mu(\text{Frob}_\mu) - 1 \) always has \( \pi_p \)-adic valuation 2. Then by Lemma \ref{lem:2.12}, we will know that \( \text{Frob}_\mu \) acts trivially on \( J[(1 - \zeta_3) \] but not on \( J[(1 - \zeta_3)^3)] \). In other words, \( \text{Frob}_\mu \) will be a nontrivial element of \( \text{Gal}(L_{3.1}/L_{2.1}) \). By Theorem \ref{thm:3.4}, we are searching for finite fields \( \mathbb{F}_R \) with \( R \equiv 1 \mod 3p \) where

\[1 - \zeta_3, 1 - \zeta_3^2 \in \mathbb{F}_R^p \]

\[\eta_{3.1} \text{ or } \eta_{3.2} \notin \mathbb{F}_R^p. \]

With the help of a computer \cite{Dev19}, it does not take long to find such \( R \). Here is a table with the smallest possible such \( R \) satisfying these conditions for \( p \in \{5, 7, 11, 13\} \).

\[
\begin{array}{c|c|c|c|c}
 p & 5 & 7 & 11 & 13 \\
 R & 2^4 & 13^2 & 43^2 & 547 \\
\end{array}
\]
Now we can wrap up with one final lemma.

**Lemma 4.14.** The cases

1. \( q = 3, p \in \{5, 7, 11, 13\}, a \in \{2, 3\}, b = 1 \)
2. \( p = 3, q \in \{5, 7, 11, 13\}, a = 1, b \in \{2, 3\} \)

are impossible.

**Proof.** Both cases are similar so we handle the first. Suppose \( a = 3 \). By our computation, there exists a nontrivial \( \gamma \in L_{3,1}/L_{2,1} \). By Lemma 2.16 we know that \( L_{3,1} = L_{2,1}(D_p) \), so \( \gamma \) must move \( D_p \) and hence it must move \( P \). Since

\[
\xi_\chi(\gamma) \in 1 + m_p^3
\]

for every \( \chi \), we know that

\[
\xi_\chi(\gamma) + \xi_\chi(\gamma^{-1}) - 1 \in 1 + m_p^6
\]

and hence \( \gamma + \gamma^{-1} - 1 \) must fix \( P \). (We are using Lemma 2.12 repeatedly.) So we can write

\[
\gamma P + \gamma^{-1} P \sim P + P,
\]

and now by Lemma 4.4 we know that \( P \) must be either \( \gamma P \) or \( \gamma^{-1} P \), which is a contradiction.

If \( a = 2 \), we can do a very similar argument by picking \( \gamma \) to be a nontrivial element of \( \text{Gal}(L_{2,1}/L_{1,1}) \).

\[\square\]

4.2. **Extension of results to** \( y^{p^i} = x^{q^j} + 1 \). Before we proceed to the case when \( p = 2 \), it is helpful to extend our results to the curve \( y^{p^i} = x^{q^j} + 1 \). From now on, we assume that \( p \) and \( q \) are distinct primes (they are no longer required to be odd).

For convenience, we make the following definition.

**Definition 4.15.** For coprime integers \( n \) and \( d \), let \( C_{n,d} \) be the curve \( y^n = x^d + 1 \) and let \( J_{n,d} \) be its Jacobian.

The results of section 2 apply generally unchanged to \( C_{n,d} \) as well. First, we let \( O_\ell \) be the ring of integers of some extension of \( \mathbb{Q}_\ell \) containing a primitive \( nd \)-th root of unity. The group \( Z \) generated by \((x, y) \mapsto (\zeta_d x, \zeta_d y)\) acts on this curve. For each character \( \chi : Z \to O_\ell^* \), we may consider the \( \chi \)-isotypic component of \( T_\ell J \otimes Z_\ell O_\ell \) and denote this by \( T_\chi \). Letting \( H_{\infty,m} = \text{Gal}(\mathbb{Q}(\mu_{nd}, J[m^{\infty}])/\mathbb{Q}) \) as before, we see that the \( H_{\infty,m} \)-action on \( T_\ell J \) must preserve the \( T_\chi \). In particular, \( H_{\infty,m} \) is abelian.

From an argument similar to those in Section 2 of [Kat81] (where Katz considers the Fermat curve \( X^N + Y^N = Z^N \)), one can show that whenever \( \chi^n \) and \( \chi^d \) are nontrivial, then \( T_\chi \) has dimension 1. Otherwise, \( T_\chi \) has dimension 0. In particular, we again get characters \( \xi_\chi : H_{\infty,\ell} \to \text{Aut} T_\chi \simeq O_\ell^* \). Moreover, techniques in Section 2 of [Kat81] will show that \( \xi_\chi(\text{Frob}_r) \) may be expressed in terms of a Jacobi sum.

For every \( h \in H_{\infty,\ell} \) we define \( \overline{h} \) in the exact same way as before (namely, let \( \sigma \in \text{Gal}(\mathbb{Q}(\mu_{nd})/\mathbb{Q}) \) be complex conjugation, and let \( \overline{h} = \sigma h \sigma^{-1} \); this is well-defined since \( H_{\infty,\ell} \) is abelian). The proofs of Proposition 2.8 works unmodified, and the same method as the proof of Corollary 4.1 now shows the following.
Corollary 4.16. Let $m$ be a positive integer. For every prime $\ell$ dividing $m$, let $\ell^{v(\ell)}$ be the highest power of $\ell$ dividing $nd$. For every prime $\ell$ dividing $m$, choose an element $\lambda_\ell \in \mathbb{Z}_\ell^\times$ such that

$$\lambda_\ell \in \begin{cases} 
\mathbb{Z}_\ell^\times & \text{if } \ell \nmid nd \\
1 + \ell^{v(\ell)}\mathbb{Z}_\ell & \text{if } \ell | nd 
\end{cases}$$

Then there exists an element $\tau$ of $H_{\infty,m}$ such that for each $\ell$ dividing $m$, this element $\tau$ acts on $J[\ell^\infty]$ acts by multiplication as $\lambda_\ell$.

