ON $BT_1$ GROUP SCHEMES AND FERMAT JACOBIANS

RACHEL PRIES AND DOUGLAS ULMER

Abstract. Let $p$ be a prime number and let $k$ be an algebraically closed field of characteristic $p$. A $BT_1$ group scheme over $k$ is a finite commutative group scheme which arises as the kernel of $p$ on a $p$-divisible (Barsotti–Tate) group. Our main result is that every $BT_1$ scheme group over $k$ occurs as a direct factor of the $p$-torsion group scheme of the Jacobian of an explicit curve defined over $\mathbb{F}_p$. To prove this, we give a careful account of three classifications of $BT_1$ group schemes, due in large part to Kraft, Ekedahl, and Oort, and we apply these classifications to study the $p$-torsion group schemes of Jacobians of Fermat curves.

1. Introduction

Fix a prime number $p$ and let $k$ be an algebraically closed field of characteristic $p$. A $BT_1$ group scheme over $k$ is a finite commutative group scheme which is the kernel of $p$ on a $p$-divisible group. (The term $BT_1$ stands for Barsotti–Tate truncated at level 1, and Barsotti–Tate is a synonym for $p$-divisible.) These are the finite commutative group schemes killed by $p$ which also satisfy $\text{Ker} F = \text{Im} V$ and $\text{Ker} V = \text{Im} F$ where $F$ and $V$ are the Frobenius and Verschiebung maps respectively. The simplest $BT_1$ group schemes are $\mathbb{Z}/p\mathbb{Z}$ and $\mu_p$.

It is also of interest to consider polarized $BT_1$ group schemes over $k$, i.e., $BT_1$ group schemes $G$ with a pairing that induces a non-degenerate, alternating pairing on the Dieudonné module of $G$. (See [Oor01, §9] for the motivation behind this unusual definition.) If $A$ is a principally polarized abelian variety of dimension $g$ over $k$, its $p$-torsion subscheme $A[\mathbb{F}_p]$ is naturally a polarized $BT_1$ group scheme of order $p^{2g}$.

The isomorphism class of a polarized $BT_1$ group scheme of order $p^{2g}$ is uniquely determined by a combinatorial invariant called the Ekedahl–Oort type. There are $2^g$ possibilities for this invariant. If $G = A[p]$ for a principally polarized abelian variety $A$, then $G$ is not an isogeny invariant, in contrast with the Newton polygon of the $p$-divisible group of $A$.

If $C$ is a smooth irreducible projective curve of genus $g$ over $k$, then its Jacobian $\text{Jac}(C)$ is a principally polarized abelian variety of dimension $g$, and thus $G = \text{Jac}(C)[p]$ is a polarized $BT_1$ group scheme of rank $p^{2g}$. By a result of Oda [Oda69], the de Rham cohomology of $C$ over $k$ determines the isomorphism class of $G$ uniquely via its Dieudonné module.

In general, it is not known which polarized $BT_1$ group schemes occur for Jacobians of curves. In fact, there are very few examples of curves for which the Ekedahl–Oort type has been computed. In this paper, one of our main results is:

Date: October 30, 2020.

2010 Mathematics Subject Classification. Primary 11D41, 11G20, 14F40, 14H40, 14L15; Secondary 11G10, 14G17, 14K15, 14H10, 14L05.

Key words and phrases. Curve, finite field, Jacobian, abelian variety, Fermat curve, Frobenius, Verschiebung, group scheme, Ekedahl–Oort type, de Rham cohomology, Dieudonné module, $p$-divisible group.
Theorem (= Theorem 11.1).

1. Every $BT_1$ group scheme over $k$ appears as a direct factor of $\text{Jac}(C)[p]$ for a suitable (and explicit) curve $C$ defined over $\mathbb{F}_p$.

2. Every polarized $BT_1$ group scheme over $k$ appears as a direct factor (with pairing) of $\text{Jac}(C)[p]$ for a suitable (and explicit) curve $C$ defined over $\mathbb{F}_p$.

We remark that parts (1) and (2) of this theorem are essentially equivalent, but neither implication is immediate.

To prove the Theorem, we study Fermat curves. We find that in almost all cases the “explicit curve” of the theorem can be taken to be a Fermat curve (and even a Fermat quotient curve, see Section 6.2). Fermat curves are natural to consider here because their de Rham cohomology with its Frobenius and Verschiebung operators admits a simple and explicit description. The Newton polygons of Fermat curves were determined by Yui [Yui80, Thm. 4.2].

For a positive integer $d$ which is prime to $p$, let $F_d$ denote the Fermat curve of degree $d$ over $k$, namely the smooth projective curve with affine equation $X^d + Y^d = 1$. We denote by $C_d$ the quotient of $F_d$ that has affine equation $y^d = x(1 - x)$. In Theorem 11.2, we prove that any $BT_1$ group scheme $G$ over $k$ occurs as a direct factor of $\text{Jac}(C_d)[p]$ for infinitely many $d$, unless $p = 2$ and $G$ contains either $\mathbb{Z}/2\mathbb{Z}$ or $\mu_2$. If $G$ has rank $p^2g$, in most cases $d$ can be taken to be any multiple of $p^{2g} - 1$. We handle the remaining cases when $p = 2$ by a fiber product argument.

Here is a more detailed outline of the paper.

In Section 2, we provide background on group schemes, Dieudonné modules, and $p$-divisible groups. Section 3 contains an exposition of the classification of $BT_1$-group schemes, including canonical filtrations, canonical types, and alternative descriptions in terms of permutations, words, and generators and relations. In Section 4, we provide several refinements involving duality, polarizations, and the Ekedahl–Oort type. In Section 5, we study homomorphisms between $BT_1$ group schemes in order to define well-known invariants like the $p$-rank and $a$-number, and some newer invariants related to the $p$-torsion group scheme of a supersingular elliptic curve, which we call the $s_{1,1}$-multiplicity and $u_{1,1}$-number.

While most of the material in Sections 2 to 5 is known, it is important to cover for several significant reasons, and we take the opportunity to correct a few minor imprecisions in the literature. First, it is useful to compare two classifications of $BT_1$ group schemes, one due to Kraft (unpublished [Kra75], see also [Moo01]) and the other due to Ekedahl and Oort [Oor01] (see also [vdG99]). The Kraft classification uses words on a two-letter alphabet $\{f, v\}$; it interacts well with direct sums and identifies the indecomposable objects in the category. The Ekedahl–Oort classification uses the interplay between $F$ and $V$ to build the canonical filtration and it is well suited to moduli-theoretic questions. Second, we give a dictionary between the two classifications and a third, given in terms of a finite set $S$ partitioned into two subsets and a permutation of $S$. This third classification is not as well known and is particularly well suited to studying the $p$-torsion group schemes of Fermat curves. Third, the work of Oort uses covariant Dieudonné theory, but the contravariant theory is more convenient for studying Fermat curves.

We then turn to studying the Fermat curve $F_d$ and its quotient curve $C_d$, for all natural numbers $d$ relatively prime to $p$. In Sections 6 and 7 we determine the polarized $BT_1$ module of $C_d$ in terms of the permutation classification and the Ekedahl–Oort structure, see Theorems 6.5 and 7.3. In Section 8 we detour to record some combinatorial preliminaries, and we recover several known
results about Fermat curves, cf. [Yui80], [KW88], [Gon97], and [MS18]. We hope that seeing these results in a unified framework will be useful to the reader.

In Sections 9 and 10, we explicitly describe the Ekedahl–Oort structure and associated invariants when \(d = p^a - 1\) and \(d = p^a + 1\), for any natural number \(a\). We call \(d = p^a + 1\) the “Hermitian case”; it was previously treated in [PW15], but the results here are more complete. We call \(d = p^a - 1\) the “encompassing case”; it is in some sense the general case because for every \(d'\), the group \(\text{Jac}(C_{d'})[p]\) is a direct factor of \(\text{Jac}(C_d)[p]\) where \(d = p^a - 1\) for some \(a\).

Finally, in Section 11, we apply the encompassing case to deduce most cases of Theorem 11.1.

In Section 11.3 we develop an extra argument for the remaining cases when \(p = 2\), and in Section 11.4 we sketch another approach to the \(p = 2\) case which relies on [LMPT Cor. 4.7].

Acknowledgements. Author RP was partially supported by NSF grant DMS-1901819, and author DU was partially supported by Simons Foundation grants 359573 and 713699.

2. Groups and modules

In this section, we set notation and briefly review certain categories of group schemes and their Dieudonné modules. Readers familiar with the general theory are invited to skip this section.

We assume familiarity with Witt vectors and finite commutative group schemes over a field \(k\) (which as always is algebraically closed of characteristic \(p > 0\)).

2.1. Group schemes of \(p\)-power order and their Dieudonné modules. In this section, we briefly review contravariant Dieudonné theory for finite group schemes of \(p\)-power order over \(k\) and for \(p\)-divisible groups. Our general reference for the assertions in this section is [Fon77].

2.1.1. Witt and co-Witt vectors. Let \(W_n\) be the ring scheme over \(k\) of Witt vectors of length \(n + 1\), and \(W = \varprojlim W_n\). Write \(F : W_n \to W_n\) for the Witt vector Frobenius: \(F(X_0, \ldots, X_n) = (X_0^p, \ldots, X_n^p)\) and \(V : W_n \to W_n\) for the Verschiebung: \(V(X_0, \ldots, X_n) = (0, X_0, \ldots, X_{n-1})\). We use the same notations for the operators \(F\) and \(V\) induced on \(W\).

To avoid notational confusion below, we write \(\sigma\) for the absolute Frobenius of \(k\), and extend it to \(W(k)\) by \(\sigma(a_0, a_1, \ldots) = (a_0^p, a_1^p, \ldots)\). Define the Dieudonné ring \(\mathbb{D} = W(k)[F, V]\) as the \(W(k)\)-algebra generated by symbols \(F\) and \(V\) with relations

\[
FV = VF = p, \quad F\alpha = \sigma(\alpha)F, \quad \text{and} \quad \alpha V = V\sigma(\alpha) \quad \text{for} \quad \alpha \in W(k). \tag{2.1.1}
\]

Let \(\mathbb{D}_k = \mathbb{D}/p\mathbb{D} \cong k\{F, V\}\).

Following [Fon77 II.1.5], define the co-Witt vectors \(CW\) as a functor of \(k\)-algebras \(R\) as follows: consider an infinite tuple \(\alpha = (\ldots, a_{-2}, a_{-1}, a_0)\) with \(a_i \in R\); given \(\alpha\) and an integer \(r \geq 0\), let \(I_r\) be the ideal of \(R\) generated by \(a_i\) with \(i \leq -r\); then \(CW(R)\) is the set of such \(\alpha\) such that \(I_r\) is nilpotent for sufficiently large \(r\). The functor \(CW\) admits a structure of \(W(k)\)-module. Finally, define endomorphisms \(F\) and \(V\) of \(CW\) by setting

\[
F(\ldots, a_{-1}, a_0) = (\ldots, a_{-1}^p, a_0^p) \quad \text{and} \quad V(\ldots, a_{-2}, a_{-1}, a_0) = (\ldots, a_{-2}, a_{-1}).
\]

With these definitions, \(CW\) is a functor with values in the category of left \(\mathbb{D}\)-modules.
2.1.2. Dieudonné modules. Recall that \( \mathbb{D} = W(k) \{ F, V \} \). If \( G \) is a finite, commutative group scheme over \( k \) of \( p \)-power order, then the Dieudonné module of \( G \) is

\[
M(G) := \text{Hom}_{k - \text{Gr}}(G, CW)
\]

(homomorphisms of \( k \)-group schemes), and it has the structure of a left \( \mathbb{D} \)-module. Then \( G \to M(G) \) is a contravariant functor which induces an anti-equivalence between the category of finite group schemes of \( p \)-power order over \( k \) and the category of left \( \mathbb{D} \)-modules that are of finite length as \( W(k) \) modules \([\text{Fon}77, \text{III.1.4}]\). The quasi-inverse functor is given by \( M \to G \) where

\[
G(R) = \text{Hom}_{\mathbb{D}}(M, CW(R))
\]

Note that \( G \) is connected if and only if \( F \) is nilpotent on \( M(G) \), and \( G \) is unipotent if and only if \( V \) is nilpotent on \( M(G) \). Here are four examples:

1. for \( \mathbb{Z}/p\mathbb{Z} \), the constant group scheme of order \( p \), \( M(\mathbb{Z}/p\mathbb{Z}) \cong \mathbb{D}/(F-1, V); \)
2. for \( \mu_p \), the kernel of Frobenius on the multiplicative group, \( M(\mu_p) \cong \mathbb{D}/(F, V-1); \)
3. for \( \alpha_p \), the kernel of Frobenius on the additive group, \( M(\alpha_p) \cong \mathbb{D}/(F, V); \) and
4. for the kernel of \( p \) on a supersingular elliptic curve over \( k \), \( M(G_{1,1}) \cong \mathbb{D}/(F-V, p). \)

2.1.3. Duality. If \( G \) is a finite, commutative group scheme over \( k \), define its Cartier dual \( G^D \) as

\[
G^D := \text{Hom}_{k - \text{Gr}}(G, \mathbb{G}_m),
\]

where \( \mathbb{G}_m \) is the multiplicative group over \( k \). If \( M \) is a left \( \mathbb{D} \)-module, the dual module \( M^* \) is

\[
M^* := \text{Hom}(M(G), \lim_n W_n),
\]

where the \( \mathbb{D} \)-module structure on the right is defined by \( (Ff)(m) = \sigma(fVm) \) and \( (Vf)(m) = \sigma^{-1}(fFm) \). A basic result of Dieudonné theory \([\text{Fon}77, \text{III.5}]\) is that

\[
M(G^D) \cong M(G)^*.
\]

2.2. \( p \)-divisible groups. A \( p \)-divisible group over \( k \) is an inductive system

\[
G_\cdot : 0 \leftrightarrow G_1 \leftrightarrow G_2 \leftrightarrow \cdots
\]

of finite group schemes of \( p \)-power order such that for all \( m, n > 0 \), the morphism \( p^m : G_{m+n} \to G_m \) factors through \( G_n \subset G_{m+n} \) and there is an exact sequence

\[
0 \to G_m \to G_{m+n} \to G_n \to 0.
\]

It follows that there is a positive integer \( h \), the height of \( G \), such that the order of \( G_n \) is \( p^{nh} \).

If \( A \) is an abelian variety of dimension \( g \) over \( k \), then setting \( G_n = A[p^n] \) (the kernel of multiplication by \( p^n \) on \( A \)) yields a \( p \)-divisible group of height \( 2g \).

Applying Cartier duality to the exact sequences above shows that the dual groups \( (G_n)^D \) form another \( p \)-divisible group called the dual of \( G \), and denoted \( G^D \).

Define \( M(G_\cdot) \), the Dieudonné module of a \( p \)-divisible group \( G \), as

\[
M(G_\cdot) := \lim_n M(G_n) = \lim_n \text{Hom}_{k - \text{Gr}}(G_n, CW).
\]

If \( G \) has height \( h \), then the \( W(k) \)-module underlying \( M(G_\cdot) \) is free of rank \( h \).
2.3. \textbf{BT}_1 \textbf{group schemes and BT}_1 \textbf{modules}. By definition, a \textit{BT}_1 \textit{group scheme} over \( k \) is a finite commutative group scheme \( G \) that is killed by \( p \) and that has the properties

\[
\ker(F : G \to G) = \text{Im}(V : G \to G) \quad \text{and} \quad \text{Im}(F : G \to G) = \ker(V : G \to G).
\]

The notation \( \text{BT}_1 \) is an abbreviation of “Barsotti–Tate of level 1” reflecting the fact \cite[Prop. 1.7]{Ill85} that \( \text{BT}_1 \) group schemes are precisely those which occur as the kernel of \( p \) on a Barsotti–Tate (=\( p \)-divisible) group.

By definition, a \( \text{BT}_1 \) module over \( k \) is a \( \mathbb{D}_k \)-module \( M \) of finite dimension over \( k \) and that has the properties

\[
\ker(F : M \to M) = \text{Im}(V : M \to M) \quad \text{and} \quad \text{Im}(F : M \to M) = \ker(V : M \to M).
\]

(Oort also calls these \( DM_1 \) modules.) Clearly, a \( \mathbb{D}_k \)-module \( M \) is a \( \text{BT}_1 \) module if and only if it is the Dieudonné module of a \( \text{BT}_1 \) group scheme over \( k \).

The group schemes \( \mathbb{Z}/p\mathbb{Z}, \mu_p, \) and \( G_{1,1} \) are \( \text{BT}_1 \) group schemes. On the other hand, \( \alpha_p \) is not, since \( \ker F = M(\alpha_p) \neq 0 = \text{Im} V \).

A \( \text{BT}_1 \) group scheme \( G \) is self-dual if there exists an isomorphism \( G \cong G^D \). Similarly, a \( \text{BT}_1 \) module \( M \) is self-dual if \( M \cong M^* \). Clearly, \( G \) is self-dual if and only if \( M(G) \) is self-dual.

One may ask that a duality \( \phi : G \to G^D \) be skew, meaning that \( \phi^D : G \cong (G^D)^D \to G^D \) satisfies \( \phi^D = -\phi \). This is equivalent to any of the following three conditions:

- the bilinear pairing \( G \times G \to \mathbb{G}_m \) induced by \( \phi \) is skew-symmetric;
- the symmetry \( M(\phi) : M(G)^* \to M(G) \) is skew (meaning \( M(\phi)^* = -M(\phi) \));
- the induced bilinear pairing \( M(G)^* \times M(G)^* \to k \) is skew-symmetric.

Interestingly, when \( p = 2 \), there exist \( \text{BT}_1 \) group schemes \( G \) with an alternating pairing such that the induced pairing on the \( M(G)^* \) is skew-symmetric but not alternating. This is the reason for the exotic definition of a polarized \( \text{BT}_1 \) group scheme mentioned in the introduction.

3. Review of classifications of \( \text{BT}_1 \) group schemes

In this section, we review bijections between isomorphism classes of \( \text{BT}_1 \) modules over \( k \) and three other classes of objects of combinatorial nature. More precisely, following Kraft \cite{Kra75}, Ekedahl, and Oort \cite{Oor01}, we will construct a diagram:

\[
\begin{array}{cccccc}
\text{BT}_1 \text{ modules} & \xrightarrow{\sim} & \text{canonical types} \\
\sim & & \sim \\
\text{multisets of primitive cyclic words on } \{f, v\} & \xrightarrow{\sim} & \text{admissible permutations of } S = S_f \cup S_v \\
\sim & & \sim \\
\text{multisets of cyclic words on } \{f, v\} & \xrightarrow{\sim} & \text{permutations of } S = S_f \cup S_v.
\end{array}
\]

where \( \text{BT}_1 \) modules are as in Section 2.3 and all other terms appearing in the diagram will be defined in Sections 3.1 to 3.3.

The top square gives bijections on isomorphism classes. The top horizontal map is the Ekedahl–Oort classification of \( \text{BT}_1 \) modules and the upper left vertical map is the Kraft classification. The
data in the center right (admissible permutations) is a convenient repackaging of the other types of data. The bottom row is a relaxation of the middle row which is useful for studying $p$-torsion group schemes of Jacobians, especially Fermat Jacobians.

There is another classification of $BT_1$ modules, due to Moonen [Moo01], involving cosets of Weyl groups. We will not use Moonen’s classification, so we omit the details. In addition, van der Geer [vdG99, §6] reformulated the combinatorial data underlying canonical types in terms of Young diagrams and partitions. We will not need this either, so we omit any further discussion.

In Section 3.1 we discuss the elementary considerations leading to the lower square. We then proceed clockwise around the upper square: in Section 3.2 we describe the arrow from words to $BT_1$ modules; in Section 3.3 we construct canonical types associated to $BT_1$ modules; and in Section 3.4 we construct permutations arising from canonical types. This order of presentation is parallel to that of our analysis of the $p$-torsion subgroups of Fermat curves in Sections 6 and 7.

