CLUSTERS AND SEMISTABLE MODELS OF HYPERELLIPTIC CURVES IN THE WILD CASE

LEONARDO FIORE AND JEFFREY YELTON

Abstract. Given a hyperelliptic curve $Y : y^2 = f(x)$ over a complete discrete valuation field K with algebraically closed residue field, we define a semistable model of Y over the ring of integers of a finite extension of K which we call the *relatively stable model* $\mathcal{Y}^{\mathrm{rst}}$ of Y. Our goal is to compute this model as the normalization in $K(Y)$ of the model $\mathcal{X}^{(\mathrm{rst})}$ of $X := \mathbb{P}^1_K$ given by the quotient of $\mathcal{Y}^{\mathrm{rst}}$ by the hyperelliptic involution. In the case of residue characteristic $p \neq 2$, the components of the special fiber $(\mathcal{X}^{(\mathrm{rst})})_s$ correspond precisely to the non-singleton *clusters* of roots of the defining polynomial f, i.e. the subsets of roots of f which are closer to each other than to the other roots of f with respect to the induced discrete valuation on the splitting field. This relationship, however, is far less straightforward in the $p = 2$ case, which is our main focus. We describe recent results showing that, when $p = 2$, for each cluster containing an *even* number of roots of f, there are 0, 1, or 2 components of $(\mathcal{Y}^{\mathrm{rst}})_s$ corresponding to it, and we discuss a direct method of finding and describing them. We also define a polynomial $F(T) \in K[T]$ whose roots allow us to find the components of $(\mathcal{Y}^{\mathrm{rst}})_s$ which are not connected to even-cardinality clusters. We finish by describing how to find relatively stable models of hyperelliptic curves in genus 1 and 2.

1. Introduction

The focus of this paper is to investigate the reduction types of hyperelliptic curves over discrete valuation fields. Given a complete discrete valuation field K of characteristic different from 2 with algebraically closed residue field, our starting point is to consider a *hyperelliptic curve* Y over K; that is, Y/K is a smooth projective curve of positive genus admitting a degree-2 morphism onto the projective line $\mathbb{P}^1_K$.

This paper is concerned with constructing a semistable model of a given hyperelliptic curve Y/K and understanding the structure of the special fiber of a semistable model of Y. As this problem is already entirely understood in the case that the residue characteristic is not 2 and the procedure in that case can be described entirely in terms of the distances between the branch points with respect to the p-adic metric on K, our primary focus will be on the case where the residue characteristic is 2. The increased complexity of the problem for this case arises from the fact that a hyperelliptic curve comes with a degree-2 map to the projective line: the fact that this degree is the same as the residue characteristic implies that we are in a "wild setting". Problems involving reduction of curves in the "wild case", in which one studies semistable models of curves with a degree-p map to the projective line over residue characteristic p, have been investigated in a number of works in recent decades (see §1.4 below), but mainly in the situation where the branch points of the map $Y \to \mathbb{P}^1_K$ are p-adically equidistant. In this article, we will consider general hyperelliptic curves over residue characteristic 2, with a particular focus on the relationship between the combinatorial data of how the branch points are "clustered" and the structure of the special fiber of a semistable model. The results presented here come from the authors' preprint [8], in which they are rigorously proved.

1.1. **Hyperelliptic curves.** In this subsection, the only assumption about the ground field K that we adhere to is that the characteristic is different from 2. In this situation, it is well known that an affine chart for a hyperelliptic curve Y/K of genus $g \geq 1$ is given by an equation of the form

$$y^2 = f(x) = c\prod_{i=1}^{d}(x - a_i), \tag{1}$$

where $f(x) \in K[x]$ is a polynomial of degree $d \in \{2g+1, 2g+2\}$ that does not have multiple roots, $c \in K^\times$ is the leading coefficient of $f(x)$, and the elements $a_i \in \bar{K}$ are the roots of f. We call f the *defining polynomial* of (this chart of) the hyperelliptic curve Y. The degree-2 morphism of Y onto the projective line is given simply by the coordinate function x; this morphism is branched precisely at each of the roots of f as well as, in the case that $d = 2g + 1$ (in other words, when f has odd degree), at the point ∞. After applying an appropriate automorphism of the projective line (i.e. a suitable change of coordinate) which moves one of the branch points to ∞, we obtain an equation of the form in (1) with $d = 2g + 1$; we adhere to this assumption about f throughout this paper. Our aim will be showing how to explicitly form semistable models of Y over finite extensions of K. We more fully explain various aspects of the problem below.

1.2. **Semistable models of curves.** For the rest of this paper, we assume that the ground field, in addition to having characteristic different from 2, is a complete discrete valuation field; we denote its ring of integers by $R \subset K$ and its residue field by k, which we assume to be algebraically closed.