We can now extend Proposition 4.5.

Proposition 4.17. Suppose $\mathcal{C}_{n,d}$ has genus $g > 1$ (i.e, $(n,d) \not\in \{(2,3),(3,2)\})$. Let $m = \text{lcm}(2,nd)$. Suppose $P$ is a torsion point on $\mathcal{C}_{n,d}$. Then we have the following.

1. If $(n,d) \not\in \{(2,5),(4,5),(5,2),(5,4)\}$ then $m(P - \infty) \sim 0$.
2. If $(n,d) \in \{(2,5),(5,2)\}$ then $3m(P - \infty) \sim 0$.
3. If $(n,d) \in \{(4,5),(5,4)\}$ then $3m(P - \infty) \sim 0$.

Proof. Choose an integer $M$ such that $M(P - \infty) \sim 0$. Assume $M$ is divisible by $m$. Let $r_1, \ldots, r_k$ be the primes that divide the prime-to-2nd part of $M$. Let the prime factorization of $nd$ be $nd = 2^{e_2}s_1^{e_{s_1}}s_2^{e_{s_2}} \cdots s_l^{e_{s_l}}$ for odd primes $s_j$.

Using Corollary 4.16 choose $\tau_1, \tau_2, \tau_3 \in \text{Gal}(\mathbb{Q}(J[M^\infty])/\mathbb{Q}(\mu_{nd}))$ such that:

$\tau_1$ acts on \[
\begin{cases} 
J[2^{\infty}] & \text{as multiplication by } 1 + 2^{\max\{1,e_2\}} \\
J[r_i^{\infty}] & \text{as multiplication by } 2 \\
J[s_j^{\infty}] & \text{as multiplication by } 1 + s_j^{e_{s_j}}
\end{cases}
\]

$\tau_2$ acts on \[
\begin{cases} 
J[2^{\infty}] & \text{as multiplication by } 1 - 2^{\max\{1,e_2\}} \\
J[r_i^{\infty}] & \text{as multiplication by } -2 \\
J[s_j^{\infty}] & \text{as multiplication by } 1 - s_j^{e_{s_j}}
\end{cases}
\]

$\tau_3$ acts on \[
\begin{cases} 
J[2^{\infty}] & \text{as multiplication by } 1 \\
J[r_i^{\infty}] & \text{as multiplication by } -1 \\
J[s_j^{\infty}] & \text{as multiplication by } 1
\end{cases}
\]

Then $\tau_1P + \tau_2P \sim \tau_3P + P$. Unless $\{\tau_1P,\tau_2P\} = \{\tau_3P,P\}$, this induces a degree $f$ map $\theta_1 : \mathcal{C}_{n,d} \to \mathbb{P}^1$ where $f \leq 2$. Without loss of generality, assume that $d$ is odd. By the Castelnuovo-Severi inequality applied with $\theta$ and the $y$-map,

$$\frac{(n-1)(d-1)}{2} = g \leq (f-1)(d-1).$$

Since $d > 1$, this implies that $n-1 \leq 2(f-1)$. The only way this could happen is if $f = 2$ and $n \in \{2,3\}$. We break up into two cases.

1. $n = 2$. This curve is hyperelliptic, so any $2 : 1$ map to $\mathbb{P}^1$ must factor through the canonical map. That would then force $\tau_1P + \tau_2P \sim 2\infty$ and $\tau_3P + P \sim 2\infty$. This means that $P \in J[2(\prod r_i)^{\infty}]$. Using Corollary 4.16 choose $\tau_4, \tau_5 \in \text{Gal}(\mathbb{Q}(J[M^\infty])/\mathbb{Q}(\mu_{nd}))$
such that:

\[ \begin{cases} J[2^\infty] & \text{as multiplication by 1} \\ J[r_i^\infty] & \text{as multiplication by 1 if } r_i = 3 \\ J[r_i^\infty] & \text{as multiplication by 3 if } r_i \neq 3 \end{cases} \]

\[ \begin{cases} J[2^\infty] & \text{as multiplication by 1} \\ J[r_i^\infty] & \text{as multiplication by 1 if } r_i = 3 \\ J[r_i^\infty] & \text{as multiplication by } -1 \text{ if } r_i \neq 3 \end{cases} \]

By construction, \( \tau_4 P + \tau_5 P \sim 2P \). Since the curve is hyperelliptic, this implies either that \( 2(P - \infty) \sim 0 \) (in which case we are done) or \( P = \tau_4 P = \tau_5 P \). In the last case, we must have \( P \in J[2 \cdot 3^\infty] \). In particular, this means that there was one \( i \) with \( r_i = 3 \), so that \( \text{gcd}(d, 3) = 1 \).

Write \( P - \infty \sim D_2 + D_3 \) where \( D_2 \in J[2] \) and \( D_3 \in J[3^\infty] \). Choose an integer \( w \geq 1 \) such that \( 3^w D_3 \sim 0 \). For contradiction, suppose that \( D_3 \) has exact order \( 3^w \), so that \( 3^w-1 D_3 \neq 0 \).