3.1. Words and permutations.

3.1.1. Words. Let $W$ be the monoid of words $w$ on the two-letter alphabet $\{f, v\}$, and write 1 for the empty word. By convention, the first (resp. last) letter of $w$ is its leftmost (resp. rightmost) letter. The complement $w^c$ of $w$ is the word obtained by exchanging $f$ and $v$ at every letter.

For a positive integer $\lambda$, write $W_\lambda$ for the words of length $\lambda$. Endow $W_\lambda$ with the lexicographic ordering with $f < v$. Thus words $w_1, w_2$ of the same length satisfy $w_1 < w_2$ if and only if there exist words $t, v, t''$ (possibly empty) such that $w_1 = tft'$ and $w_2 = tvt''$.

If $w \in W_\lambda$, we write $w = u_\lambda \cdots u_0$ where $u_i \in \{f, v\}$ for $0 \leq i \leq \lambda - 1$. Define an action of the group $\mathbb{Z}$ on $W$ by requiring that $1 \in \mathbb{Z}$ map $w = u_\lambda \cdots u_0$ to $u_0u_\lambda \cdots u_1$. If $w$ and $w'$ are in the same orbit of this action, we say $w'$ is a rotation of $w$. A cyclic word is the orbit of a word for the action of $\mathbb{Z}$. Thus two words represent the same cyclic word if and only if they are rotations of one another. Write $\overline{W}$ for the set of cyclic words. If $w \in W$, let $\overline{w} \in \overline{W}$ denote the corresponding cyclic word.

A word $w$ is primitive if $w$ is not of the from $(w')^e$ for some word $w'$ and some integer $e > 1$. If $w$ has length $\lambda > 0$, it is primitive if and only if the subgroup of $\mathbb{Z}$ fixing $w$ is exactly $\lambda \mathbb{Z}$. Write $W'$ for the set of primitive words. By definition, a cyclic word is primitive if and only if its representing words are primitive.

The three “decorations” of $W$ (length, cyclic, primitive) may be applied in any combination. So, for example, $\overline{W}_\lambda$ denotes the set of primitive cyclic words of length $\lambda$. The operation of complementation ($w \mapsto w^c$) respects length and primitivity and descends to cyclic words.

In Diagram (3.1), the middle (resp. lower) left entry denotes (isomorphism classes of) multisets of elements of $\overline{W}$ (resp. $\overline{W}$), where a multiset is a set of elements together with multiplicities.

Primitive cyclic words are a subset of all cyclic words: $\overline{W} \subset \overline{W}$. At the level of multisets, define a retraction by sending the class of a word $w = (w')^e$ with $w'$ primitive to the class of $w'$ with multiplicity $e$. The vertical arrow at the lower left of Diagram (3.1) is this retraction.

3.1.2. Permutations. Consider a finite set $S$ written as the disjoint union $S = S_f \cup S_v$ of two subsets. (We call this a partition of $S$.) Let $\pi : S \to S$ be a permutation of $S$. Two such collections of data $(S = S_f \cup S_v, \pi)$ and $(S' = S'_f \cup S'_v, \pi')$ are isomorphic if there is a bijection $\iota : S \to S'$ such that $\iota(S_f) = S'_f$, $\iota(S_v) = S'_v$, and $\iota \pi = \pi' \iota$. 

Given \((S = S_f \cup S_v, \pi)\), there is an associated multiset of cyclic words on \(\{f, v\}\) defined as follows. For \(a \in S\) with orbit of size \(\lambda\), define the word \(w_a = u_{\lambda - 1} \cdots u_0\) where
\[
u_j = \begin{cases} f & \text{if } \pi^j(a) \in S_f, \\ v & \text{if } \pi^j(a) \in S_v. \end{cases}
\]
Then \(\overline{w_a}\) depends only on the orbit of \(a\). This gives a well-defined map from orbits of \(\pi\) to cyclic words, i.e., elements of \(\overline{W}\). Taking the union over orbits, we can associate to \((S = S_f \cup S_v, \pi)\) a multiset of cyclic words. If \(S\) and \(S'\) are isomorphic, then they yield the same multiset. This construction gives one direction of the horizontal arrow at the bottom of Diagram (3.1).

Given a multiset of cyclic words, let \(S\) be the set of all words representing the given cyclic words (repeated to account for multiplicities), let \(S_f\) be the subset of those words ending with \(f\), let \(S_v\) be the subset of those words ending with \(v\), and let \(\pi\) be defined by the action of \(1 \in \mathbb{Z}\) as in Section 3.1. This defines the other direction of the horizontal arrow at the bottom of Diagram (3.1).

For example, let \(S = \{1, \ldots, 9\}, S_f = \{2, 3, 5, 6, 9\}, S_v = \{1, 4, 7, 8\}\). Let \(\pi\) be the permutation \((135)(246)(789)\). The orbit through 1 and the orbit through 2 each gives rise to the cyclic word \(\overline{ffv}\), and the orbit through 7 gives rise to the cyclic word \(\overline{fvv}\). The associated multiset is \(\{(\overline{ffv})^2, \overline{fvv}\}\) (where \((\overline{ffv})^2\) means the cyclic word \(\overline{ffv}\) taken with multiplicity 2).

3.1.3. Primitive and admissible. Recall that a word \(w\) is primitive if it is not of the form \((w')^e\) for \(e > 1\). We say that the data \((S = S_f \cup S_v, \pi)\) is admissible if the associated words \(w_a\) for \(a \in S\) are primitive. Clearly, the horizontal arrow at the bottom of Diagram (3.1) induces a bijection between multisets of primitive cyclic words and admissible permutations of \(S = S_f \cup S_v\).

We leave it to the reader to make explicit the retraction from permutations to admissible permutations that makes the lower square of Diagram (3.1) commute. The following example provides a hint: If \(S_f = \{1, 3\}, S_v = \{2, 4\}, \) and \(\pi = (1234)\), then one choice for the corresponding admissible permutation is \((12)(34)\).

3.2. Cyclic words to \(BT_1\) modules. Following Kraft [Kra75], we attach a \(BT_1\) module to a multiset of primitive cyclic words. This defines the vertical arrow at the upper left of Diagram (3.1).

3.2.1. Construction. Suppose that \(w \in \mathcal{W}\) is a primitive word, say \(w = u_{\lambda - 1} \cdots u_0\) with \(u_j \in \{f, v\}\). Let \(M(w)\) be the \(k\)-vector space with basis \(e_j\) with \(j \in \mathbb{Z}/\lambda\mathbb{Z}\) and define a \(p\)-linear map \(F : M(w) \to M(w)\) and a \(p^{-1}\)-linear map \(V : M(w) \to M(w)\) by setting
\[
F(e_j) = \begin{cases} e_{j+1} & \text{if } u_j = f, \\ 0 & \text{if } u_j = v, \end{cases} \quad \text{and} \quad V(e_j) = \begin{cases} e_j & \text{if } u_j = v, \\ 0 & \text{if } u_j = f. \end{cases}
\]
(Note that \(F(e_j) \neq 0\) exactly when the last letter of the \(j\)-th rotation of \(w\) is \(f\), and \(V(e_{j+1}) \neq 0\) exactly when the last letter of the \(j\)-th rotation of \(w\) is \(v\).) This construction yields a \(BT_1\) module which up to isomorphism only depends on the primitive cyclic word \(\overline{w}\) associated to \(w\).

Kraft proves that \(M(w)\) is indecomposable and that every indecomposable \(BT_1\) module is isomorphic to one of the form \(M(w)\) for a unique primitive cyclic word \(\overline{w}\). Thus every \(BT_1\) module \(M\) is isomorphic to a direct sum \(\bigoplus M(w_i)\) where \(\overline{w_i}\) runs through a uniquely determined multiset of primitive cyclic words.

If \(w\) is a word that is not necessarily primitive, the formulas above define a \(BT_1\) module. If \(w = (w')^e\), Kraft also proves that \(M(w) \cong M(w')^e\).
It is clear that $M(f) = M(\mathbb{Z}/p\mathbb{Z})$, $M(v) = M(\mu_p)$, and $M(fv)$ is the Dieudonné module of the kernel of $p$ on a supersingular elliptic curve. More generally, if $w$ has length $> 1$ and is primitive, then $M(w)$ is the Dieudonné module of a unipotent, connected $BT_1$ group scheme.

3.2.2. Generators and relations. Let $w$ be a primitive word with associated $BT_1$ module $M(w)$. It will be convenient to have a presentation of $M(w)$ by generators and relations. Clearly, $M(f) = \mathbb{D}_k/(F - 1, V)$ and $M(v) = \mathbb{D}_k/(F, V - 1)$.

Now suppose $w$ has length $> 1$. Then, after rotating $w$ if necessary, we may assume its last letter is $f$ and its first letter is $v$. (Both letters appear because $w$ is primitive, so in particular is not $f^m$ nor $v^n$.) We then write $w$ in exponential notation as

$$w = v^{n_1} f^{m_1} \cdots v^{n_\ell} f^{m_\ell},$$

for some positive integers $\ell, m_1, \ldots, m_\ell, n_1, \ldots, n_\ell$.

3.2.3. Lemma. The $BT_1$ module $M(w)$ admits generators $E_i$ for $i \in \mathbb{Z}/\ell\mathbb{Z}$ with relations $F^{m_i} E_{i-1} = V^{n_i} E_i$ for $i = 1, \ldots, \ell$.

Proof. Indeed, for $i = 0, \ldots, \ell - 1$, let $I(i) = \sum_{j=1}^i (m_j + n_j)$, let $I'(i) = I(i) + m_{i+1}$, and let $E_i = e_{I(i)}$. The $E_i$ generate $M(w)$ as a $\mathbb{D}_k$-module because

if $I(i) \leq j \leq I'(i)$, then $e_j = F^{j-I(i)} E_i$,

and if $I'(i) \leq j \leq I(i + 1)$, then $e_j = V^{j-I'(i)} E_{i+1}$,

and there are relations $F^{m_i} E_{i-1} = e_{I'(i-1)} = V^{n_i} E_i$. □

The following diagram illustrates this presentation of $M(w)$:

$$\begin{array}{cccc}
E_{\ell-1} = e_{I(\ell-1)} & E_0 = e_{I(0)} & E_1 = e_{I(1)} & \cdots \\
F^{m_\ell} & F^{m_1} & F^{m_2} & \\
V^{n_\ell} & V^{n_1} & V^{n_2} & \\
e_{I'(\ell-1)} & e_{I'(0)} & e_{I'(1)} & \\
\end{array}$$

3.3. $BT_1$ modules to canonical types. Following Oort [Oor01], in this section we explain how to describe the isomorphism class of a $BT_1$ module in terms of certain combinatorial data. This defines the top horizontal arrow in Diagram (3.1). Readers are invited to work through the example in Section 3.3.9 while reading this section.

Warning: Many of our formulas differ from those in [Oor01] for two reasons: first, we use the contravariant Dieudonné theory, whereas Oort uses the covariant theory; second, Oort studies a filtration defined by $F^{-1}$ and $V$, whereas we use $F$ and $V^{-1}$. The two approaches are equivalent (and exchanged under duality), but the latter is more convenient for studying Fermat curves.

3.3.1. The canonical filtration. Recall that $W$ denotes the monoid of words on $\{f, v\}$. Let $M$ be a $BT_1$ module, and define a left action of $W$ on the set of $k$-subspaces of $M$ by requiring that

$$fN := F(N) \text{ and } vN := V^{-1}(N).$$

In other words, $f$ sends a subspace $N$ to its image under $F$ and $v$ sends $N$ to its inverse image under $V$. Note that if $N_1 \subset N_2$, then $fN_1 \subset fN_2$ and $vN_1 \subset vN_2$. If $N$ is a $\mathbb{D}_k$-module, so are $fN$ and $vN$, and $fN \subset N \subset vN$. Also note that $fM = \text{Im } F = \text{Ker } V = v0$. 

Let $M$ be a $BT_1$ module. An admissible filtration on $M$ is a filtration by $\mathbb{D}^k$-modules
\[ 0 = M_0 \subsetneq M_1 \subsetneq \cdots \subsetneq M_s = M, \]
(3.3.1)
such that for all $i$, there exist indices $\phi(i)$ and $\nu(i)$ such that $fM_i = M_{\phi(i)}$ and $vM_i = M_{\nu(i)}$.

3.3.2. Definition. The canonical filtration on $M$ is the coarsest admissible filtration on $M$. Letting Equation 3.3.1 denote the canonical filtration, the blocks of $M$ are $B_i = M_{i+1}/M_i$ for $0 \leq i \leq s-1$.

The next lemma verifies that there is a coarsest admissible filtration on $M$.

3.3.3. Lemma. Suppose $w_1, w_2, t, t', t'' \in W$.

1. If $w_1 = w_2 t$, then $w_1 M \subset w_2 M$.
2. If $w_1 = t f t'$ and $w_2 = t v t''$, then $w_1 M \subset w_2 0 \subset w_2 M$.
3. The collection of all subspaces $wM$ as $w$ ranges over $W$ is totally ordered (by containment), and thus finite.

Proof. Part (1) is true since $w_1 M = w_2 t M \subset w_2 M$. Part (2) is true since $w_1 M = t f t' M \subset t f M = t v 0 \subset t v t'' 0 = w_2 0 \subset w_2 M$.

Given any two words $w$ and $w'$, parts (1) and (2) imply that either $wM \subset w'M$ or $w'M \subset wM$, so the set of subspaces of the form $wM$ is totally ordered. Since $M$ is of finite dimension over $k$, there are only finitely many distinct subspaces of the form $wM$.

The canonical filtration of $M$ is constructed by enumerating all subspaces of $M$ of the form $wM$ and indexing them in order of containment. Define $s$ to be the number of steps in the filtration and $r$ to be the integer such that $M_r = fM = v 0$. Define functions
\[ \phi : \{0, \ldots, s\} \to \{0, \ldots, r\}, \quad \nu : \{0, \ldots, s\} \to \{r, \ldots, s\}, \quad \rho : \{0, \ldots, s\} \to \mathbb{Z} \]
by $fM_i = M_{\rho(i)}$, $vM_i = M_{\nu(i)}$, and $\rho(i) = \dim_k M_i$. This data has the following properties.

3.3.4. Proposition-Definition. The data $(r, s, \phi, \nu, \rho)$ associated to the canonical filtration of $M$ is a canonical type, i.e., $s > 0$, $0 \leq r \leq s$, and the functions $\phi$, $\nu$, and $\rho$ have the following properties:

1. $\phi$ and $\nu$ are monotone nondecreasing and surjective;
2. $\rho$ is strictly increasing with $\rho(0) = 0$;
3. $\nu(i+1) > \nu(i)$ if and only if $\phi(i+1) = \phi(i)$;
4. if the equivalent conditions in (3) are true, then $\rho(i+1) - \rho(i) = \rho(\nu(i) + 1) - \rho(\nu(i))$,
   while if not, then $\rho(i+1) - \rho(i) = \rho(\phi(i) + 1) - \rho(\phi(i))$;
5. and every integer in $\{1, \ldots, s\}$ can be obtained by repeatedly applying $\phi$ and $\nu$ to $s$.

If the data $(r, s, \phi, \nu, \rho)$ comes from the canonical filtration of a $BT_1$ module, then it is clear from the definitions that $0 \leq r \leq s$, with $s > 0$, and that properties (1), (2), and (5) hold. Oort proves that properties (3) and (4) hold in [Oor01] §2.

The properties imply that $\nu(i+1) - \nu(i)$ and $\phi(i+1) - \phi(i)$ are either 0 or 1, that exactly one of them is 1, and that $\nu(i) + \phi(i) = r + i$. The next lemma is proven in [Oor01] Lemma 2.4 and will be used in later sections.

3.3.5. Lemma. Let Equation 3.3.1 denote the canonical filtration of $M$ and let $B_i = M_{i+1}/M_i$ for $0 \leq i \leq s - 1$. If $\phi(i+1) > \phi(i)$ then $F$ induces a $p$-linear isomorphism $B_i \to B_{\phi(i)}$, and if $\nu(i+1) > \nu(i)$, then $V^{-1}$ induces a $p$-linear isomorphism $B_i \to B_{\nu(i)}$. 
The key assertion is that the canonical type of $M$ determines $M$ up to isomorphism:

3.3.6. **Proposition.** If the canonical types of two $BT_1$ modules $M, M'$ are equal, then $M \cong M'$.

Oort proves a related result \cite{Oor01}*{Thm. 9.4} involving quasi-polarizations (pairings) which is more involved and only applies to self-dual $BT_1$ modules. Moonen proves the result stated here \cite{Moo01}*{§4} in the more general context where the module $M$ also has endomorphisms by a semi-simple $\mathbb{F}_p$-algebra $D$; taking $D = \mathbb{F}_p$ yields Proposition 3.3.6.

3.3.7. **Remark.** It is often more convenient to replace $\rho$ with the function $\mu : \{0, \ldots, s - 1\} \to \mathbb{Z}$ defined by $\mu(i) = \rho(i + 1) - \rho(i)$. Property (2) says $\mu$ takes positive values, and property (4) says if $\nu(i + 1) > \nu(i)$, then $\mu(i) = \mu(\nu(i))$ and if $\phi(i + 1) > \phi(i)$, then $\mu(i) = \mu(\phi(i))$.

3.3.8. **Remark.** Oort defines a canonical type to be data as above satisfying properties (1)–(4), i.e., he omits (5), and he states \cite{Oor01}*{Remark 2.8} that every canonical type comes from a $BT_1$ module. With this definition, it is true that every canonical type comes from an admissible filtration on a $BT_1$ module, but not necessarily from the canonical filtration. Here is a counterexample: Let $r = 0$, $s = 2$ and $\phi(i) = 0$, $\nu(i) = i$ and $\rho(i) = i$ for $i = 0, 1, 2$. This data comes from a filtration on the $BT_1$ module $N = M((\mu_2)^2)$ of rank 2 with $F = 0$ and $V$ bijective and it satisfies properties (1)–(4), but not (5). The canonical filtration of $N$ yields the canonical type with $r = 0$, $s = 1$, $\phi(i) = 0$, $\nu(i) = i$ for $i = 0, 1$, and $\rho(0) = 0$, $\rho(1) = 2$.

3.3.9. **An example.** Let $M$ be the $k$-vector space with basis $e_1, \ldots, e_7$ and action of $F, V$ given by

| $e$ | $e_1$ | $e_2$ | $e_3$ | $e_4$ | $e_5$ | $e_6$ | $e_7$ |
|-----|-------|-------|-------|-------|-------|-------|-------|
| $F(e)$ | 0 | 0 | 0 | $e_1$ | $e_2$ | 0 | $e_3$ |
| $V(e)$ | 0 | 0 | 0 | $e_1$ | $e_2$ | $e_3$ | $e_6$ |

Using $\langle \ldots \rangle$ to denote the span of a set of vectors, the canonical filtration of $M$ is

$$M_0 = 0 \subset M_1 = \langle e_1, e_2 \rangle \subset M_2 = \langle e_1, e_2, e_3 \rangle \subset M_3 = \langle e_1, e_2, e_3, e_4, e_5 \rangle \subset M_4 = \langle e_1, e_2, e_3, e_4, e_5, e_6 \rangle \subset M_5 = \langle e_1, e_2, e_3, e_4, e_5, e_6, e_7 \rangle = M.$$ 

The canonical type is given by $s = 5$, $r = 2$, and the functions $\phi, \nu, \rho$ below:

| $i$ | 0 | 1 | 2 | 3 | 4 | 5 |
|-----|---|---|---|---|---|---|
| $\phi(i)$ | 0 | 0 | 0 | 1 | 1 | 2 |
| $\nu(i)$ | 2 | 3 | 4 | 5 | 5 |
| $\rho(i)$ | 0 | 2 | 3 | 5 | 6 | 7 |

3.4. **Canonical types to permutations.** In this section, again following \cite{Oor01}, we explain how to use a canonical type to define a partitioned set $S = S_f \cup S_\pi$, with permutation $\pi : S \to S$, thus defining the vertical arrow at the upper right of Diagram (3.1). These results cast considerable light on the structure of $BT_1$ modules, but they will not be used explicitly in the rest of the paper, so readers are invited to skip this section on a first reading.