A *model* of C over R', where R' is the ring of integers of some finite extension $K' \supseteq K$, is a normal projective flat R'-scheme whose generic fiber is isomorphic to C over K'. In this paper, given a smooth projective geometrically connected curve C/K, we use the same letter in curly font to denote a model $\mathcal{C}/R'$ defined over the ring of integers R' of some finite extension K'/K, and we denote its special fiber by $\mathcal{C}_s/k$.

We say that a model $\mathcal{C}$ is *semistable* if its special fiber $\mathcal{C}_s$ is a reduced k-curve with at worst nodes as singularities. The following groundbreaking theorem was proved by Deligne and Mumford in [5] and then through independent arguments by Artin and Winters in [1] (see also [11, Section 10.4] for a detailed explanation of the arguments in Artin-Winters).

Theorem 1.1. *Every smooth projective geometrically connected curve C over K achieves semistable reduction over a finite extension $K' \supseteq K$, i.e. C admits a semistable model $\mathcal{C}^{\mathrm{ss}}$ over R', where R' is the ring of integers in K'.*

The above result is not constructive and does not tell us how to find a semistable model $\mathcal{C}^{\mathrm{ss}}$ or exactly how large an extension $K' \supseteq K$ is needed in order to define it. It moreover does not specify, for a given curve C/K, anything about the structure of the special fiber $(\mathcal{C}^{\mathrm{ss}})_s$. It is therefore natural to ask whether there is any general method by which we may construct a semistable model $\mathcal{Y}^{\mathrm{ss}}$ of a hyperelliptic curve Y/K defined by an equation of the form in (1).

1.3. **Special fibers of semistable models of curves.** In this paper, we are interested not only in how to construct a semistable model $\mathcal{Y}^{\mathrm{ss}}$ of a hyperelliptic curve Y, but also in how certain characteristics of the defining polynomial may determine the *structure* of the special fiber of such a semistable model. The special fiber $(\mathcal{Y}^{\mathrm{ss}})_s$ of a semistable model $\mathcal{Y}^{\mathrm{ss}}$ of a curve Y/K by definition consists of reduced components which meet each other only at nodes. Each node, viewed as a point in $\mathcal{Y}^{\mathrm{ss}}$, has a *thickness*, *i.e.* an integer $\mu > 0$ such that the completed local ring at the node is isomorphic to $R[[t_1, t_2]]/(t_1t_2 - a)$ for some $a \in R$, with $v(a) = \mu$. The structure of the special fiber $(\mathcal{Y}^{\mathrm{ss}})_s$ can be described entirely in terms of the set of its irreducible components, the genus of the normalization of each of these components, the data of which components intersect which others at how many nodes, and the thicknesses of the nodes. The sum of the genera of the normalizations of the irreducible components is known as the *abelian rank* of $(\mathcal{Y}^{\mathrm{ss}})_s$, while the number of loops in the configuration of components and their intersections (i.e. the number of loops in the dual graph of $(\mathcal{Y}^{\mathrm{ss}})_s$) is known as the *toric rank* of $(\mathcal{Y}^{\mathrm{ss}})_s$. The property of being semistable implies that the sum of these two ranks equals the genus of Y.

Replacing a semistable model $\mathcal{Y}^{\mathrm{ss}}$ of Y over R' with another semistable model of Y over R'' (where R' and R'' are the ring of integers of possibly different extensions of K) does not affect its abelian or toric rank, and therefore these ranks are intrinsic to the curve Y itself and particularly interesting to determine (meanwhile, the thicknesses of the nodes change in a predictable manner between semistable models over different extensions of R).

In order to obtain and make precise our main results about special fibers of semistable models of hyperelliptic curves, it will be necessary to define a particular semistable model of a hyperelliptic curve Y, which we call the *relatively stable model* and denote by $\mathcal{Y}^{\mathrm{rst}}$, using given the degree-2 cover $Y \to X := \mathbb{P}^1_K$. This allows us to determine and describe the model $\mathcal{Y}^{\mathrm{rst}}$ entirely in terms of an associated model $\mathcal{X}^{(\mathrm{rst})}$ of the projective line X. Any model of the projective line over a discrete valuation field K can, in turn, be described in terms of closed discs in K, reducing much of our problem to a determination of collection of discs in K which yield the relatively stable model $\mathcal{Y}^{\mathrm{rst}}/R'$ of a given hyperelliptic curve Y/K. In particular, we define the *valid discs* (in Definition 2.5) as crucial subset of this collection. All of this is discussed in §2 below. Then, in §3, we define the *cluster data* associated to a hyperelliptic curve and directly relate this to the valid discs; in the $p = 2$ setting, this is our main result and is given as Theorem 3.4. The following two sections are devoted to our method of finding all valid discs associated to a hyperelliptic curve: §4 focuses on finding valid discs which contain (clusters of) roots of f, while §5 focuses on finding centers of valid discs which do not contain roots of f. In §6 we briefly present our results which characterize the structure of the special fiber of the relatively stable model (in particular, its toric rank). Finally, in §7, we present an application of our results to the relatively stable model of a hyperelliptic curve of genus 2.