Using Corollary 4.16, choose \( \tau_6 \in \text{Gal}(\mathbb{Q}(J[M^\infty])/\mathbb{Q}(\mu_{nd})) \) such that:

\[ \begin{cases} J[2^\infty] & \text{as multiplication by 1} \\ J[3^\infty] & \text{as multiplication by 1 + 3}^{w-1} \end{cases} \]

Then \( \tau_6 P - P \sim 3^{w-1} D_3 \), so that \( 3(\tau_6 P - P) \sim 0 \). From this we get \( 3\tau_6 P \sim 3P \). By construction we know that \( \tau_6 P \neq P \) (their difference is \( 3^{w-1} D_3 \), which is nonzero by choice of \( w \)). So the linear equivalence \( 3\tau_6 P \sim 3P \) produces a nonconstant 3 : 1 map to \( \mathbb{P}^1 \), and applying Castelnuovo-Severi now gives \( g \leq (2 - 1)(3 - 1) = 2 \). That means that \( d = 5 \).

To summarize, we write \( P - \infty \sim D_2 + D_3 \) for \( D_2 \in J[2] \) and \( D_3 \in J[3] \). If \( d \neq 5 \) then \( D_3 = 0 \). In any case, the conclusion of this proposition holds.

(2) \( n = 3 \). In this case, we know that \( f \) and \( n \) are coprime, so applying Castelnuovo-Severi with the \( f : 1 \) map and the \( x \)-map to \( \mathbb{P}^1 \) gives

\[
\frac{(3 - 1)(d - 1)}{2} = g \leq (2 - 1)(3 - 1)
\]

This forces \( d = 3 \) which is impossible since \( n \) and \( d \) are coprime.

The remaining possibility is \( \{ \tau_1 P, \tau_2 P \} = \{ \tau_3 P, P \} \). Write \( P - \infty = D_{2nd} + \sum_i D_{ri} \) where \( D_{2nd} \in J[(2nd)^\infty] \) and \( D_{ri} \in J[r_i^\infty] \).

If \( P = \tau_1 P \) then \( D_{2nd} \) and the \( D_{ri} \) are also fixed by \( \tau_1 \). This forces \( D_{ri} \sim 0 \) and \( mD_{2nd} \sim 0 \). We conclude that \( m(P - \infty) \sim 0 \).

If \( P = \tau_2 P \) then as before we know that \( mD_{2nd} \sim 0 \), \( D_{ri} \sim 0 \) whenever \( r_i \neq 3 \), and \( 3D_3 \sim 0 \) if there exists an \( i \) such that \( r_i = 3 \). So we are done unless there is an \( i \) such that \( r_i = 3 \).

Suppose \( r_i = 3 \), \( 3D_3 \sim 0 \), \( D_{ri} \sim 0 \) whenever \( r_i \neq 3 \), and \( mD_{2nd} \sim 0 \). Using Corollary 4.16 choose \( \tau_7 \in \text{Gal}(\mathbb{Q}(J[M^\infty])/\mathbb{Q}(\mu_{nd})) \) such that:

\[ \begin{cases} J[(2nd)^\infty] & \text{as multiplication by 1} \\ J[3] & \text{as multiplication by } 2 \end{cases} \]

Then \( 3(P - \tau_7 P) \sim 0 \). If \( P = \tau_7 P \) then we know that \( D_3 \sim 0 \) and we are done.
If \( P \neq \tau_7 P \) yet \( 3(P - \tau_7 P) \sim 0 \), then we get a nontrivial \( 3 : 1 \) map to \( \mathbb{P}^1 \). Since \( r_i = 3 \), we know that \( nd \) is coprime to \( 3 \). Again by Castelnuovo-Severi, we obtain
\[
\frac{(n-1)(d-1)}{2} = g \leq \min\{(3-1)(n-1), (3-1)(d-1)\},
\]
which forces \( n, d \leq 5 \). Since \( d \) is odd and \( d \neq 3 \), this means \( d = 5 \). The remaining possibilities are \( n \in \{2, 4\} \). In these cases we see that \( 3m(P - \infty) \sim 0 \), which agrees with the conclusion of this proposition.

4.3. The hyperelliptic case. In this section we consider the case \( p = 2 \). Let \( \iota \) be the hyperelliptic involution.

**Theorem 4.18** [Poo01]. If \( q = 5 \), there are 18 torsion points on \( y^2 = x^q + 1 \). They come in the following families:

1. 1 point at infinity
2. 5 points of \( J[2] \). These are \((-\zeta_q^i, 0)\) for \( i \) in the range \( 0 \leq i \leq 4 \).
3. 2 points of \( J[1 - \zeta_5] \). These are \((0, \pm 1)\).
4. 10 points of \( J[(1 - \zeta_5)^3] \setminus J[(1 - \zeta_5)^2] \). These are of the form \((\zeta_5^i \sqrt{3}, \pm \sqrt{3})\) for \( i \) in the range \( 0 \leq i \leq 4 \).

**Lemma 4.19.** If \( q \geq 7 \), then any torsion point \( P \) of \( y^2 = x^q + 1 \) must satisfy either \( P = \iota P \) or \( P = \zeta_q P \).

**Proof.** From Proposition 4.17, we know that \( 2q(P - \infty) \sim 0 \). Write \( D = D_2 + D_q \) for a 2-torsion divisor \( D_2 \) and a \( q \)-torsion divisor \( D_q \). Note that \( D_2 \) is defined over \( \mathbb{Q}(\mu_q) \), so any element of \( \text{Gal}(\mathbb{Q}(\mu_q, D_q)/\mathbb{Q}(\mu_q)) \) automatically fixes \( D_2 \).