Let $(r, s, \phi, \nu, \rho)$ be as in Definition 3.3.4 and let $\Gamma = \{0, \ldots, s - 1\}$. Define $\Pi : \Gamma \to \Gamma$ by:

$$\Pi(i) = \begin{cases} 
\phi(i) & \text{if } \phi(i + 1) > \phi(i), \\
\nu(i) & \text{if } \nu(i + 1) > \nu(i).
\end{cases}$$
Property (3) of Definition 3.3.4 shows that \( \Pi \) is well defined. The monotonicity of \( \nu \) and \( \phi \) implies that \( \Pi \) is injective and thus bijective. Letting \( \mu(i) := \rho(i + 1) - \rho(i) \), property (4) shows that \( \mu \) is constant on the orbits of \( \Pi \).

We partition \( \Gamma \) as a disjoint union \( \Gamma_f \cup \Gamma_v \) where \( i \in \Gamma_f \) if and only if \( \pi(i) = \phi(i) \). Equivalently,

\[
\Gamma_f = \{ i \in \Gamma \mid \phi(i + 1) > \phi(i) \} \quad \text{and} \quad \Gamma_v = \{ i \in \Gamma \mid \nu(i + 1) > \nu(i) \}.
\]

Then \( (r, s, \phi, \nu, \rho) \) is determined by the data \( \Gamma = \Gamma_f \cup \Gamma_v, \Pi : \Gamma \to \Gamma, \) and \( \mu : \Gamma \to \mathbb{Z} \).

3.4.1. Example 3.3.9 continued. In this case, the permutation \( \Pi \) of \( \Gamma = \{0, 1, 2, 3, 4\} \) is \((0, 2)(1, 3, 4)\), the partition is given by \( \Gamma_f = \{2, 4\} \) and \( \Gamma_v = \{0, 1, 3\} \), and the associated words are

\[
w_0 = fv, \ w_1 = fvv, \ w_2 = vf, \ w_3 = vfv, \ \text{and} \ w_4 = vvf.
\]

Note that \( \mu(0) = \mu(2) = 2 \) and \( \mu(1) = \mu(3) = \mu(4) \), so \( \mu \) is constant on the orbits of \( \Pi \).

To complete the definition of the vertical arrow at the top right of Diagram (3.1), we use \( \mu \) as a set of “multiplicities” to expand \( \Gamma \) into \( S \). More precisely, define

\[
S := \{ e_{i,j} \mid i \in \Gamma, 1 \leq j \leq \mu(i) \}
\]

with partition

\[
S_f = \{ e_{i,j} \in S \mid i \in \Gamma_f \} \quad \text{and} \quad S_v = \{ e_{i,j} \in S \mid i \in \Gamma_v \}
\]

and permutation

\[
\pi : S \to S \quad \text{with} \quad \pi(e_{i,j}) := e_{\Pi(i),j}.
\]

3.4.2. Lemma. The data \( (S = S_f \cup S_v, \pi) \) is an admissible permutation.

Proof. We need to show that the cyclic words associated to \( (S = S_f \cup S_v, \pi) \) are primitive. These are precisely the cyclic words associated to \( (\Gamma = \Gamma_f \cup \Gamma_v, \Pi) \) with multiplicities given by \( \mu \), so it will suffice to check that the cyclic words associated to \( (\Gamma = \Gamma_f \cup \Gamma_v, \Pi) \) are primitive. If \( w_i \) is the word attached to \( i \in \Gamma \), then \( w_{\Pi(i)} \) is the \( j \)-th cyclic rotation of \( w_i \). Thus to show that the \( w_i \) are all primitive, it suffices to show that they are distinct.

To that end, define a left action of the monoid \( \mathcal{W} \) on the set \( \{0, \ldots, s\} \) by requiring that \( f(i) = \phi(i) \) and \( \nu(i) = \nu(i) \). If \( i \in \Gamma \), and if \( w_i \) is the word associated to \( i \), then \( w_i \) fixes \( i \) and \( i + 1 \). This is a manifestation in the canonical type of the isomorphisms from Lemma 3.3.5

\[
B_i \to B_{\Pi(i)} \to B_{\Pi^2(i)} \to \cdots \to B_i.
\]

Now assume that \( i, j \in \Gamma, i < j, \) and \( w_i = w_j = w \). We will deduce a contradiction of property (5) in Definition 3.3.4. Since \( \phi \) and \( \nu \) are nondecreasing, for all \( n > 0 \) we have

\[
w^n(s) \geq i + 1 > i \geq w^n(0) \quad \text{and} \quad w^n(s) \geq j + 1 > j \geq w^n(0),
\]

so \( w^n(s) \geq j + 1 \) and \( i \geq w^n(0) \) for all \( n > 0 \).

Choose some \( i' \) with \( i < i' \leq j \). By property (5) of Definition 3.3.4, there is a word \( w' \) with \( w'(s) = i' \). Choose \( n > 0 \) large enough that \( w^n \) is at least as long as \( w' \), and then replace \( w' \) with \( w'v^m \) where \( m \) is chosen so \( w^n \) and \( w' \) have the same length. Since \( \nu(s) = \nu(s) = s \), we still have \( w'(s) = i' \). An argument parallel to that of Lemma 3.3.3 part (2) shows that if \( w' \geq w^n \) (in the lexicographic order defined in Section 3.1.1), then \( i' = w'(s) \geq w^n(s) \geq j + 1 \), a contradiction; and if \( w' < w^n \), then \( i' = w'(s) \leq w^n(0) \leq i \), again a contradiction. We conclude that there can be no \( i < j \) with \( w_i = w_j \), and thus \( \pi \) is admissible.
3.4.3. Remark. A more thorough analysis along these lines shows that if \( M \) is a \( BT_1 \) module, then there are finitely many primitive words \( w_i \) such that \( w_i^n M \supseteq w_i^{n+1} M \) for all \( n > 1 \). Enumerate these as \( w_0, \ldots, w_{s-1} \) and choose integers \( n_i \) so that \( w_i^n M = w_i^{n+1} M \) and \( w_i^{n+1} M \) for all \( n \geq n_i \), and so that the lengths of the \( w_i^n \) are all the same. Define \( \bar{w}_i = w_i^{n_i} \). Reorder the \( w_i \) so that
\[
\bar{w}_0 < \bar{w}_1 < \cdots < \bar{w}_{s-1}.
\]
(Numbering the \( \bar{w}_i \) from \( i = 0 \) turns out to be most convenient; see the proof of Lemma \[8.5.2\].) (Let \( \bar{w}_{-1} M = 0 \).) Then the \( w_i \) are distinct, and if there are repetitions, one should add multiplicities.) Now relabel the distinct \( \tau \) function such that \( \mu(i) = \dim_k (\bar{w}_i M) / (\bar{w}_{i-1} M) \).

3.5. Words to canonical types. In this section, we give a detailed description of the map from multisets of (not necessarily primitive) cyclic words to canonical types, i.e., from the lower left to the upper right of Diagram (3.1). This will be used in Section \[7\].

For \( 1 \leq i \leq n \), let \( \bar{w}_i \) be cyclic words with multiplicities \( m_i \). Our goal is to describe the canonical type of the \( BT_1 \) module
\[
M = \bigoplus_{i=1}^n M(\bar{w}_i)^{m_i},
\]
where \( M(\bar{w}_i) \) is the Kraft module discussed in Section \[3.2\].

Let \( \lambda_i \) be the length of \( \bar{w}_i \) and choose a representative \( w_i = u_{i,\lambda_i-1} \cdots u_{i,0} \) of \( \bar{w}_i \) with \( u_{i,j} \in \{f,v\} \). The \( k \)-vector space underlying \( M \) has basis \( e_{i,j,k} \) where \( 1 \leq i \leq n \), \( j \in \mathbb{Z}/\lambda_i \mathbb{Z} \), and \( 1 \leq k \leq m_i \). Its \( \mathbb{D}_k \)-module structure is given by
\[
F(e_{i,j,k}) = \begin{cases} e_{i,j+1,k} & \text{if } u_{i,j} = f, \\ 0 & \text{if } u_{i,j} = v, \end{cases}
\]
and
\[
V(e_{i,j+1,k}) = \begin{cases} e_{i,j,k} & \text{if } u_{i,j} = v, \\ 0 & \text{if } u_{i,j} = f. \end{cases}
\]

Let \( w_{i,j} = u_{i,j-1} \cdots u_{i,0} u_{i,\lambda_i} \cdots u_{i,j} \) be the \( j \)-th rotation of \( w_i \). Choose integers \( n_i \) so that \( n_i \lambda_i = n_{i'} \lambda_{i'} \) for all \( 1 \leq i, i' \leq n \), and let \( \bar{w}_{i,j} = w_{i,j}^{n_i} \).

Let \( T \) be the multiset obtained by including each \( \bar{w}_{i,j} \) with multiplicity \( m_i \). (The \( \bar{w}_{i,j} \) need not be distinct, and if there are repetitions, one should add multiplicities.) Now relabel the distinct elements of \( T \) as \( \omega_{\ell} \) for \( 0 \leq \ell \leq s-1 \) and ordered so that \( \omega_0 < \omega_1 < \cdots < \omega_{s-1} \). Let \( \tau \) be the function such that \( \tau(i,j) = \ell \) if and only if \( \bar{w}_{i,j} = \omega_{\ell} \). Let \( \mu(\ell) \) be the multiplicity of \( \omega_{\ell} \) in \( T \).

For \( 1 \leq \ell \leq s \), let \( M_{\ell} \) be the \( k \)-subspace of \( M \) spanned by those \( e_{i,j,k} \) with \( \tau(i,j) \leq \ell - 1 \). We claim that the canonical filtration of \( M \) is
\[
0 = M_0 \subsetneq M_1 \subsetneq \cdots \subsetneq M_s = M.
\]
Indeed, it is easy to see that if \( \omega_{\ell} \) ends with \( f \) and \( \omega_{\ell'} \) is the first rotation of \( \omega_{\ell} \), then \( F \) induces a \( p \)-linear isomorphism \( M_{\ell+1}/M_\ell \to M_{\ell'+1}/M_{\ell'} \). On the other hand, if \( \omega_{\ell} \) ends with \( v \), then \( V^{-1} \) induces a \( p \)-linear isomorphism \( M_{\ell+1}/M_\ell \to M_{\ell'+1}/M_{\ell'} \). Thus the displayed filtration is an admissible filtration. Via the action of \( \mathcal{W} \) on \( M \), the word \( \omega_{\ell} \) induces a semi-linear automorphism of \( M_{\ell+1}/M_\ell \), while a power of \( \omega_{\ell} \) induces the zero map of \( M_{\ell'+1}/M_{\ell'} \) if \( \ell \neq \ell' \). It follows that \( \omega_{\ell}^n M = M_{\ell+1} \) for large enough \( n \), so this is the coarsest filtration, thus the canonical filtration.

It remains to record the values of \( r \) and the functions \( \phi, \nu, \) and \( \rho \) associated to \( M \). From what was said above, it is clear that:
\begin{itemize}
  \item $r = \# \{ \ell \mid 0 \leq \ell < s, \omega_\ell \text{ ends with } f \}$;
  \item $\phi(i) = \# \{ \ell \mid 0 \leq \ell < i, \omega_\ell \text{ ends with } f \}$;
  \item $\nu(i) = r + \# \{ \ell \mid 1 \leq \ell < i, \omega_\ell \text{ ends with } v \}$; and
  \item $\rho(i) = \sum_{i=0}^{i-1} \mu(\ell)$.
\end{itemize}

3.5.1. Example. Let $\overline{w}_1 = f v, w_2 = f f v v$, and $m_1 = m_2 = 1$. Then, taking $n_1 = 2$ and $n_2 = 1$, one finds that $T$ contains $\omega_0 = f f v v$ and $\omega_1 = v f v f$, each with multiplicity 3. The function $\tau$ is

$$
\tau(1, 0) = \tau(2, 0) = \tau(2, 2) = 0 \quad \text{and} \quad \tau(1, 1) = \tau(2, 1) = \tau(2, 1) = 1,
$$

and $\mu(0) = \mu(1) = 3$. Thus $s = 2, r = 1$, and the functions $\phi, \nu$, and $\rho$ are given by

\[
\begin{array}{c|ccc}
  i & 0 & 1 & 2 \\
  \hline
  \phi(i) & 0 & 0 & 1 \\
  \nu(i) & 1 & 2 & 2 \\
  \rho(i) & 0 & 3 & 6 \\
\end{array}
\]

4. Duality and E–O structures

In this section, we consider self-dual $BT_1$ modules and their Ekedahl–Oort structures.

4.1. Duality of $BT_1$ modules. We defined the dual of a Dieudonné module in Section \[2.1.3\]. Here we record how duality interacts with the objects in Diagram (3.1). All the assertions in this section will be left to the reader.

For a $BT_1$ module $M$, let $M^*$ is its dual. If $N \subset M$ is a $k$-subspace, then

$$
F \left( N^\perp \right) = \left( V^{-1} N \right)^\perp \quad \text{and} \quad V^{-1} \left( N^\perp \right) = \left( FN \right)^\perp.
$$

Let $0 = M_0 \subset M_1 \subset \cdots \subset M_s = M$ be the canonical filtration of $M$; setting $M_i^* = (M_{s-i})^\perp$, then the canonical filtration of $M^*$ is

$$
0 = M_0^* \subset M_1^* \subset \cdots \subset M_s^* = M^*.
$$

If the canonical data attached to $M$ is $(r, s, \phi, \nu, \rho)$, and the canonical data attached to $M^*$ is $(r^*, s^*, \phi^*, \nu^*, \rho^*)$, then $s^* = s, r^* = s - r$, and for $0 \leq i \leq s$,

$$
\phi^*(i) = s - \nu(s - i), \quad \nu^*(i) = s - \phi(s - i), \quad \text{and} \quad \rho^*(i) = \rho(s) - \rho(s - i).
$$

It follows that $M$ is self-dual if and only if the associated canonical data satisfies $s = 2r$,

$$
\phi(i) + \nu(s - i) = s, \quad \text{and} \quad \rho(i) + \rho(s - i) = \rho(s).
\quad (4.1.1)
$$

The relationship between the partitioned set with permutation associated to $M$ and to $M^*$ is $S^* = S, S_f^* = S_v, S_v^* = S_f$, and $\pi^* = \pi$. It follows that $M$ is self-dual if and only if there exists a bijection $d : S \rightarrow S$ which satisfies $d(S_f) = S_v$, and $\pi \circ d = d \circ \pi$.

As before, if $w$ is a primitive word on $\{f, v\}$, define $w^c$ to be the word obtained by exchanging $f$ and $v$. This operation descends to a well-defined involution on cyclic words and $M(w) = M(w^c)$. It follows that $M$ is self-dual if and only if the associated multiset of cyclic primitive words consists of self-dual words $(\overline{w}^c = \overline{w})$ and pairs of dual words $(\{\overline{w}, \overline{w}^c\})$.
4.2. Ekedahl–Oort classification of polarized $BT_1$ modules. A polarized $BT_1$ module is a $BT_1$ module $M$ equipped with a non-degenerate, alternating pairing $\langle \cdot, \cdot \rangle : M \times M \to k$ of Dieudonné modules (i.e., such that $\langle x, x \rangle = 0$ and $\langle Fx, y \rangle = \langle x, Vy \rangle^p$ for all $x, y \in M$). Clearly, a polarized $BT_1$ module is self-dual. Conversely, as we will see below (Corollary 4.2.3), any self-dual $BT_1$ module can be given a polarization.

In this section, we review the Ekedahl–Oort’s classification of polarized $BT_1$ modules. We compare their description of $BT_1$ modules using elementary sequences to Kraft’s description in terms of words. Both descriptions have advantages lacked by the other. The elementary sequence gives little information on the decomposition of a $BT_1$ module into indecomposables, whereas this is immediate from the Kraft description. On the other hand, $BT_1$ modules corresponding to elementary sequences are in an obvious sense defined over $\mathbb{F}_p$ and this description is more convenient for questions of moduli, as demonstrated in [Oor01] and other papers.

4.2.1. Elementary sequences. Elementary sequences are a convenient repackaging of the data of a self-dual canonical type $(r, s, \phi, \nu, \rho)$. Using Equation 4.1.1, the restrictions of $\phi$ and $\rho$ to $\{0, \ldots, r\}$ determine the rest of the data. An elementary sequence of length $g$ is a sequence of integers $\psi_1, \ldots, \psi_g$ with $\psi_{i-1} \leq \psi_i \leq \psi_{i-1} + 1$ for $i = 1, \ldots, g$. It is convenient to set $\psi_0 = 0$. The set of elementary sequences of length $g$ has cardinality $2^g$.

Given $(r, s, \phi, \nu, \rho)$, we define an elementary sequence as follows. Let $g = \rho(r)$. Set $\psi_0 = 0$, and For each $1 \leq j \leq g$, let $i$ be the unique integer $0 < i \leq r$ such that $\rho(i - 1) < j \leq \rho(i)$. Then define

$$\psi_j = \begin{cases} 
\psi_{j-1} & \text{if } \phi(i) = \phi(i - 1), \\
\psi_{j-1} + 1 & \text{if } \phi(i) > \phi(i - 1).
\end{cases}$$

(Put more vividly, the sequence of $\psi_i$ increases for $\rho(i)$ steps if $\phi(i) > \phi(i - 1)$ and it stays constant for $\rho(i)$ steps if $\phi(i) = \phi(i - 1)$.)

We leave it as an exercise for the reader to check that, given an elementary sequence, there is a unique self-dual canonical type giving rise to it by this construction.

Elementary sequences can be obtained directly from a self-dual $BT_1$ module as follows: The canonical filtration can be refined into a “final filtration,” i.e., a filtration

$$0 = M_0 \subsetneq M_1 \subsetneq \cdots \subsetneq M_{2g} = M$$

respected by $F$ and $V^{-1}$ and such that $\dim_k(M_i) = i$. The corresponding elementary sequence is then defined by $\psi_i = \dim_k(FM_i)$.

Oort proved the following theorem, see [Oor01] Thm. 9.4.

4.2.2. **Theorem.** Every elementary sequence of length $g$ arises from a polarized $BT_1$ module of dimension $2g$, and two polarized $BT_1$ modules over $k$ with the same elementary sequences are isomorphic. More precisely, there is an isomorphism of $BT_1$ modules which respects the alternating pairings. (This isomorphism is not unique in general.)

The elementary sequence attached to a $BT_1$ module is also called its Ekedahl–Oort structure.

4.2.3. **Corollary.** Every self-dual $BT_1$ module admits a polarization, i.e., a non-degenerate alternating pairing, and this pairing is unique up to (non-unique) isomorphism.
Proof. If $M$ is a self-dual $BT_1$ module, construct its canonical type, and then its elementary sequence as in Section 4.2.1. Theorem 4.2.2 furnishes a polarized $BT_1$ module with the same underlying $BT_1$ module, and this proves the existence of a polarization. For uniqueness, note that the construction of the elementary sequence does not depend on the pairing. So, given two alternating pairings on a $BT_1$ module, Theorem 4.2.2 shows there is a (not necessarily unique) automorphism of the module intertwining the pairings.

5. Homomorphisms

The Kraft description of $BT_1$ modules is well adapted to computing homomorphisms. We work out three important examples in this section.

5.1. Homs from $\mathbb{Z}/p\mathbb{Z}$ or $\mu_p$.

5.1.1. Definition. If $G$ is a $BT_1$ group scheme, the $p$-rank of $G$ is the largest integer $f$ such that there is an injection $(\mathbb{Z}/p\mathbb{Z})^f \hookrightarrow G$. Alternatively, $f$ is the dimension of the largest quotient space of $M(G)$ on which Frobenius acts bijectively.