1.4. **Related results in other works.** A hyperelliptic curve is a special case of a *superelliptic curve*, i.e. a curve defined by an equation of the form $y^n = f(x)$ for some $n \geq 2$. There have been a number of works discussing semistable models of superelliptic curves. When the exponent n in the equation for a superelliptic curve is not divisible by the residue characteristic p, the process of constructing a semistable model is relatively straightforward and is provided in [2, §3], [3, §4], [6, §4, 5] (for hyperelliptic curves, using the language of clusters), and [9] (for hyperelliptic curves, using the language of stable marked curves), as well as earlier works.

The existing results for the wild case of semistable reduction of superelliptic curves, i.e. when the defining equation is of the form $y^p = f(x)$ where p is the residue characteristic, have been far more limited. To the best of our knowledge, investigations into this case began with Coleman, who in [4] outlined an algorithm for changing coordinates in such a way that the defining equation is converted to a form whose reduction over the residue field does not describe a curve which is an inseparable degree-p cover of the line; when $p = 2$, this is more or less equivalent to our notion of *part-square decompositions* which will be introduced in §4. This idea is further developed by Lehr and Matignon in [12] and later in [10] (among several other works). The wild case is also discussed in [3, §4], in which several examples are computed and interpreted in terms of rigid analytic geometry; the working of these examples is mainly done through clever guessing rather than a direct algorithm, however.

Moreover, in each of [12] and [10], a polynomial over the ground field is defined whose roots are the centers of all discs which give rise to components of the special fiber; these polynomials (the *p-dérivée* in [12, Définition 2.4.1] and the *monodromy polynomial* in [10, Definition 3.4]) are quite distinct but each is defined similarly and plays a similar role to our polynomial $F(T) \in K[T]$ given in Definition 5.1 below, whose roots in the geometrically equidistant case certainly provide centers of all the valid discs.

Our work differs from the prior research discussed above in that our major focus is on the relationship between clusters of roots and the structure of the special fiber of a semistable model of a hyperelliptic curve when the residue characteristic is 2; to the best of our knowledge, the only specific case in terms of cluster data which has been investigated where equidistant geometry is not assumed is in the recent article [7], which treats a case involving an even number of roots clustering in pairs.

We finish this subsection by remarking that our paper does not prioritize much focus towards describing the finite extension of K over which we are constructing our semistable (relatively stable) model of Y or determining the minimal extension of K over which Y achieves semistable reduction (although in building $\mathcal{Y}^{\mathrm{rst}}$, we try to be economical in the extension of K required). However, as our results are constructive, it is fairly straightforward to compute the (necessarily totally ramified) extension K'/K over which $\mathcal{Y}^{\mathrm{rst}}$ is defined. In general, the extension K'/K is obtained from (possibly) a sequence of quadratic extensions of the subfield $K'' \subset K'$ over which the associated model $\mathcal{X}^{(\mathrm{rst})}$ of the projective line is defined using changes of coordinates $x = \alpha_i + \beta_i x_i$ for some elements $\alpha_i \in \bar{K}$ and $\beta_i \in \bar{K}^{\times}$. In practice, each scaling element β_i may be chosen to be any element of a prescribed valuation, while a given translating element α_i may be chosen to be a root of f (and thus already in the splitting field) when the corresponding valid disc contains roots of f; it is only in the case where there are valid discs not containing roots of f that one may have to choose α_i to be a root of the (generally high-degree) polynomial $F(T) \in K[T]$ defined in §5 below.

1.5. **Notational conventions.** Given an algebraic extension K'/K, an element $\alpha \in K'$ and a rational number $b \in v(K')$, we denote by $D_{\alpha,b} \subset K'$ the disc centered at α with (logarithmic) radius b, *i.e.* we let $D_{\alpha,b} = \{z \in K' \mid v(z - \alpha) \geq b\}$.

Given a polynomial $f(z) \in K'[z]$, we write $v(f)$ for its *Gauss valuation*, *i.e.* writing $f(z) = c_0 + c_1 z + \cdots + c_d z^d$, we let $v(f) = \min_i v(c_i)$.

1.6. **Acknowledgements.** The authors would like to thank Fabrizio Andreatta for proposing that the first author, as work for his Masters thesis, join the early stages of the research project of the second author, as well as for providing guidance and helpful discussions to the first author throughout his research work in the Masters program.