We break the rest of the proof into four steps.

1. In this step, we will show that \( 2(1 - \zeta_q)^3 P \sim 0 \).

   For contradiction, suppose that \( 2(1 - \zeta_q)^3 P \neq 0 \). An argument similar to the proof of Proposition 3.8 will show that \([L_{1,4} : L_{1,1}] \geq q^2\). Combining this with an argument similar to the proof of Lemma 4.11 now shows that there are at least \( 2q^3 \) distinct points of the form \( hzP \) for \( h \in \text{Gal}(\mathbb{Q}(J[2q])/\mathbb{Q}) \) and \( z \in \mathbb{Z} \). Pick \( q^3 \) of them and take care to never include both a point \( P_i \) and its hyperelliptic involute \( \iota(P_i) \); enumerate them \( P_1, \ldots, P_{q^3} \). For every tuple \( \mathbf{a} = (a_1, \ldots, a_{q^3}) \) of nonnegative integers satisfying \( a_1 + \ldots + a_{q^3} = (q - 1)/2 \), define
\[
P_{\mathbf{a}} := a_1 P_1 + \cdots + a_{q^3} P_{q^3}.
\]

   By construction, \( P_{\mathbf{a}} \) is an element of \( J[2q] \). Note that the number of such tuples \( \mathbf{a} \) is \( \binom{q^3 + (q - 1)/2 - 1}{(q - 1)/2} \) and the number of elements of \( J[2q] = (2q)^{q-1} \). We seek to show that
\[
\#\{P_{\mathbf{a}} : a_1 + \ldots + a_{q^3} = (q - 1)/2\} > \#J[2q].
\]

   To do this, we need to show the inequality \( \binom{q^3 + (q - 1)/2 - 1}{(q - 1)/2} > (2q)^{q-1} \). The numerator of the binomial coefficient is \((q^3)(q^3 + 1) \ldots (q^3 + (q - 1)/2 - 1)\), so it is bounded below by \((q^3)^{(q-1)/2}\). The denominator of the binomial coefficient is \((1)(2) \ldots ((q - 1)/2)\).

   Using the inequality \( k((q - 1)/2 - k) \leq ((q - 1)/4)^2 < (q/4)^2 \), we can bound this
In this step, we will show that we must have \( \gamma \) with an argument similar to the proof of Lemma 4.14.

For that would mean it fixes all of \( D \) and consider \( \gamma \) as a map to \( \mathbb{P}^1 \); that is, \( \varphi : \mathcal{C} \to \mathbb{P}^1 \).

The requirement that \( a_1 + \ldots + a_{q^3} = (q - 1)/2 \) implies that \( \deg \varphi \leq (q - 1)/2 \). The fact that all the \( P_i \) are distinct and that \( a \neq a' \) implies that \( \varphi \) is not the constant map.

We would like to apply the Castelnuovo-Severi inequality (Proposition 4.2) to \( \varphi \) and the canonical map \( \mathcal{C} \to \mathbb{P}^1 \). More precisely, if \( F \) is the function field of \( \mathcal{C} \), then the canonical map induces an embedding \( F_1 \subseteq F \) where \( F_1 \) is the function field of a genus 0 curve, and \( \varphi \) induces an embedding \( F_2 \subseteq F \) where \( F_2 \) is also the function field of a genus 0 curve. To apply the Castelnuovo-Severi inequality, we must check that the compositum of \( F_1 \) and \( F_2 \) in \( F \) equals \( F \). Since \( F_1 \) is of index two in \( F \), we know that \( F_1 F_2 \in \{ F_1, F \} \).

From the pigeonhole principle, we conclude that there must be two different tuples \( a \) and \( a' \) such that \( P_a = P_{a'} \). Choose a rational function \( \varphi \) such that \( \text{div} \varphi = P_a - P_{a'} \) and \( a \neq a' \).

From the previous step, we know that \( J^{(1 - \zeta_q)} \sim 0 \). Suppose for contradiction that \((1 - \zeta_q) D_q \neq \varnothing \). Then \( Q(\mu_q, D_q) \) contains at least \( L_{1,2} \). From Lemma 3.3 (3) we know that \( \text{Gal}(L_{1,2}/L_{1,1}) \) is cyclic of order \( q \); let \( \gamma \) be a generator. We will proceed with an argument similar to the proof of Lemma 4.14.

We will check that \( \gamma \) cannot fix \( P \). Since \( \gamma \) fixes \( D_2 \), we need to check that \( \gamma \) cannot fix \( D_q \). Either \( D_q \) or \((1 - \zeta_q) D_q \) lies in \( J^{(1 - \zeta_q)} \) \( \setminus J^{[1 - \zeta_q]} \), so by Lemma 2.16 we know that either \( L_{1,2} = L_{1,1}((1 - \zeta_q) D_q) \) or \( L_{1,2} = L_{1,1}((1 - \zeta_q) D_q) \). Hence \( \gamma \) cannot fix \( D_q \), for that would mean it fixes all of \( L_{1,2} \). Using Lemma 2.12, we know that for each \( \chi \), we must have