5.1.2. Lemma. If $G$ is a $BT_1$ group scheme, the $p$-rank of $G$ is equal to the multiplicity of the word $f$ in the multiset of primitive words corresponding to $M(G)$. Similarly, the largest $f$ such that $\mu_f^p$ embeds in $M(G)$ is the multiplicity of the word $v$.

Proof. From the presentation in terms of generators and relations, we see that if $w$ is a primitive word other than $f$, then there is no non-zero homomorphism from $\mathbb{Z}/p\mathbb{Z} = M(f)$ to $M(w)$. It follows that if $M(G)$ is given in the Kraft classification by a multiset of primitive cyclic words, the $p$-rank is the multiplicity of the word $f$. The assertion for $\mu_f^p$ is proved analogously.

5.2. Homs from $\alpha_p$. Let $G$ be a finite group scheme over $k$ killed by $p$. Define the foot (or socle) of $G$ to be the largest semisimple subgroup of $G$. The simple objects in the category of finite group schemes over $k$ killed by $p$ are $\mathbb{Z}/p\mathbb{Z}$, $\mu_p$, and $\alpha_p$. Thus the foot of $G$ is a direct sum of these (with multiplicity). If $G$ is connected and unipotent, then its foot is of the form $\alpha_p^\ell$, for some positive integer $\ell$. Note that $M(\alpha_p) = \mathbb{D}_k/(F, V)$.

Similarly, if $M$ is a $\mathbb{D}_k$-module of finite length, the head of $M$ is its largest semisimple quotient. We write $\mathcal{H}(M)$ for the head of $M$. If $M$ is a $\mathbb{D}_k$-module on which $F$ and $V$ are nilpotent, then $\mathcal{H}(M)$ is a $k$-vector space on which $F = V = 0$. The presentation of $M(w)$ by generators and relations makes it clear that if $w = v^{n_1}f^{m_1} \cdots v^{n_\ell}f^{m_\ell}$, then $\mathcal{H}(M(w))$ has dimension $\ell$. More precisely, the images of the generators $E_0, \ldots, E_{\ell-1}$ in $\mathcal{H}(M(w))$ are a basis.

5.2.1. Definition. If $G$ is a finite group scheme over $k$ killed by $p$, the $a$-number of $G$, denoted $a(G)$, is the largest integer $a$ such that there is an injection $\alpha_p^a \hookrightarrow G$ of group schemes. If $G$ is connected and unipotent, $p^{a(G)}$ equals the order of the foot of $G$. Similarly, if $M$ is a $\mathbb{D}_k$-module of finite length, the $a$-number of $M$, denoted $a(M)$, is the largest integer $a$ such that there is a surjection $M \twoheadrightarrow M(\alpha_p) = \mathbb{D}_k/(F, V)$ of $\mathbb{D}_k$-modules.

5.2.2. Lemma. The $a$-numbers of $BT_1$ modules have the following properties:

1. They are additive in direct sums: $a(M_1 \oplus M_2) = a(M_1) + a(M_2)$.
2. $a(M(1)) = a(M(\nu)) = 0$.
3. If $\ell, m_1, \ldots, m_\ell, n_1, \ldots, n_\ell$ are positive integers, then $a(M(v^{n_1}f^{m_1} \cdots v^{n_\ell}f^{m_\ell})) = \ell$. 

(4) \(a(M(w))\) is the number of rotations of \(w\) which start with \(v\) and end with \(f\).

Proof. Parts (1) and (2) are clear. For part (3), \(a(M) = \dim_k(\mathcal{H}(M))\) and, by the discussion above, if \(w = v^{n_1} f^{m_1} \cdots v^{n_1} f^{m_1}\), then \(a(M(w)) = \ell\). Part (4) is immediate from part (3). \(\square\)

5.3. **Homs from** \(G_{1,1}\). Let \(G_{1,1}\) be the \(BT_1\) group scheme over \(k\) whose Dieudonné module is isomorphic to \(\mathbb{D}_k/(F - V) \cong M(fv)\). Note that \(G_{1,1}\) is unique up to isomorphism and it appears “in nature” as the kernel of multiplication by \(p\) on a supersingular elliptic curve over \(k\) (e.g., see [Ulm91, Prop. 4.1]). Similarly, write \(M_{1,1} := M(G_{1,1}) = M(fv)\).

5.3.1. **Definitions.** Let \(G\) be a \(BT_1\) group scheme over \(k\) and let \(M\) be a \(BT_1\) module.

1. Define the \(s_{1,1}\)-multiplicity of \(G\) as the largest integer \(s\) such that there is an isomorphism of group schemes
   \[ G \cong G_{1,1} \oplus G'. \]
   Define the \(s_{1,1}\)-multiplicity of \(M\) as the largest integer \(s\) such that there is an isomorphism of \(\mathbb{D}_k\)-modules \(M \cong M_{1,1} \oplus M'\).

2. Define the \(u_{1,1}\)-number of \(G\) as the largest integer \(u\) such that there exists an injection
   \[ G_{1,1} \hookrightarrow G \]
   of group schemes. Define the \(u_{1,1}\)-number of \(M\) as the largest integer \(u\) such that there is a surjection \(M \twoheadrightarrow M_{1,1}^u\) of \(\mathbb{D}_k\)-modules.

The notation \(s_{1,1}\) (resp. \(u_{1,1}\)) is motivated by the word superspecial (resp. unpolarized), see Section 5.4. Note that \(s_{1,1}(M(G)) = s_{1,1}(G)\) and \(u_{1,1}(M(G)) = u_{1,1}(G)\). It is clear that \(u_{1,1}(G) \geq s_{1,1}(G)\). The \(s_{1,1}\)-multiplicity and \(u_{1,1}\)-number are additive in direct sums.

By the Kraft classification, if \(M(G)\) is described by a multiset of primitive cyclic words, then \(s_{1,1}(G)\) equals the multiplicity of the cyclic word \(fv\) in the multiset.

We want to compute the \(u_{1,1}\)-number of the standard \(BT_1\) modules \(M(w)\). Trivially,
\[ u_{1,1}(M(f)) = u_{1,1}(M(v)) = 0. \]

A straightforward exercise shows that \(u_{1,1}(M(fv)) = 1\), and more precisely that
\[ \text{Hom}_{\mathbb{D}_k}(M(fv), M(fv)) \cong \mathbb{F}_{p^2} \times k , \]
with \((c, d) \in \mathbb{F}_{p^2} \times k\) identified with the homomorphism that sends the class of \(1 \in \mathbb{D}_k/(F - V)\) to the class of \(c + dF\). The surjective homomorphisms are those where \(c \neq 0\).

We may thus assume that \(w\) has length \(\lambda > 2\). We will evaluate the \(u_{1,1}\)-number of \(M(w)\) by computing \(\text{Hom}_{\mathbb{D}_k}(M(w), M_{1,1})\) explicitly. To that end, write \(w = v^{n_1} f^{m_1} \cdots v^{n_1} f^{m_1}\). Since \(\lambda > 2\) and \(w\) is non-periodic, we may replace \(w\) with a rotated word so that \(m_1 > 1\) or \(n_1 > 1\) (or both). Roughly speaking, the following proposition says that the \(u_{1,1}\)-number of \(M(w)\) is the number of appearances in \(w\) of subwords of the form \(v^{>1}(fv)^e f^{>1}\) where \(e \geq 0\). For example, if \(w = v^{j} f^{2} v^{j} f^{1}\), then the \(u_{1,1}\)-number is 2.

5.3.2. **Proposition.** Suppose \(w = v^{n_1} f^{m_1} \cdots v^{n_1} f^{m_1}\) where \(m_1 > 1\) or \(n_1 > 1\). Define \(u\) by:

\[ u := \# \{ 1 \leq i \leq \ell \mid m_i > 1 \text{ and } n_i > 1 \} \]
\[ + \# \{ 1 \leq i < j \leq \ell \mid n_j > 1, m_j = n_{j-1} = \cdots = n_i = 1, \text{ and } m_i > 1 \}. \]

Then
(1) Hom$_{\text{df}}(M(w), M_{1,1})$ is in bijection with $k^{u+\ell}$, and
(2) the $u_{1,1}$-number of $M(w)$ is $u$.

Proof: By Lemma [3.2.3] $M(w)$ has a presentation with generators $E_0, \ldots, E_{\ell-1}$ (with indices taken modulo $\ell$) and relations $F^{m_i}E_{i-1} = V^{n_i}E_i$. Let $z_0, z_1$ be a $k$-basis of $M_{1,1}$ with $Fz_0 = Vz_0 = z_1$ and $Fz_1 = Vz_1 = 0$. Then a homomorphism $\psi : M(w) \to M_{1,1}$ is determined by its values on the generators $E_i$. Write

$$\psi(E_i) = a_{i,0}z_0 + a_{i,1}z_1.$$ 

Then $\psi$ is a $\mathbb{D}_k$-module homomorphism if and only if $F^{m_i}\psi(E_{i-1}) = V^{n_i}\psi(E_i)$ for $i = 1, \ldots, \ell$. By (2.1.1),

$$F^{m_i}\psi(E_{i-1}) = \begin{cases} a_{i-1,0}^{p}z_1 & \text{if } m_i = 1 \\
0 & \text{if } m_i > 1 \end{cases} \quad \text{and} \quad V^{n_i}\psi(E_i) = \begin{cases} a_{i,0}^{1/p}z_1 & \text{if } n_i = 1 \\
0 & \text{if } n_i > 1. \end{cases}$$

This system of equations places no constraints on $a_{i,1}$ for $i = 1, \ldots, \ell$, because $Vz_1 = Fz_1 = 0$. The constraints on $a_{i,0}$ for $i = 1, \ldots, \ell$ are: if $m_i = n_i = 1$, then $a_{i-1,0}^{p} = a_{i,0}^{1/p}$; if $m_i = 1$ and $n_i > 1$, then $a_{i-1,0} = 0$; if $m_i > 1$ and $n_i = 1$, then $a_{i,0} = 0$; if $m_i > 1$ and $n_i > 1$, then no constraint. Using that $m_1 > 1$ or $n_\ell > 1$, we find a triangular system of equations for $a_{i,0}$, and it is a straightforward exercise to show that the solutions are in bijection with $k^u$. Combining with the $k^\ell$ unconstrained values of $\{a_{i,1} \mid 1 \leq i \leq \ell\}$ yields part (1).

For part (2), the homomorphism $\psi : M(w) \to M_{1,1}$ is surjective if and only if at least one of the $a_{i,0}$ is not zero, which is equivalent to $\psi$ inducing a surjection $H(M(w)) \to H(M_{1,1}) = k$. Part (1) implies that there are $u$ independent such $\psi$ (and no more). This shows that $u$ is the largest integer such that there is a surjection $M(w) \to M_{1,1}^u$, completing the proof of (2). \[\square\]

The next proposition motivates our consideration of the $s_{1,1}$-multiplicity and the $u_{1,1}$-number.

5.3.3. Proposition. Let $A/k$ be an abelian variety and let $E/k$ be a supersingular elliptic curve.

(1) If there is an abelian variety $B$ and an isogeny $E^s \times B \to A$ of degree prime to $p$, then $s$ is less than or equal to the $s_{1,1}$-multiplicity of $A[p]$.

(2) If there is a morphism of abelian varieties $E^u \to A$ with finite kernel of order prime to $p$, then $u$ is less than or equal to the $u_{1,1}$-number of $A[p]$.

Proof: An isogeny as in (1) shows that $E[p] \cong G_{1,1}$ is a direct factor of $A[p]$ with multiplicity $s$, so $s \leq s_{1,1}(A[p])$. A morphism as in (2) shows that $E[p] \cong G_{1,1}$ appears in $A[p]$ with multiplicity at least $u$, so $u \leq u_{1,1}(A[p])$. \[\square\]

5.4. Connection to the superspecial rank of Achter and Pries. In [AP15 Def. 3.3], Achter and Pries consider a "principally quasi-polarized" $BT_1$ group scheme (i.e., a $BT_1$ group scheme $G$ such that $M(G)$ is equipped with a non-degenerate, alternating pairing), and they define its superspecial rank as the largest integer $s$ such that there is an injection $G^s_{1,1} \hookrightarrow G$ such that the polarization on $G$ restricts to a non-degenerate pairing on $G^s_{1,1}$. In this situation, the pairing allows them to define a complement (see [AP15 Lemma 3.4]), so that $G \cong G^s_{1,1} \oplus G'$ (a direct sum of polarized $BT_1$ group schemes). Thus their superspecial rank is at most the $s_{1,1}$-multiplicity we defined above. It follows from Corollary [4.2.3] that a self-dual $BT_1$ group scheme $G$ equipped with a decomposition $G \cong G_{1,1} \oplus G'$ (just of $BT_1$ group schemes) automatically admits a principal
quasi-polarization compatible with the direct sum decomposition. Therefore, the superspecial rank of a principally quasi-polarized $BT_1$ group scheme is equal to its $s_{1,1}$-multiplicity.

They also define an *unpolarized superspecial rank*, which is the same as our $u_{1,1}$-number, and prove a result [AP15, Lemma 3.8] which is closely related to and implied by Proposition 5.3.2.

We end this section with a correction to [AP15, Thm. 3.14]. Let $K$ be the function field of an irreducible, smooth, proper curve $X$ over $k$, let $J_X$ be the Jacobian of $X$, and let $E$ be a supersingular elliptic curve over $k$ which we regard as a curve over $K$ by base change. Let $Sel(K,p)$ denote the Selmer group for the multiplication-by-$p$ isogeny of $E/K$. This group is defined and studied in detail in [Ulm91].

5.5. **Proposition.** With notation as above, let $a$, $u_{1,1}$, and $s_{1,1}$ be the $a$-number, $u_{1,1}$-number, and $s_{1,1}$-multiplicity of $J_X[p]$ respectively. Then $Sel(K,p)$ is isomorphic to the product of a finite group and a $k$-vector space of dimension $a + u_{1,1} - s_{1,1}$.

**Proof.** Applying [Ulm91, Props. 6.2 and 6.4] with $C = X$, $D = E$, $\Delta = 1$, and $n = 1$ shows that $Sel(K,p)$ differs by a finite group from $Hom_D(H^{1}_{dR}(X),M_{1,1})$. By [Oda69, Cor. 5.11], $H^{1}_{dR}(X) \cong M(J_X[p])$. Write $M(J_X[p])$ as a sum of indecomposable $BT_1$ modules $M(w)$ for suitable cyclic words $w$. In Section 4.3, we computed $H_w = Hom_D(M(w),M_{1,1})$ for a primitive cyclic word $w$. Recall that $H_w = 0$ if $w = f$ or $w = v$ and $H_w \cong \mathbb{F}_{p^2} \times k$ if $w = fv$. Also $H_w \cong k^{u+\ell}$, if $w \not\in \{1,f,v,fv\}$ where $u = u_{1,1}(M(w))$ and $\ell$ is the $a$-number of $M(w)$. The result follows from the additivity of $a$-numbers, $u_{1,1}$-numbers, and $s_{1,1}$-multiplicities. \hfill $\square$

5.6. **Examples.** For small genus, we give tables of elementary sequences (“E–O”), written as lists of integers, matched with the self-dual multisets of primitive cyclic words (“K”), together with their $p$-ranks, $a$-numbers, $s_{1,1}$-multiplicities, and $u_{1,1}$-numbers. The notation $(w)^e$ stands for the cyclic word $w$ with multiplicity $e$.

From Section 4.2.1 for the $BT_1$ module corresponding to the elementary sequence $[\psi_1, \ldots, \psi_g]$, one can see that the $p$-rank is the largest $i$ such that $\psi_i = i$ and the $a$-number is $g - \psi_g$. For the $BT_1$ module corresponding to a multiset of cyclic words, the $p$-rank is the multiplicity of the word $f$ by Lemma 5.1.2 and the $a$-number can be computed using Lemma 5.2.2.

For the Krafl classification of a $BT_1$ module, the $s_{1,1}$-multiplicity is the multiplicity of the cyclic word $fv$ and the $u_{1,1}$-number can be computed using Proposition 5.3.2. We do not know how to compute the $s_{1,1}$-multiplicity or $u_{1,1}$-number directly from the elementary sequence.

\begin{center}
\begin{tabular}{|c|c|c|c|c|}
\hline
\multicolumn{5}{|c|}{\textbf{\textit{g = 1}}} \\
\hline
E–O & K & $p$-rank & $a$-number & $s_{1,1}$-mult. & $u_{1,1}$-number \\
\hline
[0] & $\{fv\}$ & 0 & 1 & 1 & 1 \\
[1] & $\{f,v\}$ & 1 & 0 & 0 & 0 \\
\hline
\end{tabular}
\end{center}

\begin{center}
\begin{tabular}{|c|c|c|c|c|}
\hline
\multicolumn{5}{|c|}{\textbf{\textit{g = 2}}} \\
\hline
E–O & K & $p$-rank & $a$-number & $s_{1,1}$-mult. & $u_{1,1}$-number \\
\hline
[0,0] & $\{(fv)^2\}$ & 0 & 2 & 2 & 2 \\
[0,1] & $\{ffv\}$ & 0 & 1 & 0 & 1 \\
[1,1] & $\{f,v,fv\}$ & 1 & 1 & 1 & 1 \\
[1,2] & $\{(f)^2,(v)^2\}$ & 2 & 0 & 0 & 0 \\
\hline
\end{tabular}
\end{center}
The E–O structure [0, 1] is the first where the \( u_{1,1} \)-number is greater than the \( s_{1,1} \)-multiplicity. If \( G \) is a polarized \( BT_1 \) group scheme of order \( p^{2g} \) with positive \( p \)-rank, then \( G \cong G' \oplus \mathbb{Z}/p\mathbb{Z} \oplus \mu_p \) where \( G' \) is a polarized \( BT_1 \) group scheme of order \( p^{2g-2} \). Thus the rows of the table for \( g \) for \( G \) with positive \( p \)-rank can be deduced from the table for \( g - 1 \). In passing from genus \( g - 1 \) to genus \( g \): the elementary sequence changes from \([\psi_1, \ldots, \psi_{g-1}] \) to \([1, \psi_1 + 1, \ldots, \psi_{g-1} + 1] \); the multiplicity of the words \( f \) and \( v \) increases by 1; the \( p \)-rank increases by 1; and the \( a \)-number, \( s_{1,1} \)-multiplicity, and \( u_{1,1} \)-number stay the same. In light of this, when \( g = 3 \) and \( g = 4 \), we only include the \( BT_1 \) group schemes with \( p \)-rank 0 in the table.

\[
\begin{array}{|c|c|c|c|c|}
\hline
\text{E–O} & \text{K} & \text{p-rank} & \text{a-number} & \text{s}_{1,1}\text{-mult.} & \text{u}_{1,1}\text{-number} \\
\hline
[0, 0, 0] & \{(fv)^3\} & 0 & 3 & 3 & 3 \\
[0, 0, 1] & \{fv, ffvv\} & 0 & 2 & 1 & 2 \\
[0, 1, 1] & \{fv, vff\} & 0 & 2 & 0 & 0 \\
[0, 1, 2] & \{fffvvv\} & 0 & 1 & 0 & 1 \\
\hline
\end{array}
\]

The E–O structure [0, 1, 1] is the first which is decomposable as a \( BT_1 \) group scheme but indecomposable as a polarized \( BT_1 \) group scheme.

\[
\begin{array}{|c|c|c|c|c|c|}
\hline
\text{E–O} & \text{K} & \text{p-rank} & \text{a-number} & \text{s}_{1,1}\text{-mult.} & \text{u}_{1,1}\text{-number} \\
\hline
[0, 0, 0, 0] & \{(fv)^4\} & 0 & 4 & 4 & 4 \\
[0, 0, 0, 1] & \{(fv)^2, ffvv\} & 0 & 3 & 2 & 3 \\
[0, 0, 1, 1] & \{ffvfvf\} & 0 & 3 & 0 & 1 \\
[0, 0, 1, 2] & \{(ffvv)^2\} & 0 & 2 & 0 & 2 \\
[0, 1, 1, 1] & \{fv, fffv, vff\} & 0 & 3 & 1 & 1 \\
[0, 1, 1, 2] & \{fv, fffvfv\} & 0 & 2 & 1 & 2 \\
[0, 1, 2, 2] & \{fffv, fvvv\} & 0 & 2 & 0 & 0 \\
[0, 1, 2, 3] & \{fffvvv\} & 0 & 1 & 0 & 1 \\
\hline
\end{array}
\]

6. Fermat Jacobians

In this section, we apply the technology from Sections 2, 3, and 4 to study \( p \)-torsion group schemes of Jacobians of Fermat curves.