2. Relatively stable models of Galois covers of the projective line

In this section, we assume $Y \to X$ to be a Galois covering map of curves over K with Galois group G and that Y has positive genus, keeping in mind that the situation we care about is when Y is a hyperelliptic curve with its degree-2 map to $X := \mathbb{P}^1_K$. Given a curve Y equipped with such a covering map, we are able to define a useful generalization of the stable model of Y (which, unlike the stable model, is also defined when $g(Y) = 1$), which we provide in the following subsection.

2.1. **Definition of the relatively stable model.** In order to define this particular semistable model of Y, we need to make some preliminary definitions.

Definition 2.1. *Let $\mathcal{Y}$ be a semistable model of Y which is also acted upon by G, and let V be a component of the special fiber $\mathcal{Y}_s$.*

(a) We say that V is (-1)-line *if we have $V \cong \mathbb{P}^1_k$ and if there is exactly 1 node of $\mathcal{Y}_s$ lying on V.*
(b) Write $\tilde{V}$ for the normalization of V. We say that V is a horizontal (-2)-curve *if we have $\tilde{V} \cong \mathbb{P}^1_k$, if there are exactly 2 points $P_1, P_2 \in \tilde{V}(k)$ lying over the intersection of V with the set of nodes of $\mathcal{Y}_s$, and if G fixes both P_1 and P_2.*

Definition 2.2. *A model of $\mathcal{Y}$ of Y is said to be* relatively stable *with respect to the Galois cover $Y \to X$ if it satisfies the following properties:*

(i) $\mathcal{Y}$ is semistable and also acted upon by G, so that the cover $Y \to X$ extends to a map $\mathcal{Y} \to \mathcal{X} := \mathcal{Y}/G$;
(ii) no component of the special fiber $\mathcal{Y}_s$ is a (-1)-line or a (-2)-curve; and
(iii) there is no smooth point of the special fiber $\mathcal{X}_s$ whose inverse image in $\mathcal{Y}_s$ is a node.

Remark 2.3. We record the following facts about semistable models of a covering $Y \to X$.

(a) A relatively stable model $\mathcal{Y}^{\mathrm{rst}}$ of Y, if it exists, is unique, and remains relatively stable under any extension of the ground field K.

(b) Suppose that $g(Y) \geq 2$ or that $g(X) = 0$. Then the curve Y has a relatively stable model over the ground field K if and only if it has semistable reduction over K. Moreover, if $Y \to X$ is the trivial map, then the relatively stable model is just the stable model of Y.

(c) Suppose that Y/K has a relatively stable model $\mathcal{Y}^{\mathrm{rst}}/R$, and write $\mathcal{X}^{(\mathrm{rst})} = \mathcal{Y}^{\mathrm{rst}}/G$ so that the covering $Y \to X$ extends to a map $\mathcal{Y}^{\mathrm{rst}} \to \mathcal{X}^{(\mathrm{rst})}$. Since no node in $(\mathcal{Y}^{\mathrm{rst}})_s$ maps to a smooth point in $(\mathcal{X}^{(\mathrm{rst})})_s$ by property (iii) in Definition 2.2, we have that the inverse image of the set of nodes in $(\mathcal{X}^{(\mathrm{rst})})_s$ coincides with the set of nodes in $(\mathcal{Y}^{\mathrm{rst}})_s$.

We end this subsection by pointing out a simple but important property of the relatively stable model, which will allow us to build relatively stable models of hyperelliptic curves directly from models of the projective line.

Proposition 2.4. *Suppose that Y has a relatively stable model $\mathcal{Y}^{\mathrm{rst}}$ and define the map $\mathcal{Y}^{\mathrm{rst}} \to \mathcal{X}^{(\mathrm{rst})}$ as above. If a component V of the special fiber $(\mathcal{X}^{(\mathrm{rst})})_s$ is a smooth k-curve, then its inverse image in $(\mathcal{Y}^{\mathrm{rst}})_s$ also is. In particular, if all irreducible components of $(\mathcal{X}^{(\mathrm{rst})})_s$ are smooth k-curves, then the same is true of the irreducible components of $(\mathcal{X}^{(\mathrm{rst})})_s$.*

2.2. **Discs and models of the projective line.** We now resume our assumption that $Y \to X := \mathbb{P}^1_K$ is a hyperelliptic curve given by the equation in (1) with its degree-2 map to the projective line given by the standard coordinate x. We assume that Y has a relatively stable model $\mathcal{Y}^{\mathrm{rst}}$ over the ground field K and define the model $\mathcal{X}^{(\mathrm{rst})}$ of the projective line X as above.