\[
\xi_\chi(\gamma) = 1 + \pi_q u_\chi
\]
for some \( u_\chi \in \mathcal{O}_q \). Now observe that
\[
2\xi_\chi(\gamma) - 2\xi_\chi(\gamma)^3 + \xi_\chi(\gamma)^4 = (2 + 2\pi_qu_\chi) \\
+ (-2 - 6\pi_qu_\chi - 6\pi_q^2u_\chi^2 - 2\pi_q^3u_\chi^3) \\
+ (1 + 4\pi_qu_\chi + 6\pi_q^2u_\chi^2 + 4\pi_q^3u_\chi^3 + \pi_qu_\chi^4) \\
= 1 + 2\pi_q^3u_\chi^3 + \pi_q^4u_\chi^4,
\]
so by another application of Lemma 2.12 we see that the endomorphism \( 2\gamma - 2\gamma^3 + \gamma^4 \) must fix \( D_q \). It also fixes \( D_2 \) since \( \gamma \) acts trivially on \( D_2 \), so we see that \( P \sim (2\gamma - 2\gamma^3 + \gamma^4)P \). Rewriting this as
\[
P + 2\gamma^3P \sim 2\gamma P + \gamma^4 P,
\]
we obtain a 3:1 map to \( \mathbf{P}^1 \) since by construction, \( \gamma \) has order \( q \) and \( \gamma \) does not fix \( P \). By the Castelnuovo-Severi inequality applied to this degree 3 map to \( \mathbf{P}^1 \) and the degree 2 canonical map to \( \mathbf{P}^1 \), we get
\[
g \leq (2 - 1)(3 - 1) = 2,
\]
contradicting the fact that \( g = (q - 1)/2 = (7 - 1)/2 \geq 3 \).

(3) In this step, we will show that either \( 2P \sim 0 \) or \( (1 - \zeta_q)P \sim 0 \).

Suppose that \( P \neq \zeta_qP \). We will show that \( 2P \sim 0 \).

From the previous step, we know that \( 2(1 - \zeta_q)P \sim 0 \), so \( 2P \sim 2\zeta_qP \). Since \( P \neq \zeta_qP \), we obtain from \( 2P \sim 2\zeta_qP \) a 2:1 map \( \varphi \) to \( \mathbf{P}^1 \). If this map did not factor through the 2:1 canonical map, the Castelnuovo-Severi inequality would imply that \( g \leq (2 - 1)(2 - 1) = 1 \), which is a contradiction. Hence this map must factor through the 2:1 canonical map, so as in step (1) we see that \( 2P \sim 0 \).

(4) To finish, the case that \( 2P \sim 0 \) would imply \( P = \zeta P \). The case that \( (1 - \zeta_q)P \sim 0 \) would imply \( P = \zeta_qP \).

\[\square\]

4.4. Some remaining curves. In the previous section we observed that \( C_{2,5} \) had some “unexpected” torsion points in \( J[(1 - \zeta_5)^3] \setminus J[(1 - \zeta_5)^2] \). Since \( C_{2,5} \) is an elliptic curve, all the torsion points will be on the curve.

In order to deal with \( C_{2,3} \) and \( C_{2,5} \) having “extra” torsion points, we will instead look at three covers of each curve. We will consider

(1) the covers \( C_{4,3}, C_{8,3}, \) and \( C_{2,9} \) of \( C_{2,3} \)
(2) the covers \( C_{4,5}, C_{2,15}, \) and \( C_{2,25} \) of \( C_{2,5} \)

The map from each cover \( C_{n,d} \) to \( C_{p,q} \) is \( \varphi_{n,d} : C_{n,d} \rightarrow C_{p,q} \) given by \((x, y) \mapsto (x^{d/q}, y^{n/p})\).

From Proposition 4.17, we have an upper bound \( N_{n,d} \) on the order of any torsion point on \( C_{n,d} \). If \( P \) is a point of \( C_{n,d} \) satisfying \( N_{n,d}(P - \infty) \sim 0 \), then \( N_{n,d}(\varphi_{n,d}(P) - \infty) \sim 0 \) as well, so \( P \in \varphi^{-1}_{n,d}(C_{p,q} \cap J_{p,q}[N_{n,d}]) \). Now we have reduced to a finite problem since \( J_{p,q}[N_{n,d}] \) has order \( N(\varphi^{-1}_{n,d}(C_{p,q} \cap J_{p,q}[N_{n,d}])) \) has order at most \( \langle n/p, d/q \rangle N_{n,d}(p-1)(q-1) \).

We finish the analysis of these curves with the aid of a computer. The first task is to determine the set \( \varphi^{-1}_{n,d}(C_{p,q} \cap J_{p,q}[N_{n,d}]) \) by writing down \( x \)- and \( y \)-coordinates of points up to precision \( 10^{50} \).
When \((p, q) = (2, 5)\) we already know (from Theorem 4.18) the set \(C_{p,q} \cap J_{p,q}[N_{n,d}]\), and it is simple to pull these back along \(\varphi_{n,d}\) to get a list of potential torsion points on \(C_{n,d}\). We enumerate these points \(P_1, P_2, \cdots\).