6.1. \( BT_1 \) modules associated to curves. Let \( C \) be an irreducible, smooth, projective curve of genus \( g \) over \( k \), and let \( J = J_C \) be its Jacobian. In [Oda69 §5], Oda gives \( H^{1}_{dR}(C) = H^{1}_{dR}(J) \) the structure of a \( BT_1 \) module. More precisely, if we choose a covering of \( C \) by affine open subsets \( \{U_i\} \), then we can represent a class \( c \) in \( H^{1}_{dR}(C) \) by a hyper-cocycle

\[
(f_{ij}, \omega_i) \quad \text{with} \quad f_{ij} \in \mathcal{O}_C(U_i \cap U_j), \quad \omega_i \in \Omega^1_C(U_i), \quad \text{and} \quad df_{ij} = \omega_i - \omega_j.
\]

Furthermore, \( Fc \) is represented by \((f_{ij}^p, 0)\) and \( Vc \) is represented by \((0, \text{Car}(\omega_i))\) where \( \text{Car} \) is the Cartier operator. In particular, writing \( H \) for \( H^{1}_{dR}(C) \), then

\[
\text{Im}(V : H \to H) = \text{Ker}(F : H \to H) \cong H^0(C, \Omega^1_C),
\]

(6.1.1)
and
\[ \text{Im}(F : H \to H) = \text{Ker}(V : H \to H) \cong H^1(C, \mathcal{O}_C). \]
Moreover, Oda proves [Oda69, Cor. 5.11] that \( H \) is the Dieudonné module of the \( p \)-torsion of \( J \), i.e., there is a canonical isomorphism of \( \mathbb{D}_k \)-modules
\[ H^1_{dR}(C) \cong M(J[p]). \] (6.1.2)

6.2. **Fermat curves.** For each positive integer \( d \) not divisible by \( p \), let \( F_d \) be the Fermat curve of degree \( d \), i.e., the smooth, projective curve over \( k \) with affine model
\[ F_d : \quad x^d + y^d = 1, \]
and let \( J_{F_d} \) be its Jacobian.
Let \( C_d \) be the smooth, projective curve over \( k \) with affine model
\[ C_d : \quad y^d = x(1 - x). \] (6.2.1)
Then \( C_d \) is a quotient of \( F_d \). (Substitute \( x^d \) for \( x \) and \( xy \) for \( y \) in the equation for \( C_d \).) The map \( F_d \to C_d \) is the quotient of \( F_d \) by a subgroup of \( (\mu_d)^2 \subset \text{Aut}(F_d) \) of index \( d \). There are other quotients of \( F_d \), having equations \( x^e = u^s(1 - u)^r \) for \( e \mid d \), \( \gcd(e, r, s) = 1 \), and \( r + s \leq e - 1 \). All of the interesting features of \( J_{F_d}[p] \) are already present in \( J_{C_d}[p] \), so we restrict to studying this case for simplicity. For the convenience of the reader, we state (at Remark 6.9 below) how our combinatorial description of \( J_{C_d}[p] \) carries over to the case of \( J_{F_d}[p] \).

6.3. **Cohomology of \( C_d \).** The Riemann-Hurwitz formula shows that the genus of \( C_d \) is
\[ g(C_d) = \lfloor (d - 1)/2 \rfloor = \begin{cases} (d - 1)/2 & \text{if } d \text{ is odd,} \\ (d - 2)/2 & \text{if } d \text{ is even.} \end{cases} \]
Moreover, \( C_d \) admits an action of \( \zeta \in \mu_d \) with \( \zeta : (x, y) \mapsto (x, \zeta y) \).

We next describe \( H^1_{dR}(C_d) \) in a form conducive to studying it as a \( \mathbb{D}_k \)-module. First, write
\[ H^1_{dR}(C_d) = \bigoplus_{a \in \mathbb{Z}/d\mathbb{Z}} H_a, \]
where \( H_a \) is the subspace of \( H^1_{dR}(C_d) \) where every \( \zeta \in \mu_d \) acts by multiplication by \( \zeta^a \). Since the action of \( \mu_d \) on \( C_d \) induces the trivial action on \( H^2_{dR}(C_d) \), the cup product induces a perfect duality between \( H_a \) and \( H_{-a} \), and a trivial pairing between \( H_a \) and \( H_b \) if \( b \neq -a \mod d \).
Let
\[ S = \begin{cases} \mathbb{Z}/d\mathbb{Z} \setminus \{0\} & \text{if } d \text{ is odd,} \\ \mathbb{Z}/d\mathbb{Z} \setminus \{0, d/2\} & \text{if } d \text{ is even.} \end{cases} \]
Note that multiplication by \( p \) induces a permutation of \( S \). For the rest of the paper, whenever we make an archimedean statement about an element \( a \in S \) (e.g., \( 0 < a < d/2 \)), we make sense of it by implicitly lifting \( a \) to its least positive residue.

6.4. **Proposition.**
(1) If \( a \in \mathbb{Z}/d\mathbb{Z} \), then \( \dim_k(H_a) = 1 \) if \( a \in S \) and \( H_a = 0 \) if \( a \notin S \).
(2) \( H^0(C_d, \Omega^1_{C_d}) = \bigoplus_{0 < a < d/2} H_a \).
(3) If \( 0 < a < d/2 \), then \( FH_a = 0 \) and \( V \) induces an isomorphism \( V : H_{pa} \to H_a \).
(4) If \( d/2 < a < d \), then \( V H_a = 0 \) and \( F \) induces an isomorphism \( F : H_a \to H_{pa} \).
Proof. For $0 < a < d/2$, a simple calculation shows that the 1-form $g^a dx/y^d$ on the affine model $(6.2.1)$ extends to a global 1-form on $C_d$, and its class in $H^1_{dR}(C_d)$ lies in $H_a$. This shows that $\dim \ker(H_a) \geq 1$ for $0 < a < d/2$. Because of the perfect duality between $H_a$ and $H_{-a}$, we see that $\dim \ker(H_a) \geq 1$ for $d/2 < a < d$. Since $g(C_d) = [(d-1)/2]$, it follows that $\dim \ker(H_a) = 1$ for $a \in S$ and $H_a = 0$ for $a \notin S$. This proves parts (1) and (2).

By definition, $FH_a \subset H_{pa}$ and $VH_{pa} \subset H_a$. By $(6.1.1)$, $F$ kills $H^0(C_d, \Omega^1_{C_d}) = \oplus_{0 < a < d/2} H_a$. Since $\dim(\ker F) = g$, the map $F : H_a \to H_{pa}$ is injective, and thus bijective, for $d/2 < a < d$. Similarly, by $(6.1.1)$, $\text{Im } V = H^0(C_d, \Omega^1_{C_d})$. So $V : H_{pa} \to H_a$ is surjective, and thus bijective, for $0 < a < d/2$ and zero for $d/2 < a < d$. This proves parts (3) and (4).

Now let $S_f, S_v \subset S$ be given by

$$S_f = \{ a \mid d/2 < a < d \} \quad \text{and} \quad S_v = \{ a \mid 0 < a < d/2 \},$$

and let $\pi : S \to S$ be the permutation induced by multiplication by $p$.

6.5. Theorem. The $BT_1$ module $H^1_{dR}(C_d)$ is isomorphic to that obtained from the data $S = S_f \cup S_v$ and $\pi : S \to S$ via Diagram (3.1).

By Corollary (4.2.3), given any non-degenerate, alternating form on the $BT_1$ module defined by $(S = S_f \cup S_v, \pi)$, we may choose the isomorphism in the theorem so that it intertwines the given form with the polarization on $H^1_{dR}(C_d)$ induced by the cup product.

Proof. Consider the set $O$ of orbits of $\pi$ on $S$. As in Section (3.1.2) we associate a cyclic word $\overline{w}_o$ to each orbit $o$. The $BT_1$ module associated to $S = S_f \cup S_v$ and $\pi$ is the direct sum

$$M = \bigoplus_{o \in O} M(\overline{w}_o).$$

It suffices to prove that, for each orbit $o$, there is an isomorphism of $\mathbb{D}_k$-modules,

$$M(\overline{w}_o) \cong \bigoplus_{a \in o} H_a,$$

where the right hand side has the structure of $\mathbb{D}_k$-module induced by that of $H^1_{dR}(C_d)$.

Consider an orbit $o \in O$ and let $\lambda$ be its length. Choosing a base point $o \in o$, we find a word $w_o = u_{\lambda-1} \cdots u_0$ representing the cyclic word $\overline{w}_o$. The letters in $w_o$ yield $p$-linear isomorphisms (induced by $F$ or $V^{-1}$)

$$H_a \cong H_{pa} \cong \cdots \cong H_{p^{\lambda}a} = H_a.$$

The composed isomorphism $H_a \cong H_a$ is $p^\lambda$-linear. Let $e_0 \in H_a$ be a non-zero element sent to itself under the composed isomorphism $H_a \cong H_a$. (Such an $e_0$ exists by the $p^\lambda$-linearity.) For $j \in \mathbb{Z}/\lambda\mathbb{Z}$, let $e_j \in H_{p^ja}$ be the image of $e_0$ under the displayed isomorphism, so we have

$$F e_j = \begin{cases} e_{j+1} & \text{if } u_j = f, \\ 0 & \text{if } u_j = v, \end{cases} \quad \text{and} \quad V e_{j+1} = \begin{cases} e_j & \text{if } u_j = v, \\ 0 & \text{if } u_j = f. \end{cases}$$

Comparing with the definition of $M(\overline{w}_o)$ in Section (3.2), we see that $\oplus_{a \in o} H_a$ is isomorphic as a $\mathbb{D}_k$-module to $M(\overline{w}_o)$. This completes the proof of the theorem. □
6.6. Remark. The decomposition of $H = H^1_{dR}(C_d)$ into subspaces corresponding to orbits,

$$H = \bigoplus_o (\oplus_{a \in o} H_a),$$

is not in general a decomposition of $H$ into indecomposable $\mathbb{D}_k$-modules. Indeed, if $o$ is an orbit giving rise to a non-primitive word, the summand corresponding to $o$ is decomposable. (See Section 3.2 and the next example.)

6.7. Example. Let $p = 5$, $d = 624 = 5^4 - 1$ and $o = \{84, 420, 228, 516\}$. Then the cyclic word $w_o$ is represented by $(vf)^2$, and $\bigoplus_{a \in o} H_a$ is isomorphic to the direct sum of two copies of $M(vf)$, but there is no such isomorphism with each summand being a direct sum of eigenspaces $H_a$.

Theorem 6.5 and (6.1.2) ([Oda69, Cor. 5.11]) imply this result about $J = J_d$, the Jacobian of $C_d$.

6.8. Corollary. The Dieudonné module $M(J_d[p])$ is the $BT_1$ module associated to the data

$$S = \begin{cases} \mathbb{Z}/d\mathbb{Z} \setminus \{0, d/2\} & \text{if } d \text{ is even} \\ \mathbb{Z}/d\mathbb{Z} \setminus \{0\} & \text{if } d \text{ is odd.} \end{cases}$$

$$S_f = \{a \in S \mid d/2 < a < d\}, \quad S_v = \{a \in S \mid 0 < a < d/2\},$$

and the permutation $\pi : S \to S$ given by $\pi(i) = pa$.

6.9. Remark. An analysis completely parallel to the above shows that the Dieudonné module of $J_{F_d}[p]$ (where $F_d$ is the Fermat curve of degree $d$) is associated to the data

$$S = \{(a, b) \in (\mathbb{Z}/d\mathbb{Z})^2 \mid a \neq 0, b \neq 0, a + b \neq 0\} ,$$

$$S_f = \{(a, b) \in S \mid a + b > d\}, \quad S_v = \{(a, b) \in S \mid a + b < d\},$$

and the permutation $\pi(a, b) = (pa, pb)$. (This uses a generalization of [Dum95, §5–6], from $d = q + 1$ to a general $d$, in place of Proposition 6.4.)

6.10. A Shimura variety perspective. Another calculation of the isomorphism type of $J_d[p]$ can be extracted from the literature as follows.

In [Moo04], Moonen considers $A_g$, the moduli space of principally polarized abelian varieties of dimension $g$, and its Shimura subvarieties of PEL type. The space (or rather stack) $A_g$ carries two stratifications, by Newton polygons and by Ekedahl–Oort types. Moonen proves that for each irreducible component of each Shimura subvariety of PEL type, the open stratum of the induced Newton stratification and that of the induced E–O stratification coincide. Moreover, the Newton polygon and the isomorphism class of polarized $BT_1$ group scheme that appear on this stratum can be computed from the data defining the Shimura subvariety. This Newton polygon is minimal among those that appear on the component, and it is called the “$\mu$-ordinary” Newton polygon.

Now consider $J_d$. The elementary calculation of Proposition 6.4 parts (1) and (2) (which goes back at least to [Wei76]) shows that $J_d$ is an abelian variety of CM type (because it admits endomorphisms by a quotient $E$ of the group ring $\mathbb{Q}[\mu_d]$), and it reveals the CM type of $J_d$. The Shimura subvariety of $A_g$ corresponding to abelian varieties with endomorphisms by $E$ is zero dimensional, so each point of it lies in the open stratum of the two induced stratifications. The result of Moonen then allows one to compute the isomorphism type of $J_d[p]$ in terms of the
Shimura data, i.e., in terms of the data given in Proposition 6.4 parts (1) and (2). The output of this calculation is equivalent to our Theorem 6.5. We refer to [Moo04 §1] for more details.

7. Ekedahl–Oort structures of Fermat Jacobians: Generalities

In this section, we work out the overall form of the Ekedahl–Oort structure of the Jacobian of a Fermat curve. Specific cases will be considered in later sections. As before, \(d\) is an integer prime to \(p\). Let \(C_d\) be the Fermat quotient curve with affine model \(y^d = x(1 - x)\) and \(J_d\) be its Jacobian.

7.1. Words, patterns, and multiplicities. From Corollary 6.8 recall that \(S = \mathbb{Z}/d\mathbb{Z} \setminus \{0\}\) if \(d\) is odd and \(S = \mathbb{Z}/d\mathbb{Z} \setminus \{0, d/2\}\) if \(d\) is even. We partitioned \(S\) as \(S = S_f \cup S_v\) where

\[
S_f = \{a \in S \mid d/2 < a < d\} \quad \text{and} \quad S_v = \{a \in S \mid 0 < a < d/2\}.
\]

There is an action on \(S\) of the cyclic subgroup \(\langle p \rangle \subset (\mathbb{Z}/d\mathbb{Z})^*\) in the group of units modulo \(d\).

If \(a \in S\) and \(\lambda > 0\) is minimal such that \(p^\lambda a = a\), then the word for \(a\) is \(w = t_{\lambda - 1} \cdots t_0\) where

\[
t_j = \begin{cases} f & \text{if } p^j a \in S_f, \\ v & \text{if } p^j a \in S_v. \end{cases}
\] (7.1.1)

Taking the class of \(w\) in the set of cyclic words on \(\{f, v\}\) gives a well-defined map from orbits of \(\langle p \rangle\) on \(S\) to cyclic words, and we used this above to study the Kraft classification of \(J_d[p]\).

We introduce a slight variation, the pattern for \(a \in S\). Fix \(\ell = |\langle p \rangle|\) to be the multiplicative order of \(p\) modulo \(d\). Let \(\mathcal{W}_\ell\) be the set of words of length \(\ell\) on \(\{f, v\}\). Define a map \(\text{Pat} : S \to \mathcal{W}_\ell\) as follows: for \(a \in S\), define \(\text{Pat}(a) = w = t_{\ell-1} \cdots t_0\) where \(t_j\) is defined as in (7.1.1). If the word for \(a\) has length \(\ell\) (which happens when \(\gcd(a, d) = 1\)), then the pattern of \(a\) and the word of \(a\) are the same, whereas for \(a\) with shorter words, the pattern is a power of the word. The virtue of working with patterns is that we can consider the set of words \(\mathcal{W}_\ell\) of fixed length \(\ell\). (In the notation of Section 3.5 if \(w_{i,j}\) is the word of \(a \in S\), then \(\tilde{w}_{i,j}\) is a power of the pattern of \(a\).)

For example, take \(d = 9\) and \(p = 2\), so that \(\ell = 6\). The orbits of \(\langle p \rangle\) are \(1 \to 2 \to 4 \to 8 \to 7 \to 5 \to 1\) and \(3 \to 6 \to 3\). For \(a = 3\), the word is \(vf\) and \(\text{Pat}(3) = fvfvf\). For \(a = 6\), the word is \(vf\) and \(\text{Pat}(6) = vfvf\). Otherwise, \(\text{Pat}(a)\) is the same as the word and given by:

\[
\begin{array}{cccccccc}
\text{a} & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\
\text{Word} & fffvvv & vfffvv & vvfffv & fffvff & ffvfff & vfvfff & vffvff \\
\end{array}
\]

For \(w \in \mathcal{W}_\ell\), define the multiplicity of \(w\) to be the cardinality of its inverse image under \(\text{Pat}:
\mu(w) := |\text{Pat}^{-1}(w)|.
\]

These multiplicities will be useful to determine the Ekedahl–Oort structure of \(J_d[p]\).

7.2. Ekedahl–Oort structure of \(J_d[p]\). If \(M\) is a \(BT_1\) module, recall that a final filtration is a filtration \(0 \subset M_1 \subset \cdots \subset M_2q\) respected by \(F\) and \(V^{-1}\) having the property that \(\dim_k(M_j) = j\). The Ekedahl–Oort structure is given by the elementary sequence of integers \([\psi_1, \ldots, \psi_g]\) where \(\psi_j = \dim_k(FM_j)\). Let \(\psi_0 = 0\) and recall that \(\psi_{j-1} \leq \psi_j \leq \psi_{j-1} + 1\) for \(j = 1, \ldots, g\).

To simplify the description of the E–O structure, we introduce the following notation: \(\nearrow^m\) stands for an increasing sequence of integers of length \(m\) and \(\rightarrow^n\) stands for a constant sequence of integers of length \(n\). Thus

\[
[\nearrow^3 \rightarrow^2] = [1, 2, 3, 3, 3] \quad \text{and} \quad [\rightarrow^2 \nearrow^3] = [0, 0, 1, 2, 3].
\]
Now enumerate all the elements of \( W \) that start with \( f \) in lexicographic order:
\[
w_0 = f^\ell, \quad w_1 = f^{\ell-1}v, \quad w_2 = f^{\ell-2}vf, \quad w_3 = f^{\ell-2}vv, \ldots, \quad w_{2\ell-1} = fv^{\ell-1}.
\] (7.2.1)

Let \( \mu_j = \mu(w_j) \).