Fortunately, it is very straightforward to describe models of the projective line, which we do in the following manner. Each *smooth* model of $X = \mathbb{P}^1_K$ over the ring of integers R is isomorphic to $\mathbb{P}^1_R$ and is determined by a choice of coordinate $x_{\alpha,\beta} := \beta^{-1}(x - \alpha)$. Any two such models are isomorphic over R if and only if we have $v(\alpha_2 - \alpha_1) \geq v(\beta_1) = v(\beta_2)$, where x_{α_1,β_1} and x_{α_2,β_2} are the corresponding coordinates. This condition is equivalent to the equality of the discs $D_{\alpha_i,v(\beta_i)}$ for $i = 1, 2$, so the smooth models of X over R are in bijection with the set of discs in K. Any model $\mathcal{X}$ of X over R is then the compositum of a finite collection of smooth models of X over R (that is, $\mathcal{X}$ is the minimal model of X which dominates each smooth model in the collection), and this corresponds to a finite collection of such discs in K.

Given any disc $D \subset K$, we write $\mathcal{X}_D/R$ for the corresponding smooth model of X and, for any model $\mathcal{X}$ of X, write $\mathcal{X}_D \leq \mathcal{X}$ if $\mathcal{X}$ dominates $\mathcal{X}_D$. Proposition 2.4 along with the definition of the relatively stable model implies that knowledge of the model $\mathcal{X}^{(\mathrm{rst})}$ of the projective line determines the relatively stable model $\mathcal{Y}^{\mathrm{rst}}$: for each (smooth) component of the special fiber $(\mathcal{X}^{(\mathrm{rst})})_s$, which corresponds to a smooth model $\mathcal{X}_D$ of X, its normalization $\mathcal{Y}_D$ in $K(Y)$ corresponds to a component (resp. 2 components) of $(\mathcal{Y}^{\mathrm{rst}})_s$ if $(\mathcal{Y}_D)_s$ is an irreducible curve (resp. 2 disjoint lines each isomorphic to $\mathbb{P}^1_k$). Our task of computing the relatively stable model of Y therefore essentially comes down to identifying which discs $D \subset K$ satisfy $\mathcal{X}_D \leq \mathcal{X}^{(\mathrm{rst})}$.

We now define a term which we will use throughout the rest of the paper in order to refer to components of the relatively stable model of a hyperelliptic curve.

Definition 2.5. *With the above notation and assumptions, the disc $D \subseteq K$ is a* valid disc *if it satisfies $\mathcal{X}_D \leq \mathcal{X}^{(\mathrm{rst})}$ and if the quadratic cover $(\mathcal{Y}_D)_s \to (\mathcal{X}_D)_s$ is separable.*

We note that our notion of *valid disc* differs from the one in [6], although in both cases valid discs are used to build a particular semistable model of Y with desired properties.

Remark 2.6. The discs D such that $\mathcal{X}_D \leq \mathcal{X}^{(\mathrm{rst})}$ which are not valid discs are in some sense inessential in the context of understanding the invariants of the special fibers of semistable models of Y listed in §1.3: for such a disc D, the strict transform of $\mathcal{X}_D$ in $(\mathcal{Y}^{\mathrm{rst}})_s$ is just a projective line only which needs to be included as a component of $(\mathcal{Y}^{\mathrm{rst}})_s$ because contracting it would result in a singular point with ≥ 3 branches, which would violate semistability. Our focus for the rest of the paper is therefore on the valid discs D and the components of $(\mathcal{Y}^{\mathrm{rst}})_s$ which correspond to them.

3. The cluster data associated to a hyperelliptic curve

In the $p \neq 2$ setting, the valid discs as defined in the previous section correspond precisely to certain subsets of roots of the defining polynomial of our hyperelliptic curve Y which we call the *clusters* associated to Y, and our primary focus in this paper is relating valid discs to clusters in the $p = 2$ setting.

3.1. **Cluster data.** One naïve attempt to construct a semistable model of a hyperelliptic curve Y is to use the scheme $\mathcal{Y}$ over the ring of integers R using the equation in (1) (as long as the coefficients of this equation lie in R, which one can always assume after appropriately scaling x and y). While one may verify through elementary computations that this model $\mathcal{Y}$ always fails to be semistable over residue characteristic 2, in the $p \neq 2$ setting, the semistability of $\mathcal{Y}$ can be easily seen to be equivalent to the conditions that

(i) the roots of f are *equidistant* (i.e. the valuations of the difference between the roots are all equal) so that $\mathcal{Y}_s$ is smooth; and
(ii) the roots of f are equidistant except for certain pairs of roots of f which are closer to each other with respect to the discrete valuation of K (so that each pair maps to a root of multiplicity 2 of the reduced polynomial $\bar{f}$ and produces a node of $\mathcal{Y}_s$).

This suggests that the construction of semistable models of Y over finite extensions of K is closely related to the data of which subsets of roots of the defining polynomial f "cluster" together with respect to the distance function on the splitting field of f. This notion is made precise in [6] by defining the *cluster data* associated to a hyperelliptic curve Y over a discrete valuation field K as follows.