When \((p, q) = (2, 3)\), we know that \(C_{2,3}\) is an elliptic curve so we must determine \(C_{p,q}[N_{n,d}]\) and pull these back to points of \(C_{n,d}\). Let \(u_k\) be the \(k\)-division polynomial of \(C_{2,3}\); this polynomial has the property that its roots are the \(x\)-coordinates of the \(k\)-torsion points of \(C_{2,3}\). The coefficients of \(u_k\) are integers. Since \(N_{n,d}\) is even, we know furthermore that \(u_{N_{n,d}}\) will be of the form \(u_{N_{n,d}}(x) = x v_{N_{n,d}}(x^3)\) due to the action of \(\zeta_3\). The degree of \(v_{N_{n,d}}\) is \(N_{n,d}^2/6\). Let \(\{x_0, x_1, \cdots, x_{N_{n,d}^2/6}\}\) be the roots of \(x v_{N_{n,d}}(x)\). For each \(r\), there may be a torsion point on \(C_{n,d}\) whose \(x\)-coordinate is \((x_r)^{1/d}\). We can choose the \(d\)th root arbitrarily, because a choice of a different \(d\)th root does not affect whether or not the corresponding point is \(N_{n,d}\)-torsion (since they will all be in the same orbit of the \(Z\)-action). Letting \(P_r\) be any point on \(C_{n,d}\) whose \(x\)-coordinate is \((x_r)^{1/d}\), the divisor \((P_r - \infty)\) is potentially torsion.

For each \((n, d)\in \{(4, 3), (8, 3), (2, 9), (2, 15), (4, 5), (2, 25)\}\), we first determine the coordinates of each \(P_r\) up to accuracy \(10^{10}\). To test whether \(N_{n,d}(P_r - \infty)\) is principal, we use the magma package hcpersids. This package computes the periods and the Abel-Jacobi map for any superelliptic curve. Using this package, we compute the Abel-Jacobi images of \(N_{n,d}(P_r - \infty)\in J_{n,d}(C)\). By testing whether this image is sufficiently close to zero, we determine whether or not \(P_r\) has the potential to be a torsion point.

The results are as follows.

1. The only torsion points on \(C_{2,9}\) are the 12 superelliptic branch points \(\{(\zeta_3^i, 0) : 0 \leq i \leq 8\} \cup \{(0, \pm 1)\} \cup \{\infty\}\).
2. The only torsion points on \(C_{8,3}\) are the 12 superelliptic branch points \(\{(\zeta_3^i, 0) : 0 \leq i \leq 2\} \cup \{(0, \zeta_3^i) : 0 \leq i \leq 7\} \cup \{\infty\}\).
3. The only torsion points on \(C_{2,15}\) are the 18 superelliptic branch points \(\{(\zeta_5^i, 0) : 0 \leq i \leq 14\} \cup \{(0, \pm 1)\} \cup \{\infty\}\).
4. The only torsion points on \(C_{2,25}\) are the 28 superelliptic branch points \(\{(\zeta_5^i, 0) : 0 \leq i \leq 24\} \cup \{(0, \pm 1)\} \cup \{\infty\}\).
5. The only torsion points on \(C_{4,5}\) are the 10 superelliptic branch points \(\{(\zeta_5^i, 0) : 0 \leq i \leq 4\} \cup \{(0, \zeta_5^i) : 0 \leq i \leq 3\} \cup \{\infty\}\).
6. On \(C_{4,3}\), the program returned that the point \((2, \sqrt{3})\) had the potential to be a torsion point; its Abel-Jacobi image was quite close to zero. Further analysis using the IsPrincipal feature of magma led to the following:

\[
12(2, \sqrt{3}) = 12\infty = \text{div}((-4\sqrt{3}y + 12)x^2 + (18y^2 - 8\sqrt{3}y - 6)x + y^4 - 12\sqrt{3}y^3 + 18y^2 - 4\sqrt{3}y + 9)
\]

No other torsion points were found, so the complete list of the 20 torsion points of \(C_{4,3}\) is \(\{(2\zeta_3^i, \sqrt{3}\zeta_3^i) : 0 \leq i \leq 2, 0 \leq j \leq 3\} \cup \{(\zeta_3^i, 0) : 0 \leq i \leq 2\} \cup \{(0, \zeta_3^i) : 0 \leq j \leq 3\} \cup \{\infty\}\).

4.5. Main Theorem. In summary, we have the following result.

**Theorem 4.20.** Suppose \(n, d\) are coprime integers with \(n, d \geq 2\). The point at infinity of \(C_{n,d}\), and points of \(C_{n,d}\) whose \(x\)- or \(y\)-coordinate is zero are all torsion points. These are the only torsion points except in the following cases.
(1) \((n, d) \in \{(2, 3), (3, 2)\}\). Then \(C_{n,d}\) is an elliptic curve, so it has infinitely many torsion points.

(2) \((n, d) \in \{(2, 5), (5, 2)\}\). The only other torsion points on \(C_{2,5}\) are \(\{(\zeta_5^i \sqrt{4}, \pm \sqrt{5}) : 0 \leq i \leq 4\}\). The curves \(C_{2,5}\) and \(C_{5,2}\) are isomorphic via \((x, y) \in C_{2,5} \mapsto (\zeta_4 y, -x) \in C_{5,2}\), so torsion points on \(C_{5,2}\) are similar.

(3) \((n, d) \in \{(3, 4), (4, 3)\}\). The only other torsion points on \(C_{4,3}\) are \(\{(2\zeta_4^i, \pm \sqrt{3}) : 0 \leq i \leq 4\}\). As before, torsion points on the isomorphic curve \(C_{3,4}\) are similar.

**Proof.** Suppose \(n, d\) are both odd. Pick an odd prime \(p\) dividing \(n\) and an odd prime \(q\) dividing \(d\). Then \(p \neq q\) since \(n\) and \(d\) are coprime. There is a the map from \(y^n = x^d + 1\) to \(y^p = x^d + 1\) given by \((x, y) \mapsto (x^{d/q}, y^{n/p})\) that sends torsion points to torsion points. By our work in Section 4.4, we know that the only torsion points on the latter curve are those whose \(x\)- or \(y\)-coordinate is zero, and also the point at \(\infty\). The preimages of these points on the original curve are also points whose \(x\)-coordinate or \(y\)-coordinate is zero, and also the point at \(\infty\).