7.3. Theorem. Let \( \ell = |(p)| \) and let \( \mu_0, \ldots, \mu_{2\ell-1} \) be as above. Let \( J_d \) be the Jacobian of the Fermat quotient curve \( C_d \). The Ekedahl–Oort structure of \( J_d[p] \) is given by
\[
\left[ \bigoplus_{\ell=0}^{\mu_1} \bigoplus_{\ell=2}^{\mu_2} \cdots \bigoplus_{\ell=2\ell-1} \right].
\]

Proof. This follows from Section 3.5 and Corollary 6.8. Indeed, for the \( BT_1 \) module \( M(J_d[p]) \) and for \( 0 \leq j < 2\ell-1 \), the subspace \( M_{j+1} \) in the canonical filtration is the span of classes in \( H_{dR}^1(C_d) \) indexed by \( a \in S \) such that \( \operatorname{Pat}(a) \leq w_j \). In particular, \( \dim(M_{j+1}) = \sum_{i=0}^{j} \mu_i \). By Definition 3.3.2, \( B_j = M_{j+1}/M_j \), so \( \dim(B_j) = \mu_j \). If \( j \) is even, then \( w_j \) ends with \( f \) so \( B_j \) is mapped isomorphically onto its image by \( F \). If \( j \) is odd, then \( w_j \) ends with \( v \) so \( B_j \) is killed by \( F \). Thus the \( \mu_j \) are the values of \( \rho(j+1) - \rho(j) \) in the canonical type of \( J_d[p] \), and they give the lengths of the runs where the elementary sequence is increasing (\( j \) even) or constant (\( j \) odd).

Let \( \mu(f-v) \) be the sum of the multiplicities of all words of length \( \ell \) starting with \( f \) and ending with \( v \); let \( \mu(-vf) \) be the sum of the multiplicities of all words of length \( \ell \) ending with \( vf \).

7.4. Theorem. The \( BT_1 \) group scheme \( J_d[p] \) has

1. \( p \)-rank equal to \( \mu(w_0) = \mu(f^\ell) \),
2. \( a \)-number equal to \( \sum_{j=1}^{2\ell-2} \mu_{2j-1} = \mu(f-v) = \mu(-vf) = \mu(-vf) \),
3. \( s_{1,1} \)-multiplicity equal to \( \mu((fv)^{\ell/2}) \) if \( \ell \) is even and to 0 if \( \ell \) is odd,
4. \( u_{1,1} \)-number equal to
\[
\mu(-v^2f^2) + \mu(-v^2fv^2) + \cdots + \mu(-v^2(fv)^{(\ell-4)/2}v^2) + \begin{cases} m((fv)^{\ell/2}) & \text{if } \ell \text{ is even,} \\ 0 & \text{if } \ell \text{ is odd.} \end{cases}
\]

Proof. The \( p \)-rank of \( J_d[p] \) is the dimension of the largest quotient of \( M = M(J_d[p]) \) on which Frobenius acts bijectively. It is clear from the Kraft description of \( M \) that this quotient is isomorphic to \( \oplus H_a \) where the sum is over those \( a \in S \) with pattern \( f^\ell \). This proves (1).

By Section 5.2, the element \( a \in S \) contributes to the \( a \)-number if and only if its word (or, equivalently, pattern) starts with \( f \) and ends with \( v \). In (7.2.1), \( j \) is odd if and only if the word starts with \( f \) and ends with \( v \), so the \( a \)-number is \( \sum \mu_{2j-1} = \mu(f-v) \). Moreover, \( \operatorname{Pat}(a) \) has the form \(-vf\) if and only if \( \operatorname{Pat}(pa) \) has the form \( f-v \), so \( \mu(f-v) = \mu(-vf) \). Also \( \operatorname{Pat}(a) \) has the form \(-vf\) if and only if \( \operatorname{Pat}(d-a) \) has the form \(-f-v \), so \( \mu(-vf) = \mu(-vf) \). This proves (2).

Similarly, (3) and (4) follow from Section 5.3 and Proposition 5.3.2.

8. \( p \)-ranks, \( a \)-numbers, and combinatorics

As before, \( C_d \) is the smooth, projective curve over \( k \) with affine model \( y^d = x(1-x) \) and \( J_d \) is its Jacobian. We continue our study of the group scheme \( J_d[p] \) by recovering some results previously proven for Fermat curves and recording some elementary combinatorics for later.
8.1. New and old. If \( d' \) divides \( d \), then there is a natural quotient morphism \( \pi : C_d \to C_{d'} \) of degree \( d/d' \), which is prime to \( p \). The induced composition

\[
J_{d'} \xrightarrow{\pi} J_d \xrightarrow{\pi} J_{d'}
\]

is multiplication by \( d/d' \) and therefore induces an isomorphism on \( J_{d'}[p] \). Thus \( J_{d'}[p] \) is a direct factor of \( J_d[p] \). Let \( J_d[p]^{old} \) be the sum of the images of \( J_{d'}[p] \) and \( J_d[p]^{new} \) be the intersection of the kernels of \( J_d[p] \to J_{d'}[p] \), in both cases over all proper divisors \( d' \) of \( d \). Then

\[
J_d[p] = J_d[p]^{new} \oplus J_d[p]^{old}.
\]

Recall that an abelian variety is said to be ordinary if its \( p \)-rank is equal to its dimension.

8.2. Proposition. Suppose \( d > 2 \) and \( p \nmid d \). Then \( J_d \) is ordinary if and only if \( d \) divides \( p - 1 \).

An analogue of Proposition 8.2 for the Fermat curve \( F_d \) was proven by Yui \cite{Yui80, Thm. 4.2} using exponential sums; (see also \cite{Gon97} Prop. 5.1) for a proof using the Cartier operator when \( d \) is prime. The “if” direction of our proposition can be deduced from Yui’s result.

Proof. By Theorem 7.4(1), \( J_d \) is ordinary if and only if \( \mu(f^\ell) = g \). If \( d \) divides \( p - 1 \), then \( \ell = 1 \), the orbits of \( \langle p \rangle \) on \( S \subset \mathbb{Z}/d\mathbb{Z} \) are singletons, and \( \mu(f) = |S_f| = g \), so \( J_d \) is ordinary.

For the converse, it suffices to show that if \( H \subset (\mathbb{Z}/d\mathbb{Z})^\times \) is a subgroup such that \( hS_v = S_v \) (equivalently \( hS_f = S_f \)) for all \( h \in H \), then \( H = \{1\} \). Suppose that \( H \) is such a subgroup, \( 1 < a < d \), and the class of \( a \) lies in \( H \). If \( a > d/2 \), then \( a \) sends \( 1 \in S_v \) to \( a \notin S_v \), a contradiction. If \( a < d/2 \), then there is an integer \( b \) in the interval \( (d/2a, d/a) \), so \( b \in S_v \), and \( d/2 < ab < d \), so the class of \( ab \) lies in \( S_f \), again a contradiction. Thus \( H = \{1\} \). \( \square \)

By Proposition 8.2, \( J_d \) is ordinary if and only if \( C_d \) is a quotient of \( C_{p-1} \) by a subgroup of \( \mu_{p-1} \).

At the opposite extreme, Proposition 8.3 below gives examples where the \( p \)-rank is zero. More generally, if \( J_d \) is supersingular then the \( p \)-rank of \( J_d \) is zero (the converse is not necessarily true for \( g \geq 3 \)). The curve \( J_d \) is supersingular when \( d \) is prime and the order of \( p \) in \( (\mathbb{Z}/d\mathbb{Z})^\times \) is even, see \cite{Gon97} Prop. 5.1. The behavior of the \( p \)-rank of \( J_d \) for arbitrary \( p \) and \( d \) seems rather erratic and we are unable to say much about it in general.

An abelian variety is superspecial if its \( a \)-number is equal to its dimension. This is equivalent to the abelian variety being isomorphic to a product of supersingular elliptic curves. If \( d = 2 \), then \( C_d \) is rational, so superspecial; we exclude this trivial case in the next result. This result shows that \( J_d \) is superspecial if and only if \( C_d \) is a quotient of \( C_{p+1} \) by a subgroup of \( \mu_{p+1} \).

8.3. Proposition. Suppose \( d > 2 \) and \( p \nmid d \). Then \( J_d \) is superspecial if and only if \( d \) divides \( p + 1 \).

An analogue of Proposition 8.3 for the Fermat curve \( F_d \) was proven by Kodama and Washio \cite{KW88, Cor. 1, p. 192} by a direct calculation of the Cartier operator. The “if” direction of our proposition can be deduced from that result.

Proof. By Theorem 7.4(2), \( J_d \) is superspecial if and only if \( \mu((f^v)^{g/2}) = g \). This is the case if and only if \( p \) exchanges \( S_f \) and \( S_v \), i.e., \( pS_f = S_c \) and \( pS_v = S_f \). Note also that if this statement holds for \( (p, d) \) then it holds for \( (p, d') \) for any divisor \( d' > 2 \) of \( d \).

Suppose that \( p \) exchanges \( S_f \) and \( S_v \). From the proof of Proposition 8.2 (applied to \( H = \langle p^2 \rangle \)), \( p \) has order 2 modulo \( d \), and the same holds for \( p \) modulo \( d' \) for any divisor \( d' \) of \( d \). Since 1 does not exchange \( S_f \) and \( S_v \), the order of \( p \) is exactly 2 modulo any \( d' > 2 \) dividing \( d \). If \( d' \) is an odd
prime power, this implies $p \equiv -1 \pmod{d'}$. More generally, let $d' = 2^e$ be the largest power of 2 dividing $d$. We claim that $p \equiv -1 \pmod{d'}$ and so $p \equiv -1 \pmod{d}$, as required. If $d' = 1$ or 2, then $p \equiv -1 \pmod{d'}$ and we are finished. If $d' = 4$, there is a unique class of order exactly 2 modulo $d'$, namely $-1$, and again $p \equiv -1 \pmod{d'}$. Finally, if $e > 2$, then there are three elements of order exactly 2 modulo $d'$, but only one of them reduces to an element of order 2 modulo $2^{e-1}$, namely $-1$. Again we find $p \equiv -1 \pmod{d'}$, and this completes the proof.

In contrast to the case of $p$-ranks, we are able to give a simple formula for $a$-numbers. Part (1) below appeared in [EP13] Thm. 1.3. An analogue of parts (1), (3), and (4) below for $F_q$ was proven by Montanucci and Speziali [MS18] by a direct calculation of the Cartier operator.

8.4. Proposition. Suppose $d > 2$ and $p$ does not divide $d$.

(1) If $p = 2$, the $a$-number of $J_d$ is $\frac{d-1}{4}$ if $d \equiv 1 \pmod{4}$ and is $\frac{d+1}{4}$ if $d \equiv 3 \pmod{4}$.

(2) If $p$ is odd, the $a$-number of $J_d$ is

$$\sum_{j=1}^{(p-1)/2} \left( \left\lfloor \frac{2jd}{2p} \right\rfloor - \left\lfloor \frac{(2j-1)d}{2p} \right\rfloor \right) = \frac{(p-1)}{p} \frac{d}{4} - \sum_{j=1}^{(p-1)/2} \left( \left\langle \frac{2jd}{2p} \right\rangle - \left\langle \frac{(2j-1)d}{2p} \right\rangle \right).$$

Here $\langle \cdot \rangle$ denotes the fractional part.

(3) If $p$ is odd and $d \equiv \pm 1 \pmod{2p}$, then the $a$-number of $J_d$ is

$$\frac{(p-1)}{p} \frac{(d \pm 1)}{4}.$$

(4) If $p$ is odd and $d \equiv p \pm 1 \pmod{2p}$, then the $a$-number of $J_d$ is

$$\frac{(p-1)}{p} \frac{(d \pm (p-1))}{4}.$$

If $p$ is large and $d$ is large with respect to $p$, Proposition 8.4 says that the $a$-number is close to $\frac{(p-1)}{p} \frac{d}{4}$ which is close to $g/2$. When $p$ is odd, the second expression in part (2) shows that the difference is $< (p-1)/2$ in absolute value. Numerical experiments suggest that

$$\left| a(J_d) - \frac{(p-1)}{p} \frac{d}{4} \right| \leq \frac{(p-1)^2}{4p},$$

and part (4) shows that this upper bound is achieved when $d \equiv p \pm 1 \pmod{2p}$.

Proof. According to Theorem 7.4 the $a$-number of $J_d[p]$ is $\mu(-fv)$, the number of elements $a \in S$ such that the pattern of $a$ ends with $fv$. These are precisely the elements with $a \in S_v$ and $pa \in S_f$, and we may count them using archimedean considerations.

More precisely, $a \in S_v$ means $0 < a < d/2$ and $pa \in S_f$ means that (the least positive residue of) $pa$ satisfies $d/2 < pa < d$. If $p = 2$, these two conditions are true when $d/4 < a < d/2$ and there are $\left[ d/2 \right] - \left[ d/4 \right]$ such integers, which proves part (1).

For part (2), the same line of reasoning shows that the elements of $S$ which contribute to the $a$-number are represented by integers satisfying one of the inequalities

$$\frac{d}{2p} < a < \frac{2d}{2p}, \quad \frac{3d}{2p} < a < \frac{4d}{2p}, \quad \ldots, \quad \frac{(p-2)d}{2p} < a < \frac{(p-1)d}{2p},$$

so that
and the number of such integers is the left hand side of the displayed equation in part (2). The equality in the displayed equation in part (2) is immediate from the definitions of $\lfloor . \rfloor$ and $\langle . \rangle$.

Parts (3) and (4) follow from part (2) and an exercise which we leave to the reader. \(\square\)

8.5. **Counting lemmas.** The following combinatorial results will be used later in this section.

Fix \(d\) and let \(\ell\) be the order of \(\langle p \rangle \subset (\mathbb{Z}/d\mathbb{Z})^\times\). Given a pattern of length \(\ell\), say \(w = u_{\ell-1} \cdots u_0\), we say that \(0 \leq j < \ell - 1\) is a break of \(w\) if \(u_{j+1} \neq u_j\), and we say \(j = \ell - 1\) is a break of \(w\) if \(u_0 \neq u_{\ell-1}\). If \(k\) is the number of breaks of \(w\), then \(k\) is even and \(0 \leq k \leq \ell\). Moreover, a pattern \(w\) is determined by its set of breaks and by its last letter \(u_0\), and there are \(2\binom{\ell}{k}\) words of length \(\ell\) with \(k\) breaks. One checks that the sum of these numbers for \(0 \leq k \leq \ell\) equals \(2^\ell\).

We also consider “self-dual” words of length \(\ell = 2\lambda\), i.e., words of the form \(w^e \cdot w\) where \(w\) has length \(\lambda\). Such a word is determined by its last half \(w\), and we may encode \(w\) by specifying its last letter and its “breaks” as above, ignoring the last potential break: If \(w = u_{\lambda-1} \cdots u_0\) we say that \(0 \leq j < \lambda - 1\) is a break if \(u_{j+1} \neq u_j\). (We ignore \(j = \lambda - 1\) because whether or not \(u_{\lambda-1}\) is a break of \(w^e \cdot w\) is already determined by the other data.) There are \(2\binom{\lambda-1}{k}\) words \(w\) with \(k\) breaks, and again one checks that the sum of these numbers for \(0 \leq k \leq \lambda - 1\) is \(2^\lambda\).

The next lemma will provide a check when we compute the multiplicities of given patterns.

8.5.1. **Lemma.**

\begin{align*}
E_\ell &:= 2 \sum_{k=0 \atop k \text{ even}}^{\ell} \binom{\ell}{k} \left( \frac{p-1}{2} \right)^k \left( \frac{p+1}{2} \right)^{\ell-k} = p^\ell + 1, \\
O_\ell &:= 2 \sum_{k=0 \atop k \text{ odd}}^{\ell} \binom{\ell}{k} \left( \frac{p-1}{2} \right)^k \left( \frac{p+1}{2} \right)^{\ell-k} = p^\ell - 1. \\
2 \sum_{k=0 \atop k \text{ even}}^{\lambda-1} \binom{\lambda-1}{k} \left( \frac{p-1}{2} \right)^{k-1} \left( \frac{p+1}{2} \right)^{\lambda+1-k} + 2 \sum_{k=0 \atop k \text{ odd}}^{\lambda-1} \binom{\lambda-1}{k} \left( \frac{p-1}{2} \right)^k \left( \frac{p+1}{2} \right)^{\lambda-k} & = p^\lambda - 1.
\end{align*}

**Proof.** From the definitions, it is clear that \(E_1 = p + 1\) and \(O_1 = p - 1\). We proceed by induction. Using the identity \(\binom{\ell+1}{k} = \binom{\ell}{k} + \binom{\ell}{k-1}\), we find that
\[
\frac{p+1}{2} E_\ell + \frac{p-1}{2} O_\ell = E_{\ell+1},
\]
and
\[
\frac{p-1}{2} E_\ell + \frac{p+1}{2} O_\ell = O_{\ell+1}.
\]
If parts (1) and (2) of the lemma are known up to \(\ell\), then we find
\[
E_{\ell+1} = \frac{p+1}{2} (p^\ell + 1) + \frac{p-1}{2} (p^\ell - 1) = p^{\ell+1} + 1.
\]
and
\[ O_{\ell+1} = \frac{p-1}{2} (p^{\ell} + 1) + \frac{p+1}{2} (p^{\ell} - 1) = p^{\ell+1} - 1. \]

This proves parts (1) and (2). For part (3), one checks that the displayed quantity is
\[ \frac{p-1}{2} E_{\lambda-1} + \frac{p+1}{2} O_{\lambda-1} = \frac{p-1}{2} (p^{\lambda-1} + 1) + \frac{p+1}{2} (p^{\lambda-1} - 1) = p^\lambda - 1. \]

\[ \square \]

Fix an integer \( \ell \geq 1 \) and list all words of length \( \ell \) which begin with \( f \) in lexicographic order:
\[ w_0 = f^\ell, w_1 = f^{\ell-1} v, w_2 = f^{\ell-2} v f, \ldots, w_{2^{\ell-1}-1} = f v^{\ell-1}. \]

Let \( k(i) \) be the number of breaks of \( w_i \) (in the second sense used above, i.e., not looking at possible wrap-around breaks).

8.5.2. Lemma. The function \( k(i) \) has the properties: (i) \( k(0) = 0 \) and (ii) if \( 2^j \leq i < 2^{j+1}, \) then \( k(i) = k(2^{j-1} - 1 - i) + 1, \) and it is characterized by these properties. Also, \( k(i) \equiv i \pmod{2}. \)

Proof. Clearly \( k(0) = 0. \) In the list of words, \( w_i \) is the binary representation of the integer \( i \) where \( f \) stands for 0, \( v \) stands for 1, and the leftmost letters are the most significant digits. With this interpretation, we see that if \( 2^j \leq i < 2^{j+1}, \) then \( w_i \) has the form \( f^{\ell-j-1} v t, \) and \( w_{2^{j-1}-1-i} = f^{\ell-j} (t^c), \) and it is visible that \( w_i \) has one more break than \( w_{2^{j-1}-1-i} \) does. This proves the second property of \( k. \) The two properties clearly characterize \( k. \) Finally, the congruence is immediate from the recursion \( k(i) = k(2^{j-1} - 1 - i) + 1 \) and the base case \( k(0) = 0. \)

\[ \square \]

Note that the function \( i \mapsto k(i) \) is independent of \( \ell \) if \( 2^{\ell-1} > i. \) Its first few values are:

| \( i \) | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
|--------|---|---|---|---|---|---|---|---|---|---|----|----|----|----|----|----|
| \( k(i) \) | 0 | 1 | 2 | 2 | 1 | 2 | 3 | 2 | 1 | 2 | 3 | 4 | 3 | 2 | 3 | 2 | 1 |

9. \( J_d[p] \) IN THE “ENCOMPASSING” CASE

In this section, we fix \( d = p^\ell - 1, \) which we call the “encompassing” case. The reason is that if \( p \nmid d', \) then \( d' \) divides \( p^\ell - 1 \) for some \( \ell, \) and so \( J_{d'}[p] \) appears in \( J_{p^{\ell-1}}[p]. \) Let \( S = S_f \cup S_v \) and \( \pi \) be defined as before.