Definition 3.1. *Given a hyperelliptic curve Y/K defined by an equation of the form given in (1), let $\mathcal{R}$ denote the set of roots of f and assume that we have $\mathcal{R} \subset K$. A* cluster *of $\mathcal{R}$ is a nonempty subset $\mathfrak{s} \subseteq \mathcal{R}$ satisfying $\mathfrak{s} = \mathcal{R} \cap D$ for some disc $D \subset K$. The* depth $d(\mathfrak{s})$ *of a cluster $\mathfrak{s}$ is given by*

$$d(\mathfrak{s}) = \min_{z,z' \in \mathfrak{s}} \{v(z - z')\}. \tag{2}$$

The set of pairs $(\mathfrak{s}, d(\mathfrak{s}))$, where $\mathfrak{s}$ varies among all clusters of $\mathcal{R}$, is called the cluster data *of $\mathcal{R}$, or the* cluster data *associated to the hyperelliptic curve Y.*

One advantage of defining the relatively stable model of a hyperelliptic curve as we have done is that in the $p \neq 2$ setting, the components of this model are determined by the cluster data of the hyperelliptic curve in the simplest way possible.

Theorem 3.2. *In the $p \neq 2$ setting, there is a one-to-one correspondence between non-singleton clusters and valid discs, given by mapping a cluster $\mathfrak{s}$ to the minimal disc $D_{\mathfrak{s}}$ which contains it. In other words, the valid discs are precisely those discs that minimally cut out the clusters of $\mathcal{R}$.*

In the situation of Theorem 3.2, the model $\mathcal{Y}^{\mathrm{rst}}$ is formed in the following way. Given a cluster $\mathfrak{s}$ of roots of the defining polynomial f, choose an element $\alpha \in \mathfrak{s}$ and an element $\beta \in K^\times$ satisfying $v(\beta) = d(\mathfrak{s})$. Then the corresponding smooth model $\mathcal{X}_D$ of the projective line X is given by the coordinate $x_{\alpha,\beta} := \beta^{-1}(x - \alpha)$. Meanwhile, let $f_{\alpha,\beta}(x) \in K[x]$ be the polynomial such that $f_{\alpha,\beta}(x_{\alpha,\beta}) = f(x)$. After scaling y by an element of appropriate evaluation in the (unique) quadratic extension of K, from the standard equation in (1) we obtain an equation of the form $y^2 = \gamma^{-1} f_{\alpha,\beta}(x_{\alpha,\beta})$, where $\gamma \in K^\times$ is an element whose valuation equals the minimal valuation among the coefficients of the polynomial $f_{\alpha,\beta}$. This new equation defines a model $\mathcal{Y}_{D_{\mathfrak{s}}}$ of Y, defined over an extension $R' \supseteq R$ which is at most quadratic, which is the normalization of the model of

X corresponding to the disc $D_{\mathfrak{s}}$ in the function field $K'(Y)$ (where K' is the fraction field of R'). The desired semistable model $\mathcal{Y}^{\mathrm{ss}}$ is comprised of these normalizations. The idea is illustrated by the following example.

Example 3.3. Let $K = \mathbb{Q}_p^{\mathrm{unr}}$ for some $p \geq 5$ and

$$f(x) = x(x-p^3)(x-p)(x-1)(x-1+p^4)(x-2)(x-3). \tag{3}$$

The set of roots of f is $\mathcal{R} := \{0, p^3, p, 1, 1-p^4, 2, 3\}$. The *clusters* of these roots (i.e. the subsets $\mathfrak{s}$ consisting of roots which are closer to each other than they are to the roots in $\mathcal{R} \smallsetminus \mathfrak{s}$) are

$$\mathfrak{s}_0 := \mathcal{R},\ \mathfrak{s}_1 := \{0, p^3, p\},\ \mathfrak{s}_2 := \{0, p^3\},\ \mathfrak{s}_3 := \{1, 1-p^4\},$$

as well as each of the singleton subsets of $\mathcal{R}$ (which we ignore). The data of these clusters is represented by the following diagram.

$$\text{cluster data of } \mathcal{R} : \quad \boxed{\boxed{\boxed{0 \qquad p^3}_{\,\mathfrak{s}_2} \qquad p}_{\,\mathfrak{s}_1} \qquad \boxed{1 \qquad 1-p^4}_{\,\mathfrak{s}_3} \quad 2 \qquad 3}_{\,\mathfrak{s}_0}$$