Without loss of generality, now suppose that \(n\) is even. If \(n\) had an odd prime factor, we can use the same argument in the first paragraph. So assume that \(n = 2^n\). If \(d\) has a prime factor \(q \geq 7\), then using Lemma 4.19 and the same argument in the first paragraph, we are done.

The final case is \(n = 2^n\) and \(d = 3^i 5^k\). If \(j \geq 2\), then \(C_{n,d}\) maps to \(C_{2^j, 5^k}\). From our work in Section 4.4, we know that the only torsion points of \(C_{2,5}\) are the superelliptic branch points, so the same argument as in the first paragraph works. The case \(k \geq 2\) is similar. So we may assume \(d \in \{3, 5, 15\}\).

If \(d = 3\), then if \(n \geq 8\) we can map \(C_{n,d} \to C_{8,3}\). The latter’s torsion points are only the superelliptic branch points by our work in Section 4.4, so we are done (by an argument similar to the first paragraph). If \(n \in \{2, 4\}\), we have analyzed these cases in Section 4.4.

If \(d = 5\), then if \(n \geq 4\) we can map \(C_{n,d} \to C_{4,5}\). The latter’s torsion points are only the superelliptic branch points by our work in Section 4.4, so we are done (by an argument similar to the first paragraph). We know the torsion points of \(C_{2,5}\) from Theorem 4.18.

If \(d = 15\), then we can always map \(C_{n,d} \to C_{2,15}\). The latter’s torsion points are only the superelliptic branch points by our work in Section 4.4, so we are done (by an argument similar to the first paragraph). \(\square\)

5. Torsion Points on a Generic Superelliptic Curve

The aim of this section is to prove the following result.

**Theorem 5.1.** Suppose \(n, d \geq 2\) are coprime and satisfy \(n + d \geq 7\). Let \(C\) be the curve defined by the equation

\[
y^n = \prod_{x=1}^{d} (x - a_i)
\]

over \(k := \mathbb{Q}(a_1, \ldots, a_d)\), and suppose \(C\) is embedded into its jacobian \(J\) using the unique point at infinity. The points \((a_i, 0)\) and \(\infty\) are torsion points.

(1) If \(d \geq 3\), there are no other torsion points.
(2) If $d = 2$ and $n \neq 5$, the only other torsion points are

$$\left\{ \left( \frac{a_1 + a_2}{2}, -\zeta_i \sqrt{\left( \frac{a_1 - a_2}{2} \right)^2} \right) : 0 \leq i \leq n - 1 \right\}.$$ 

(3) If $(n, d) = (5, 2)$, the only other torsion points are

$$\left\{ \left( \frac{a_1 + a_2}{2}, -\zeta_i \sqrt{\left( \frac{a_1 - a_2}{2} \right)^2} \right) : 0 \leq i \leq 4 \right\} \cup \left\{ \left( \pm \frac{(a_2 - a_1) \sqrt{5} + (a_1 + a_2)}{2}, \zeta_i \sqrt{(a_2 - a_1)^2} \right) : 0 \leq i \leq 4 \right\}.$$ 

This extends Theorem 7.1 of [PS14] from $n = 2$ to all $n$. To prove this result, we need a few more results about torsion points on certain curves.

5.1. The curves $y^n = x^d + x$.

**Proposition 5.2.** Suppose $n, d \geq 2$ are coprime, $P$ is a torsion point whose order divides $d$, and $P \neq \infty$. Then $d = 2$ or $(n, d) = (2, 3)$.

**Proof.** Suppose $(x_P, y_P)$ is such a point and $\text{div}(a(x, y) = d(x_P, y_P) - d\infty$ where $a(x, y)$ is a polynomial in $x$ and $y$. By considering the pole order at $\infty$, we conclude that $a(x, y)$ can be chosen to be of the form $a(x, y) = y - g(x)$ where $\text{deg} g < d/n$. Then

$$\text{div} \prod_{i=0}^{n-1} (\zeta_i^n y - g(x)) = d \sum_{i=0}^{n-1} (x_P, \zeta_i^n y_P) - nd\infty = \text{div}(x - x_P)^d,$$

from which we conclude that $\prod_{i=0}^{n-1} (\zeta_i^n y - g(x))$ and $(x - x_P)^d$ must be scalar multiples of each other. Using $y^n = x^d + x$, we rewrite

$$\prod_{i=0}^{n-1} (\zeta_i^n y - g(x)) = y^n - g(x)^n = x^d + x - g(x)^n.$$ 

Since $\text{deg} g < d/n$, comparing the coefficient of $x^d$ results in

$$x^d + x - g(x)^n = (x - x_P)^d.$$ 

Perform the change of variables $x' = x - x_P/2$ and define $h(x) := g(x + x_P/2)$. Then we get

$$h(x')^n = \left( x' + \frac{x_P}{2} \right)^d + x' + \frac{x_P}{2} - \left( x' - \frac{x_P}{2} \right)^d.$$ 

Therefore,

$$h(x)^n + (-1)^d h(-x')^n = x' \left( 1 - (-1)^d \right) + \frac{x_P}{2} \left( 1 + (-1)^d \right).$$

Now we break up into cases depending on the parity of $d$.