9.1. \( p \)-adic digits. Elements \( a \in S \) correspond to \( p \)-adic expansions
\[ a = a_0 + a_1 p + \cdots + a_{\ell-1} p^{\ell-1}. \quad (9.1.1) \]

where \( a_i \in \{0, \ldots, p - 1\} \) and we exclude the following cases: all \( a_i = 0 \) (when \( a = 0); \) all \( a_i = p - 1 \) (when \( a = d); \) and, if \( p \) is odd, the case all \( a_i = (p - 1)/2 \) (when \( a = d/2). \) Multiplication by \( p \) corresponds to permuting the digits cyclically.
9.2. **Multiplicities.** Given \( a \in S \), let \( w_a = \text{Pat}(a) = u_{\ell-1} \cdots u_0 \) be the pattern of \( a \). Let \((9.1.1)\) be the \( p \)-adic expansion of \( a \). Then \( a \in S_v \) (meaning \( u_0 = v \)) if and only if \( a < d/2 \). This is true if and only if

\[
\begin{align*}
a_{\ell-1} &< (p - 1)/2, \\
a_{\ell-1} &= (p - 1)/2 \quad \text{and} \quad a_{\ell-2} < (p - 1)/2, \\
a_{\ell-1} &= a_{\ell-2} = (p - 1)/2 \quad \text{and} \quad a_{\ell-3} < (p - 1)/2, \\
&\vdots
\end{align*}
\]

In other words, the condition is that the first \( p \)-adic digit to the left of \( a_{\ell-1} \) (inclusive) which is not \((p - 1)/2\) is in fact less than \((p - 1)/2\).

Similarly \( a \in S_f \) (meaning \( u_0 = f \)) if and only if \( a > d/2 \). This is true if and only if the first \( p \)-adic digit to the left of \( a_{\ell-1} \) (inclusive) which is not \((p - 1)/2\) is in fact greater than \((p - 1)/2\). The other letters \( u_j \) of \( w_a \) are determined similarly by looking at the \( p \)-adic digits of \( a \) to the left of \( a_{\ell-1-j} \). (Finding the first digit \( \neq (p - 1)/2 \) may require wrapping around.)

For example, if \( \ell = 4 \) and \( p > 3 \), then \( \text{Pat}(a) = ffvf \) when

\[
a = (p - 1)/2 + (p - 2)p + 0p^2 + (p - 2)p^3.
\]

The following proposition records the “multiplicities” of each pattern.

9.3. **Proposition.** Let \( d = p^\ell - 1 \) and define \( S = S_f \cup S_v \) as usual.

1. For \( w \in \mathcal{W}_\ell \), write \( \mu(w) \) for the number of elements \( a \in S \) with \( \text{Pat}(a) = w \). Then

\[
\mu(f^{\ell}) = \mu(v^{\ell}) = \begin{cases} 
\left(\frac{p+1}{2}\right)^{\ell} - 2 & \text{if } p \text{ is odd}, \\
0 & \text{if } p = 2.
\end{cases}
\]

If \( w \) has \( k > 0 \) breaks, then

\[
\mu(w) = \begin{cases} 
\left(\frac{p-1}{2}\right)^k \left(\frac{p+1}{2}\right)^{\ell-k} & \text{if } p \text{ is odd}, \\
1 & \text{if } p = 2.
\end{cases}
\]

2. More generally, if \( p = 2 \) and \( t \) is a word of length \( \ell' \leq \ell \), then

\[
\mu(-t) = \begin{cases} 
2^{\ell-\ell'} & \text{if } t \neq f^{\ell'} \text{ and } t \neq v^{\ell'}, \\
2^{\ell-\ell'} - 1 & \text{if } t = f^{\ell'} \text{ or } t = v^{\ell'}.
\end{cases}
\]

3. If \( p > 2 \), then

\[
\mu(-f^{e}) = \mu(-v^{e}) = \left(\frac{p+1}{2}\right)^{\ell} - 2,
\]

\[
\mu(-v^{e_1} \cdots f^{e_1}) = \mu(-f^{e_1} \cdots v^{e_1}) = \left(\frac{p+1}{2}\right)^{\sum e_j - k} \left(\frac{p-1}{2}\right)^{k-1} \left( \frac{p^{\ell+1-\sum e_j} - 1}{2} \right),
\]

and

\[
\mu(-f^{e_1} v^{e_1} \cdots f^{e_1}) = \mu(-v^{e_1} f^{e_1} \cdots v^{e_1}) = \left(\frac{p+1}{2}\right)^{p\sum e_j - k-1} \left(\frac{p-1}{2}\right)^{k} \left( \frac{p^{\ell+1-\sum e_j} + 1}{2} \right).
\]
Proof. If \( p = 2 \), then \((p - 1)/2\) is not a valid binary digit, so each letter of \( w_a \) is determined by the corresponding binary digit of \( a \) (with \( 1 \mapsto f \) and \( 0 \mapsto v \)). The cases with all digits being 0 or 1 (namely \( v^\ell \) or \( f^\ell \)) are disallowed. This confirms the case \( p = 2 \) of the proposition.

If \( p > 2 \), we first note that the number of elements \( a \in S \) whose pattern is \( f^\ell \) or \( v^\ell \) is \( (p+1)\ell/2 - 2 \). Indeed, for \( f^\ell \) (resp. \( v^\ell \)), we may choose each \( a_j \) freely with \((p - 1)/2 < a_j \leq p - 1 \) (resp. \( 0 \leq a_j \leq (p - 1)/2 \)) except that we may not take them all to be \((p - 1)/2 \) nor all \( p - 1 \) (resp. 0). This gives part (1) of the proposition for \( p \) odd.

Finally, assume that \( p > 2 \) and \( w \) is a pattern with breaks. Then the inequalities just before the proposition show that an element \( a \) with pattern \( w \) should have \( a_j \) satisfying:

\[
\begin{align*}
 a_j &\leq (p - 1)/2 \quad \text{if } u_{t-1-j} = v \text{ and } j \text{ is not a break of } w, \\
 a_j &< (p - 1)/2 \quad \text{if } u_{t-1-j} = v \text{ and } j \text{ is a break of } w, \\
 a_j &\geq (p - 1)/2 \quad \text{if } u_{t-1-j} = f \text{ and } j \text{ is not a break of } w, \\
 a_j &> (p - 1)/2 \quad \text{if } u_{t-1-j} = f \text{ and } j \text{ is a break of } w .
\end{align*}
\]

The count displayed in the statement of the proposition is then immediate. \( \square \)

As a check, note that to specify a word there are two choices for the last letter and \( \binom{\ell}{k} \) choices for the breaks (which must be even in number). If \( p > 2 \), then by Lemma \[8.5.1\](1), the sum of the multiplicities over all words is \( |S| \) because

\[
2 \sum_{k=0}^{p-1} \binom{\ell}{k} \left( \frac{p-1}{2} \right)^k \left( \frac{p+1}{2} \right)^{-k} - 4 = p^\ell - 3 = |S| .
\]

Part (1) of the next result is implied by \[EPT3\] Thm. 1.3.

9.4. Theorem. Let \( d = p^\ell - 1 \). Let \( J_d \) be the Jacobian of the smooth, projective curve \( C_d \) with affine model \( y^d = x(1 - x) \).

(1) If \( p = 2 \), then the Ekedahl–Oort structure of \( J_d \) has the form

\[
[0, 1, 1, 2, 2, \ldots, 2^{\ell-2} - 1, 2^\ell - 2 - 1].
\]

(2) If \( p > 2 \), then the Ekedahl–Oort structure of \( J_d \) has the form

\[
[\mu_0 \rightarrow \mu_1 \rightarrow \cdots \rightarrow \mu_{2\ell-1-1}]
\]

where \( \mu_0 = (\ell+1)^\ell/2 - 2 \) and for \( 1 \leq i \leq 2^{\ell-1} - 1 \), letting \( k(i) \) be the function in Lemma \[8.5.2\]

\[
\mu_i = \begin{cases} 
\left( \frac{p-1}{2} \right)^{k(i)} \left( \frac{p+1}{2} \right)^{\ell-k(i)} & \text{if } i \text{ is even,} \\
\left( \frac{p-1}{2} \right)^{k(i)+1} \left( \frac{p+1}{2} \right)^{-k(i)-1} & \text{if } i \text{ is odd.}
\end{cases}
\]

Proof. This follows immediately from Theorem \[7.3\] the calculation of multiplicities in Proposition \[9.3\] and the evaluation of the number of breaks in Lemma \[8.5.2\] (One should note that the \( k \) appearing in Proposition \[9.3\] for \( w_i \) is \( k(i) \) if \( i \) is even, and it is \( k(i) + 1 \) if \( i \) is odd.) \( \square \)

Here are some examples for \( p > 2 \):

- If \( \ell = 1 \), \( C_d \) has genus \( (p - 3)/2 \) and is ordinary. The list of words is just \( f \). The elementary sequence has one segment of length \( (p - 3)/2 \) which is strictly increasing:

\[
[\mathcal{S}(p-3)/2] = [1, 2, \ldots, (p - 3)/2].
\]
• If $\ell = 2$, the list of words is $ff, f v$. The elementary sequence has an increasing section of length $(p + 1)^2/4 - 2$ and a constant section of length $(p - 1)^2/4$, i.e., it is:
$$\begin{array}{l}
\nearrow^{(p+1)^2/4-2} \searrow^{(p-1)^2/4}
\end{array} = [1, 2, \ldots, (p + 1)^2/4 - 2, \ldots, (p + 1)^2/4 - 2].$$
• If $\ell = 3$, the list of words is $f^3, f^2 v, f v f, f v v$. The elementary sequence has four segments and, letting $m = (p + 1)^3/8 - 2$ and $n = (p + 1)(p - 1)^2/8$, has the form
$$\begin{array}{l}
\nearrow^m \searrow^n \nearrow^n \searrow^n
\end{array}.$$
• If $\ell = 4$, the word $f^3 v$ occurs; this is the smallest example whose $BT_1$ group scheme was not previously known to occur as a factor of the $p$-torsion of a Jacobian, for all $p$.

When $p = 2$, part (1) of the next result follows from the Deuring–Shafarevic formula \cite{Sub75} and part (2) can be found in \cite{EP13} Thm. 1.3.

9.5. Proposition. Let $\ell$ be a positive integer and let $d = p^\ell - 1 > 2$. Then $J_d[p]$ has:
(1) $p$-rank equal to
$$\begin{cases}
\frac{(p+1)^\ell}{2} - 2 & \text{if } p > 2, \\
0 & \text{if } p = 2,
\end{cases}$$
(2) $a$-number equal to
$$\begin{cases}
\frac{p-1}{2} \frac{p^\ell - 1}{2} & \text{if } p > 2, \\
2^{\ell-2} & \text{if } p = 2 \text{ (and } \ell > 1),
\end{cases}$$
(3) $s_{1,1}$-multiplicity equal to
$$\begin{cases}
0 & \text{if } \ell \text{ is odd} \\
\left(\frac{p^2-1}{4}\right)^{\ell/2} & \text{if } \ell \text{ is even and } p > 2, \\
1 & \text{if } \ell \text{ is even and } p = 2,
\end{cases}$$
(4) and $u_{1,1}$-number equal to
$$\begin{cases}
\sum_{j=0}^{\lceil(\ell-4)/2\rceil} \left(\frac{p+1}{2}\right)^2 \left(\frac{p-1}{2}\right)^{2j+1} \left(\frac{p^\ell - 3 - 2j - 1}{2}\right) + \left(\frac{p^2-1}{4}\right)^{\ell/2} & \text{if } \ell \text{ is odd, if } p > 2, \\
0 & \text{if } \ell \text{ is even, if } p > 2,
\end{cases}$$
$$
\sum_{j=0}^{\lceil(\ell-4)/2\rceil} 2^{\ell-4-2j} + \begin{cases}
0 & \text{if } \ell \text{ is odd, if } p = 2,
1 & \text{if } \ell \text{ is even, if } p = 2.
\end{cases}
$$

Proof. This follows from Theorem 7.4 using the multiplicities furnished by Proposition 9.3. \qed

10. $J_d[p]$ in the Hermitian case

Fix an integer $\lambda \geq 1$ and let $d = p^\lambda + 1$. In this case, the Fermat curve of degree $d$ is isomorphic (over $\overline{F}_p$) to the Hermitian curve $y_1^{q+1} = x_1^q + x_1$ where $q = p^\lambda$. It is well-known that the Hermitian curve is supersingular and its Ekedahl–Oort structure was studied in \cite{PW15}. In this case, the curve $C_d$ is a quotient of the Hermitian curve, thus also supersingular.
10.1. $p$-adic digits. Let $S = \mathbb{Z}/d\mathbb{Z} \setminus \{0\}$ if $p = 2$ and $S = \mathbb{Z}/d\mathbb{Z} \setminus \{0, d/2\}$ if $p > 2$. Let $S_v = \{b \in S \mid 0 < b < d/2\}$ and $S_f = \{b \in S \mid d/2 < b < d\}$. Let $\pi : S \to S$ be induced by multiplication by $p$. Let 
\[ T = \{(b_1, \ldots, b_\lambda) \mid 0 \leq b_j \leq p - 1 \text{ and not all } b_j = (p-1)/2 \text{ if } p \text{ is odd}\} .\]

There is a bijection $T \to S$ given by 
\[ (b_1, \ldots, b_\lambda) \mapsto b = 1 + \sum_{j=1}^{\lambda} b_j p^{j-1} .\]

Under this bijection, the permutation $\pi$ is given by 
\[ (b_1, \ldots, b_\lambda) \mapsto (p - 1 - b_\lambda, b_1, \ldots, b_{\lambda-1}) .\]

An element $b$ belongs to $S_v$ if and only if $b < (p^\lambda + 1)/2$. This is true if and only if, in the tuple, 
the entry $b_j$ with largest $j$ such that $b_j \neq (p-1)/2$ has the property that $b_j < (p-1)/2$.

10.2. Multiplicities. The multiplicative order of $p$ modulo $d$ is $\ell = 2\lambda$. We define a map $\text{Pat}' : S \to \mathcal{W}_\lambda$ as follows: Given $b \in S$, the pattern $\text{Pat}'(b)$ of $b$ is the word $w = u_{\lambda-1} \cdots u_0$ given by 
\[ u_j = \begin{cases} f & \text{if } p^j b \in S_f, \\
 & \\
v & \text{if } p^j b \in S_v. \end{cases} \]

(The notation $\text{Pat}'$ is used to distinguish this from the pattern in the encompassing case.) If the word for $b$ has length $\ell$ (the maximum length), then it is $\text{Pat}'(b)^c \cdot \text{Pat}'(b)$ (where the $c$ stands for the complementary word). For any $b \in S$, the word for $b$ has a power with length $\ell$ and this power equals $\text{Pat}'(b)^c \cdot \text{Pat}'(b)$. Note that since $p^\lambda = -1 \pmod{d}$, the word of $b$ is “self-dual”, i.e., of the form $t^{\ell} t$, for every $b \in S$.

For a word $w$ of length $\lambda$, let $\mu'(w)$ be the number of elements $b \in S$ with $\text{Pat}'(b) = w$. For a word $t$ of length $\leq \lambda$, let $\mu'(-t)$ be the number of elements $b \in S$ with $\text{Pat}'(b) = t' \cdot t$ for some $t'$, in other words, the number of $b$ with pattern ending in $t$.

10.3. Proposition.

(1) If $p = 2$ and $w$ is a word of length $\lambda$, then $\mu'(w) = 1$.

(2) If $p = 2$ and $t$ is a word of length $\lambda' \leq \lambda$, then $\mu'(-t) = p^{\lambda - \lambda'}$.

(3) Suppose $e_1, \ldots, e_\lambda$ are positive integers with $\sum e_i = \lambda$.

If $p > 2$ and $k$ is odd, then 
\[ \mu'(v^{e_k} v^{e_{k-1}} \cdots v^{e_1}) = \mu'(v^{e_k} v^{e_{k-1}} \cdots v^{e_1}) = \left(\frac{p+1}{2}\right)^{-k} \left(\frac{p-1}{2}\right)^k , \]

and if $p > 2$ and $k$ is even, then 
\[ \mu'(v^{e_k} \cdots v^{e_1}) = \mu'(f^{e_k} \cdots v^{e_1}) = \left(\frac{p+1}{2}\right)^{\lambda+1-k} \left(\frac{p-1}{2}\right)^{k-1} . \]
(4) More generally, given integers \( e_1, \ldots, e_k > 0 \), let \( \lambda' = \sum e_i \) and suppose \( \lambda' \leq \lambda \). If \( p > 2 \) and \( k \) is odd, and \( t \) is a word of the form \( t = f^{e_k} v^{e_{k-1}} \cdots v^{e_2} f^{e_1} \), then

\[
\mu'(t) = \mu'(t^-) = \left( \frac{p+1}{2} \right)^{\lambda'-k} \left( \frac{p-1}{2} \right)^{k-1} \left( \frac{p^{\lambda+1-\lambda'} - 1}{2} \right),
\]

and if \( p > 2 \) and \( k \) is even, and \( t \) is a word of the form \( t = v^{e_k} \cdots f^{e_1} \), then

\[
\mu'(t) = \mu'(t^-) = \left( \frac{p+1}{2} \right)^{\lambda'-k} \left( \frac{p-1}{2} \right)^{k-1} \left( \frac{p^{\lambda+1-\lambda'} + 1}{2} \right).
\]

Part (3) of Proposition 10.3 contradicts [PW15, Lemma 4.3], which we believe is in error.

**Proof.** If \( p = 2 \), then \( \text{Pat}' \) induces a bijection between \( S \) and \( \mathcal{W}_\lambda \), and parts (1) and (2) are then immediate. Now assume \( p > 2 \) and note that part (3) is a special case of part (4), so we will prove the latter. It is clear that \( \mu'(t) = \mu'(t^-) \).

Suppose \( k \) is odd and \( t = f^{e_k} v^{e_{k-1}} \cdots v^{e_2} f^{e_1} \). Write \( f^{e_k} \cdots f^{e_1} = u_{\lambda-1} \cdots u_0 \) with \( u_j \in \{ f, v \} \). Then \( b \in S \) has pattern \(-t\) if and only if the \( p\)-adic digits \( (b_1, \ldots, b_\lambda) \) satisfy, for \( \lambda + 1 - \lambda' < j \leq \lambda \),

\[
\begin{align*}
&b_j \leq (p-1)/2 \quad \text{if } u_{\lambda-j} = v \text{ and } j \text{ is not a break of } t, \\
&b_j < (p-1)/2 \quad \text{if } u_{\lambda-j} = v \text{ and } j \text{ is a break of } t, \\
&b_j \geq (p-1)/2 \quad \text{if } u_{\lambda-j} = f \text{ and } j \text{ is not a break of } t, \\
&b_j > (p-1)/2 \quad \text{if } u_{\lambda-j} = f \text{ and } j \text{ is a break of } t,
\end{align*}
\]

and the number corresponding to the tuple \( \beta = (b_1, \ldots, b_{\lambda+1-\lambda'}) \) is large, namely

\[
p^{\lambda+1-\lambda'} + 1 > 1 + \sum_{j=1}^{\lambda+1-\lambda'} b_j p^{j-1} > \frac{p^{\lambda+1-\lambda'} + 1}{2}.
\]

So there are \( (p^{\lambda+1-\lambda'} - 1)/2 \) choices for \( \beta \). Taking the product with the number of possibilities for \( b_j \) for \( \lambda + 1 - \lambda' < j \leq \lambda \) yields the quantity in the statement.

Similarly, if \( k \) is even and \( t = v^{e_k} \cdots f^{e_1} = u_{\lambda-1} \cdots u_0 \) with \( u_j \in \{ f, v \} \), then \( b \in S \) has pattern \(-t\) if and only if the \( p\)-adic digits \( (b_1, \ldots, b_\lambda) \) satisfy (10.3.1), for \( \lambda + 1 - \lambda' < j \leq \lambda \) and the number corresponding to the tuple \( \beta \) is small, namely

\[
0 < 1 + \sum_{j=1}^{\lambda+1-\lambda'} b_j p^{j-1} \leq \frac{p^{\lambda+1-\lambda'} + 1}{2}.
\]

So there are \( (p^{\lambda+1-\lambda'} + 1)/2 \) choices for \( \beta \). Again, taking the product with the number of possibilities for \( b_j \) for \( \lambda + 1 - \lambda' < j \leq \lambda \) yields the quantity in the statement. \( \square \)

As a check, to specify a word there are two choices for the last letter and \( \binom{k}{2} \) choices for the breaks (which may be even or odd in number). If \( p > 2 \), then by Lemma 8.5.1(3), the sum of the multiplicities over all words is \( p^\lambda - 1 = |S| \) as it should be.