Corresponding to each cluster $\mathfrak{s}_i$ is a smooth model of X given by the coordinate x_i, and these coordinates may be defined as

$$x = x_0 = px_1 = p^3 x_2 = p^4(x_3 - 1). \tag{4}$$

We define corresponding coordinates y_i by scaling y by suitable elements of $\mathbb{Q}_p(\sqrt{p})$ as

$$y = y_0 = p^{3/2} y_1 = p^{7/2} y_2 = p^4 y_3.$$

We now define corresponding models $\mathcal{Y}_i / \mathbb{Z}_p^{\mathrm{unr}}[\sqrt{p}]$ of $Y / \mathbb{Q}_p^{\mathrm{unr}}(\sqrt{p})$ for $i = 0, 1, 2, 3$, given by the below equations.

$$\begin{aligned}
&\mathcal{Y}_0 : y_0^2 = f(x) = f(x_0) \\
&\mathcal{Y}_1 : y_1^2 = p^{-3} f(x) = x_1(x_1 - p^2)(x_1 - 1)(px_1 - 1)(px_1 - 1 + p^4)(px_1 - 2)(px_1 - 3) \\
&\mathcal{Y}_2 : y_2^2 = p^{-7} f(x) = x_2(x_2 - 1)(p^2 x_2 - 1)(p^3 x_2 - 1)(p^3 x_2 - 1 + p^4)(p^3 x_2 - 1)(p^3 x_2 - 2) \\
&\mathcal{Y}_3 : y_3^2 = p^{-8} f(x) = (p^4 x_3 - 1)(p^4 x_3 - 1 - p^3)(p^4 x_3 - 1 - p)(x_3)(x_3 - 1)(p^4 x_3 - 2)
\end{aligned} \tag{5}$$

Their respective reductions (that is, their special fibers $(\mathcal{Y}_i)_s$) over the residue field $\overline{\mathbb{F}_p}$ are as follows.

$$\begin{aligned}
&(\mathcal{Y}_0)_s : y_0^2 = x_0^3 (x_0 - 1)^2 (x_0 - 2)(x_0 - 3) \\
&(\mathcal{Y}_1)_s : y_1^2 = 6x_1^2 (x_1 - 1) \\
&(\mathcal{Y}_2)_s : y_2^2 = -6x_2 (x_2 - 1) \\
&(\mathcal{Y}_3)_s : y_3^2 = 2x_3 (x_3 - 1)
\end{aligned} \tag{6}$$

The desingularizations of each of these special fibers give rise to the components of the special fiber of the desired semistable model $\mathcal{Y}^{\mathrm{ss}}$: here $(\mathcal{Y}_0)_s$ contributes a smooth component V_0 of genus 1; $(\mathcal{Y}_1)_s$ contributes a line V_1 which intersects V_0 at a single node; $(\mathcal{Y}_2)_s$ contributes a line V_2 which intersects V_1 at 2 nodes; and $(\mathcal{Y}_3)_s$ contributes a line V_3 which intersects V_0 at 2 nodes. The toric rank of $(\mathcal{Y}^{\mathrm{ss}})_s$ is 2, coming from the fact that there are exactly 2 even-cardinality clusters.

3.2. **The relationship between clusters and the relatively stable model when $p = 2$.** When the residue characteristic of K is 2, it is natural to ask whether a semistable model of Y can be constructed by a procedure governed entirely by the associated cluster data in this way, or in other words, whether there is some analog of Theorem 3.2 for working over residue characteristic 2. In short, the answer is "no", as it turns out that in the $p = 2$ setting, valid discs do not correspond in this way in a one-on-one manner with clusters, as is shown in particular by parts (a) and (b) of Theorem 3.4 below.

Theorem 3.4. *Assume all of the above set-up for a hyperelliptic curve Y/K of genus g given by an equation of the form $y^2 = f(x) \in K[x]$, where the polynomial f has degree $2g+1$, and assume that the residue characteristic of K is 2. Let $\mathcal{Y}^{\mathrm{rst}}/R'$ be the relatively stable model of Y, where R' is the ring of integers of an appropriate finite field extension $K' \supseteq K$. Let $\mathcal{R} \subset \bar{K}$ denote the set of roots of f. For any even-cardinality cluster of roots $\mathfrak{s} \subsetneq \mathcal{R}$, we write $\mathfrak{s}'$ for the minimal cluster which properly contains $\mathfrak{s}$ (which always exists since $\mathcal{R}$ itself is a cluster and has odd cardinality).*

The clusters of roots in $\mathcal{R}$ and the valid discs associated to Y are related in the following manner.