(1) $d$ is even. Then the equation \([3]\) becomes $h(x')^n + h(-x')^n = x_P$. This means that $\gcd(h(x'), h(-x'))^n$ divides $x_P$, so $h(x')$ and $h(-x')$ are forced to be coprime. Moreover, factoring gives $x_P = h(x')^n + h(-x')^n = \prod_{i=0}^{n-1} (h(x') + \zeta_i^n \cdot \zeta_n h(-x'))$. In particular, $h(x') + \zeta_{2n} h(-x')$ and $h(x') + \zeta_{2n} \cdot \zeta_n h(-x')$ are forced to be constants,
meaning that \( h(x') \) and \( h(-x') \) are constants as well. Hence \( g(x) \) is constant; this forces \( x^d + x - (x - x_P)^d \) to be a constant. If \( d \geq 3 \), then considering the coefficient of \( x^{d-1} \) forces \( x_P = 0 \), which is a contradiction since \( x^d + x - (x - 0)^d = x \) is nonconstant. Hence \( d = 2 \).

(2) \( d \) is odd. We proceed as in the previous case to obtain \( h(x')^n - h(-x')^n = 2x' \).

As before, this means that \( h(x') \) and \( h(-x') \) are coprime. Factoring gives \( 2x' = \prod_{i=0}^{n-1}(h(x') - \zeta_i h(-x')) \). If \( n \geq 3 \), then considering the degree of each factor shows that at least two of them must be constants, which will force \( h(x') \) and \( h(-x') \) to be constant, and we can repeat the same argument as before to conclude \( d = 2 \) (which is a contradiction since \( d \) is odd). We are left with \( n = 2 \) and \( 2x' = (h(x') + h(-x'))(h(x') - h(-x')) \). Since \( h(x') + h(-x') \) is an odd polynomial and \( h(x') - h(-x') \) is an even polynomial, we see that \( h(x') - h(-x') \) is a constant while \( h(x') + h(-x') \) is a multiple of \( x' \). This just means that \( \deg h = 1 \), which will also mean that \( \deg g = 1 \).

Writing \( g(x) = ax + b \) and recalling that \( n = 2 \), we obtain

\[
x^d + x - (x - x_P)^d = (ax + b)^2.
\]

Considering the coefficient of \( x^{d-1} \), we conclude that either \( x_P = 0 \) or \( d = 3 \). The former is impossible since it would force \( x = (ax + b)^2 \), so we conclude that \((n,d) = (2,3)\).

5.2. A couple curves for which \( n + d = 7 \). Using the same method as in Section 4.4, one can show the following.

**Proposition 5.3.**

(1) The only torsion points on \( y^3 = x^4 + x^2 + 1 \) whose order divides 12 are \( \infty \) and those where \( y = 0 \).

(2) The only torsion points on \( y^4 = x^3 + x + 1 \) whose order divides 12 are \( \infty \) and those where \( y = 0 \).

**Proof.** In the first case, the curve \( y^3 = x^4 + x^2 + 1 \) is a 2 : 1 cover of the elliptic curve \( y^3 = x^2 + x + 1 \). As in Section 4.4, we can find all points on this elliptic curve whose order divides 12, pull them back to \( y^3 = x^4 + x^2 + 1 \), and test if any of these points have order dividing 12: the only ones that appear are the \((1 - \zeta_3)\)-torsion points and \( \infty \).

Similarly, the curve \( y^4 = x^3 + x + 1 \) is a 2 : 1 cover of the elliptic curve \( y^2 = x^3 + x + 1 \) and the same technique works. \( \square \)

5.3. **Proof of Theorem 5.1.** If \( d = 2 \), then the curve \( y^n = (x - a_1)(x - a_2) \) is isomorphic over \( k \) to \( y^n = x^2 - 1 \) via the isomorphism

\[
(x, y) \in C_{n,2} \mapsto \left(\frac{(a_2 - a_1)x + (a_1 + a_2)}{2}, \sqrt{\frac{(a_2 - a_1)^2}{4} y}\right) \in C.
\]

Using our classification of torsion points on \( C_{n,2} \) in Theorem 4.20, we obtain parts (2) and (3) of Theorem 5.1.

Now suppose \( d \geq 3 \) and that \( P \) is a torsion point of \( C \). We need to show that the \( y \)-coordinate of \( P \) is zero. Without loss of generality, we can assume that \( n = p \) for some prime \( p \). Suppose that \( P \) is a torsion point of exact order \((1 - \zeta_p)^a m\), where \( m \) is coprime to \( p \). By exact order, we mean that \((1 - \zeta_p)^{a - 1} m' P \neq 0 \) for any divisor \( m' \) of \( m \).
Then \( J[(1 - \zeta_p)^a m] \) is a finite étale cover of \( \text{Spec} k \), so the image of \( P \in J[(1 - \zeta_p)^a m] \) under any specialization will also be a torsion point of exact order \((1 - \zeta_p)^a m\). Combining this observation with Theorem 4.20, we see that if \( n + d \geq 8 \), then either (i) \((a, m) = (0, 1)\), (ii) \((a, m) = (1, 1)\), or (iii) \( a = 0, m \mid d \). The last case is impossible by Proposition 5.2. Case (i) corresponds to \( \infty \) and case (ii) corresponds to the points where \( y = 0 \).

The case \( n + d = 7 \) remains. We have taken care of \((n, d) = (5, 2)\) in the case \( d = 2 \). The case \((n, d) = (2, 5)\) follows from Theorem 7.1 of [PS14]. For \((n, d) \in \{(3, 4), (4, 3)\}\), we combine Theorem 4.20(3), Proposition 5.3, and our observation that torsion points specialize to torsion points of the same order to finish.