The case \( p = 2 \) of the next result is a special case of [EP13 Thm. 1.3].

10.4. **Theorem.** Let \( d = p^\lambda + 1 \) and let \( J_d \) be the Jacobian of the smooth, projective curve \( C_d \) defined by \( y^d = x(1-x) \).
(1) If \( p = 2 \), then the Ekedahl–Oort structure of \( J_d \) is \([0], [0, 1], [0, 1, 1, 2] \) for \( \lambda = 1, 2, 3 \) respectively, and for \( \lambda > 3 \), it has the form
\[
[0, 1, 1, 2, 2, \ldots, 2^{\lambda-2} - 1, 2^{\lambda-2} - 1, 2^{\lambda-2}].
\]

(2) If \( p > 2 \), then the Ekedahl–Oort structure of \( J_d \) has the form
\[
[\rightarrow \mu'_0 \rightarrow \mu'_1 \ldots \rightarrow \mu'_{2^{\lambda-1}-1}]
\]
where, letting \( k(i) \) be the function described in Lemma 8.5.2.

\[
\mu'_i = \begin{cases} 
\left( \frac{p+1}{2} \right)^{\lambda-k(i)-1} \left( \frac{p-1}{2} \right)^{k(i)+1} & \text{if } i \text{ is even,} \\
\left( \frac{p+1}{2} \right)^{\lambda-k(i)} \left( \frac{p-1}{2} \right)^{k(i)} & \text{if } i \text{ is odd.}
\end{cases}
\]

Proof. This follows immediately from Theorem 7.3, the calculation of multiplicities in Proposition 10.3, and the evaluation of the number of breaks in Lemma 8.5.2. (One has to note that the \( k \) in Proposition 10.3 for \( w_i \) is \( k(i) + 1 \).)

Here are some examples for \( p > 2 \):

- If \( \ell = 1 \), the curve \( C_d \) has genus \( (p-1)/2 \) and is superspecial: the list of relevant words is just \( *v \), and the elementary sequence is one constant segment of length \( (p-1)/2 \):
  \[
  [\rightarrow (p-1)/2] = [0, \ldots, 0].
  \]

- If \( \ell = 2 \), the list of relevant words is \( *fvv, *vvf \), and the elementary sequence has a constant section of length \( (p^2-1)/4 \) and an increasing section of length \( (p^2-1)/4 \):
  \[
  [\rightarrow (p^2-1)/4 \rightarrow (p^2-1)/4] = [0, \ldots, 0, 1, 2, \ldots, (p^2-1)/4].
  \]

- If \( \ell = 3 \), the list of relevant words is \( *f^3v^3, *f^2v^3f, *f^3v^3, f^3v^3f^2 \) and the elementary sequence has four segments and has the form

\[
[\rightarrow m \rightarrow n \rightarrow m],
\]
where \( m = (p+1)^2(p-1)/8 \) and \( n = (p-1)^3/8 \).

10.5. Proposition. Let \( \lambda \) be a positive integer and let \( d = p^\lambda + 1 \). Then \( J_d[p] \) has:

(1) \( p \)-rank 0,
(2) \( a \)-number equal to
\[
\begin{cases} 
1 & \text{if } p = 2 \text{ and } \lambda = 1, \\
2^{\lambda-2} & \text{if } p = 2 \text{ and } \lambda > 1, \\
\left( \frac{p-1}{2} \right) \left( \frac{p^{\lambda-1}+1}{2} \right) & \text{if } p > 2,
\end{cases}
\]
(3) \( s_{1,1} \)-multiplicity equal to
\[
\begin{cases} 
0 & \text{if } \lambda \text{ is even,} \\
1 & \text{if } \lambda \text{ is odd and } p = 2, \\
\left( \frac{p-1}{2} \right)^\lambda & \text{if } \lambda \text{ is odd and } p > 2,
\end{cases}
\]
(4) and $u_{1,1}$-number equal to

$$
\begin{cases}
0 & \text{if } \lambda \text{ is even}, \\
1 & \text{if } \lambda \text{ is odd},
\end{cases}
\sum_{j=0}^{(\lambda-4)/2} 2^{\lambda-4-2j} + \begin{cases}
0 & \text{if } \lambda = 1, \\
1 & \text{if } \lambda > 1,
\end{cases}
$$

if $p = 2$, and

$$
\begin{cases}
0 & \text{if } \lambda \text{ is even}, \\
\left(\frac{p-1}{2}\right)^\lambda & \text{if } \lambda \text{ is odd},
\end{cases}
\sum_{j=0}^{(\lambda-4)/2} \left(\frac{p+1}{2}\right)^2 \left(\frac{p-1}{2}\right)^{2j+1} \left(\frac{p^{\lambda-3-2j} + 1}{2}\right)
+ \begin{cases}
0 & \text{if } \lambda = 1, \\
\left(\frac{p+1}{2}\right)^2 \left(\frac{p-1}{2}\right)^{\lambda-2} & \text{if } \lambda > 1 \text{ and odd}, \\
\left(\frac{p+1}{2}\right)^2 \left(\frac{p-1}{2}\right)^{\lambda-1} & \text{if } \lambda \text{ even},
\end{cases}
$$

if $p > 2$.

We note that part (1) may also be obtained from the Deuring–Shafarevich formula \([Sub75]\), and the analogue of the $a$-number calculation in part (2) for the Fermat curve of degree $p^\lambda + 1$ is given in \([Gro90]\ Prop. 14.10]\).

**Proof.** This follows from Theorem \([7,4]\) using the multiplicities in Proposition \([10,3]\), although there are some subtleties. The $p$-rank is 0, since for every $b \in S$, the associated word is self-dual, i.e., of the form $t^c \cdot t$, so no there is no $b$ whose word is a power of $f$.

By Theorem \([7,4]\) the $a$-number is the number of $b$ whose word ends with $v f$. If $\lambda = 1$, such a $b$ has pattern $f$, and by Proposition \([10,3]\) the number of such $b$ is 1 if $p = 2$ and $(p-1)/2$ if $p > 2$. If $\lambda > 1$, then the word of $b$ ends with $v f$ if and only if the pattern of $b$ ends with $v f$, and by Proposition \([10,3]\) the multiplicity is $p^{\lambda-2}$ if $p = 2$ and $(p-1)(p^{\lambda-1} + 1)/4$ if $p > 2$, as desired.

By Theorem \([7,4]\) the $s_{1,1}$-multiplicity of $J_d[p]$ is the number of $b$ with $\text{Pat}'(b)^c \cdot \text{Pat}'(b) = (f v)^\lambda$. This is of the form $t^c \cdot t$ only when $\lambda$ is odd. When $\lambda$ is odd, Proposition \([10,3]\) gives the multiplicity as 1 when $p = 2$ and as $((p-1)/2)^\lambda$ when $p > 2$, as desired.

The $u_{1,1}$-number has contributions from: the $s_{1,1}$-multiplicity (the multiplicity of $f v \cdots f$ when $\lambda$ is odd); the multiplicity of the word $v^2(fv)^j f^2$ when $\lambda \geq 2j + 4$; and the multiplicity of $(fv)^{(\lambda-2)/2} f^2$ when $\lambda$ is even or $v(fv)^{(\lambda-3)/2} f^2$ when $\lambda > 1$ is odd. \(\square\)

### 11. Universality

In this section, we use the “encompassing case” in Section \([9]\) to prove the following result:

**11.1. Theorem** (Universality).

1. Every $BT_1$ group scheme over $k$ appears as a direct factor of $J\mathcal{C}[p]$ for a suitable (and explicit) curve $C$ defined over $\mathbb{F}_p$.

2. Every polarized $BT_1$ group scheme over $k$ appears as a direct factor (with pairing) of $J\mathcal{C}[p]$ for a suitable (and explicit) curve $C$ defined over $\mathbb{F}_p$.

We will prove Theorem \([11,1]\) in the rest of this section.

By Corollary \([4,2,3]\) a self-dual $BT_1$ group scheme has a unique polarization up to isomorphism, so part (2) follows from part (1). The following result establishes Theorem \([11,1]\) for $p > 2$ and for $p = 2$ except for factors of $\mathbb{Z}/2\mathbb{Z}$ and $\mu_2$, with the curve $C$ being a quotient of a Fermat curve.
11.2. **Theorem.** If \( p > 2 \), then every \( BT_1 \) group scheme \( G \) appears in \( J_d[p] \) for a suitable \( d \), and in fact it appears to arbitrarily high multiplicity. If \( p = 2 \), the same holds if and only if \( G \) has no factors of \( \mathbb{Z}/2\mathbb{Z} \) and \( \mu_2 \). The same conclusions hold for \( J_{F_d'} \), the Jacobian of the Fermat curve of degree \( d \).

**Proof.** By the Kraft classification, an indecomposable \( BT_1 \) group scheme \( G \) corresponds to a primitive cyclic word \( \overline{w} \) on the two-letter alphabet \( \{f, v\} \). Choose a representative word \( w \) for \( \overline{w} \). If \( w \) has length \( \ell \), take \( d = p^\ell - 1 \). Then the multiplicity of \( G \) in \( J_d \) equals the number \( a \in S \) whose word is \( w \). We computed the multiplicity in Proposition 9.3 and it is positive except when \( p = 2 \) and \( w = f \) or \( w = v \). Thus \( G \) is a factor of \( J_d[p] \) unless \( p = 2 \) and \( w = f \) or \( w = v \).

To assure high multiplicities, we may replace \( w \) with \( w^e \) and repeat the argument, concluding that \( w \) appears in some \( J_d[p] \) with multiplicity a non-zero multiple of \( e \). Direct sums of powers of indecomposables also appear because if \( G_1 \) appears in \( J_d[p] \) and \( G_2 \) appears in \( J_d'[p] \) and \( G_1 \) and \( G_2 \) have no common indecomposable factors, then \( G_1 \oplus G_2 \) appears in \( J_{d'}[p] \). The non-appearence of \( \mu_2 \) and \( \mathbb{Z}/2\mathbb{Z} \) follows from the multiplicities in Proposition 9.3 and the fact that every \( G \) that appears in some \( J_d[p] \) appears when \( d \) has the form \( d = p^\ell - 1 \) for a suitable \( \ell \).

The existence results for \( J_d[p] \) imply those for \( J_{F_d'}[p] \) because there is an inclusion \( J_d[p] \hookrightarrow J_{F_d'}[p] \) (see Section 6.2). For the non-appearence of \( \mathbb{Z}/2\mathbb{Z} \) and \( \mu_2 \) in \( J_{F_d'}[p] \) when \( p = 2 \), we must show that no orbit of \( \langle 2 \rangle \) on the set \( S \) in Remark 6.9 lies entirely in \( S_v \); by symmetry, this implies that no orbit lies entirey in \( S_f \) either. Suppose \( (a, b) \in S_v \), i.e., \( a + b < d \) (where the inequality implicitly uses the least positive residues of \( a \) and \( b \)). Replacing \( (a, b) \) by \( (a', b') = (2^e a, 2^e b) \) for some \( e \geq 0 \), we may assume that \( d/2 < a' + b' < d \). Then \( d < 2a' + 2b' < 2d \), so \( (2a', 2b') \in S_f \). Thus every orbit of \( \langle 2 \rangle \) on \( S \) that meets \( S_v \) also meets \( S_f \), completing the proof. \( \square \)

11.3. **The case \( p = 2 \).** To finish the proof of Theorem 11.1, it remains to treat the case where \( p = 2 \) and \( G \) is a \( BT_1 \) group scheme over \( k \) with factors of \( \mathbb{Z}/2\mathbb{Z} \) or \( \mu_2 \). Write

\[
G \cong (\mathbb{Z}/2\mathbb{Z})^f \oplus (\mu_2)^{f_2} \oplus G',
\]

where \( G' \) is a \( BT_1 \) group scheme with no factors of \( \mathbb{Z}/2\mathbb{Z} \) and no factors of \( \mu_2 \). By Theorem 11.2, \( G' \) appears in \( J_d[p] \) for a suitable (odd) value of \( d \). We choose one such value of \( d \).

Let \( r \) be an odd positive integer and let \( X_r \) be the smooth, projective curve over \( k \) defined by

\[
X_r : \quad (x^2 - x)(z^r - 1) = 1.
\]

One computes that \( X_r \) has genus \( r - 1 \), and by [Sub75, Prop. 3.2], it is ordinary, i.e.,

\[
J_{X_r'}[p] \cong (\mathbb{Z}/2\mathbb{Z} \oplus \mu_2)^{r-1}.
\]

Choose \( r \geq \max\{f, f_2\} + 1 \) and odd, and define \( C \) as the fiber product of the degree \( r \) projection \( X_r \to \mathbb{P}^1 \) and the degree \( d \) projection \( C_d \to \mathbb{P}^1 \). Since \( d \) and \( r \) are odd, \( J_{X_r'}[p] \) and \( J_d[p] \) are direct factors of \( J_C'[p] \). They are disjoint since \( X_r \) is ordinary, whereas \( C_d \) has \( p \)-rank zero. Thus \( G \subset (\mathbb{Z}/2\mathbb{Z} \oplus \mu_2)^{r-1} \oplus G' \) is a direct factor of \( J_C'[p] \), completing the proof of Theorem 11.1.

11.4. **Another approach to \( p = 2 \).** We sketch another approach to adding factors of \( \mathbb{Z}/2\mathbb{Z} \) and \( \mu_2 \) to \( G' \) which is closely related to [Pri19, §6] and the material discussed in Section 6.10.

Assume \( p = 2 \) and \( G \) is given as in (11.3.1), where \( G' \) is a \( BT_1 \) group scheme with no factors of \( \mathbb{Z}/2\mathbb{Z} \) or \( \mu_2 \). Choose \( d \) such that \( G' \) appears as a direct factor of \( J_d[p] \). We will sketch an argument showing that, for arbitrarily large \( f \), there are cyclic covers \( C \) of \( \mathbb{P}^1 \) such that

\[
J_C[p] \cong (\mathbb{Z}/2\mathbb{Z} \oplus \mu_2)^r \oplus J_d[p].
\]
This allows us to deduce Theorem 11.1 from Theorem 11.2 except that \( C \) is no longer explicit, and we have no control over its field of definition other than that it is a finite field. To do this, we repeatedly apply [LMPT, Cor. 4.7] to show that for any multiple \( f \) of \( d - 1 \), there is a cyclic cover \( \tilde{C} \) of \( \mathbb{P}^1 \) of degree \( d \) such that the Newton polygon of \( \tilde{C} \) is that of \( C_d \) together with \( f \) extra segments of slopes 0 and 1. The curve \( \tilde{C} \) is obtained by deforming \( C_d \) to a cover with \( f/(d - 1) \) new pairs of branch points. Because \( C_d \) has complex multiplication, when the new branch points are "sufficiently general", the Newton polygon of \( \tilde{C} \) is minimal among those for cyclic covers with the same monodromy data, and so \( \tilde{C} \) is \( \mu \)-ordinary (as in Section 6.10).

Since \( \tilde{C} \) is \( \mu \)-ordinary, by [Moo04, §1], the isomorphism class of \( J_{\tilde{C}}[p] \) is uniquely determined and may be computed from the monodromy data. (We refer to [LMPT19, §4.5] for examples of such computations.) In this case, \( J_{\tilde{C}}[p] \cong (\mathbb{Z}/2\mathbb{Z} \oplus \mu_2)^f \oplus J_d[p] \), and thus \( G \) appears in \( J_{\tilde{C}}[p] \).

References

[AP15] J. D. Achter and R. Pries, Superspecial rank of supersingular abelian varieties and Jacobians, J. Théor. Nombres Bordeaux 27 (2015), 605–624.

[Dum95] N. Dummigan, The determinants of certain Mordell-Weil lattices, Amer. J. Math. 117 (1995), 1409–1429.

[EP13] A. Elkin and R. Pries, Ekedahl-Oort strata of hyperelliptic curves in characteristic 2, Algebra Number Theory 7 (2013), 507–532.

[Fon77] J.-M. Fontaine, Groupes \( p \)-divisibles sur les corps locaux, Société Mathématique de France, Paris, 1977. Astérisque, No. 47-48.

[Gon97] J. González, Hasse-Witt matrices for the Fermat curves of prime degree, Tohoku Math. J. (2) 49 (1997), 149–163.

[Gro90] B. H. Gross, Group representations and lattices, J. Amer. Math. Soc. 3 (1990), 929–960.

[Ill85] L. Illusie, Déformations de groupes de Barsotti-Tate (d’après A. Grothendieck), 1985, pp. 151–198. Seminar on arithmetic bundles: the Mordell conjecture (Paris, 1983/84).

[Kra75] H. Kraft, Kommutative algebraische \( p \)-Gruppen (mit anwendungen auf \( p \)-divisible Gruppen und Abelsche Varietäten), 1975. Manuscript, Universität Bonn, 86 pp.

[KW88] T. Kodama and T. Washio, Hasse-Witt matrices of Fermat curves, Manuscripta Math. 60 (1988), 185–195.

[LMPT19] W. Li, E. Mantovan, R. Pries, and Y. Tang, Newton polygons arising from special families of cyclic covers of the projective line, Res. Number Theory 5 (2019), Paper No. 12, 31 pp.

[LMPT] Newton polygon stratification of the Torelli locus in PEL-type Shimura varieties. Preprint, arXiv:1811.00604.

[Moo01] B. Moonen, Group schemes with additional structures and Weyl group cosets, Moduli of abelian varieties (Texel Island, 1999), 2001, pp. 255–298.

[Moo04] Serre-Tate theory for moduli spaces of PEL type, Ann. Sci. École Norm. Sup. (4) 37 (2004), 223–269.

[MS18] M. Montanucci and P. Speziali, The \( \alpha \)-numbers of Fermat and Hurwitz curves, J. Pure Appl. Algebra 222 (2018), 477–488.

[Oda69] T. Oda, The first de Rham cohomology group and Dieudonné modules, Ann. Sci. École Norm. Sup. (4) 2 (1969), 63–135.

[Oor01] F. Oort, A stratification of a moduli space of abelian varieties, Moduli of abelian varieties (Texel Island, 1999), 2001, pp. 345–416.

[Pri19] R. Pries, Current results on Newton polygons of curves, Open problems in arithmetic algebraic geometry, 2019, pp. 179–207.

[PW15] R. Pries and C. Weir, The Ekedahl-Oort type of Jacobians of Hermitian curves, Asian J. Math. 19 (2015), 845–869.

[Sub75] D. Subrao, The \( p \)-rank of Artin-Schreier curves, Manuscripta Math. 16 (1975), 169–193.

[Ulm19] D. Ulmer, On the Brauer-Siegel ratio for abelian varieties over function fields, Algebra Number Theory 13 (2019), 1069–1120.

[Ulm91] D. L. Ulmer, \( p \)-descent in characteristic \( p \), Duke Math. J. 62 (1991), 237–265.
[vdG99] G. van der Geer, *Cycles on the moduli space of abelian varieties*, Moduli of curves and abelian varieties, 1999, pp. 65–89.

[Wei76] A. Weil, *Sur les périodes des intégrales abéliennes*, Comm. Pure Appl. Math. 29 (1976), 813–819.

[Yui80] N. Yui, *On the Jacobian variety of the Fermat curve*, J. Algebra 65 (1980), 1–35.

Department of Mathematics, Colorado State University, Fort Collins, CO 80523 USA

*Email address: pries@math.colostate.edu*

Department of Mathematics, University of Arizona, Tucson, AZ 85721 USA

*Email address: ulmer@math.arizona.edu*