(a) Given a valid disc $D \subseteq \bar{K}$, the cardinality of $D \cap \mathcal{R}$ is even (and we may have $D \cap \mathcal{R} = \varnothing$).
(b) If a cluster $\mathfrak{s}$ has even cardinality, there are either 0, 1, or 2 valid discs $D \subseteq R'$ such that either $D \cap \mathcal{R} = \mathfrak{s}$ or D is the smallest disc containing $\mathfrak{s}'$.
(c) Let $\mathfrak{s}$ be an even-cardinality cluster of relative depth $m := \delta(\mathfrak{s})$. There exists a rational number $B_{f,\mathfrak{s}} \in \mathbb{Q}_{\geq 0}$, which is independent of the relative depth of $\mathfrak{s}$ in the sense of Remark 3.5 below, such that
   (i) if $m > B_{f,\mathfrak{s}}$, the number of valid discs as in part (b) is "2";
   (ii) if $m = B_{f,\mathfrak{s}}$, the number of valid discs as in part (b) is "1"; and
   (iii) if $m < B_{f,\mathfrak{s}}$, the number of valid discs as in part (b) is "0".
(d) Given an even-cardinality cluster $\mathfrak{s}$, the bound $B_{f,\mathfrak{s}}$ from part (d) satisfies $B_{f,\mathfrak{s}} \leq 4v(2)$. If we furthermore assume that $\mathfrak{s}$ and $\mathfrak{s}'$ each have a maximal subcluster of odd cardinality (e.g. a maximal subcluster which is a singleton), we have the inequality

$$B_{f,\mathfrak{s}} \geq \Big(\frac{2}{|\mathfrak{s}|-1} + \frac{2}{2g+1-|\mathfrak{s}|}\Big)v(2). \tag{7}$$

Remark 3.5. The rational number $B_{f,\mathfrak{s}}$ given in part (c) of the above theorem does not depend on the depth $\delta(\mathfrak{s})$ in the following sense. Given a center α of the minimal disc $D_{\mathfrak{s}}$ which contains $\mathfrak{s}$, let $\mathfrak{s}_{[\lambda]} = \{\lambda(a-\alpha)+\alpha \mid a \in \mathfrak{s}\}$ for some $\lambda \in \bar{K}^{\times}$ such that $v(\lambda) > -\delta(\mathfrak{s})$, so that $\mathfrak{s}_{[\lambda]}$ is a scaled version of $\mathfrak{s}$ and is a cluster in $\mathcal{R}_{[\lambda]} := \mathfrak{s}_{[\lambda]} \sqcup (\mathcal{R} \smallsetminus \mathfrak{s})$ with relative depth $\delta(\mathfrak{s}_{[\lambda]}) = \delta(\mathfrak{s}) + v(\lambda)$. Then we have $B_{f_{[\lambda]},\mathfrak{s}_{[\lambda]}} = B_{f,\mathfrak{s}}$. In this sense, loosely speaking, we may view $B_{f,\mathfrak{s}}$ as a sort of "threshold" for the depth of $\mathfrak{s}$ at which we obtain 1 valid disc linked to $\mathfrak{s}$ and above which we obtain 2 valid discs linked to $\mathfrak{s}$.

Theorem 3.4 can be viewed as a vast generalization of the results in [14], where the second author explicitly constructed semistable models of elliptic curves with a cluster of cardinality 2 and depth m (as well as elliptic curves with no even-cardinality clusters). The threshold for m above which there are 1 or 2 valid discs containing that cardinality-2 cluster which is found in [14] comes as the following easy corollary to the above theorem; we remark that this corollary can be deduced also from standard formulas for the j-invariant of an elliptic curve (specifically, the particular choice of power of 2 multiplied to the rest of the formula, which influences the valuation of the j-invariant in residue characteristic 2, as can be seen for instance in the Legendre case as in [13, Proposition III.1.7]).

Corollary 3.6. *Suppose that we are in the $g = 1$ case of the situation in Theorem 3.4 and that $\mathfrak{s}$ is a cluster of cardinality 2. Then we have $B_{f,\mathfrak{s}} = 4v(2)$.*

Proof. The parent cluster of $\mathfrak{s}$ (*i.e.* the minimal cluster strictly containing $\mathfrak{s}$) is $\mathfrak{s}' = \mathcal{R}$, which has cardinality 3. It is clear that both $\mathfrak{s}$ and $\mathfrak{s}'$ have a singleton child cluster (i.e., a maximal subcluster consisting of only one root). Now, Theorem 3.4(d) gives that $B_{f,\mathfrak{s}} \leq 4v(2)$ and

$$B_{f,\mathfrak{s}} \geq (\frac{2}{1} + \frac{2}{1})v(2) = 4v(2). \tag{8}$$

The equality $B_{f,\mathfrak{s}} = 4v(2)$ follows. □

Examples of computations which yield the desired model $\mathcal{Y}^{\mathrm{rst}}$ in the case that $m \leq 4v(2)$ are given as [14, Examples 2 and 3]. An example for the $m > 4v(2)$ case is given as [14, Example 9], except that there a semistable model whose special fiber has a single (nodal) component, rather than the relatively stable model $\mathcal{Y}^{\mathrm{rst}}$, is found.

4. Finding valid discs containing clusters

The standard approach to finding semistable models of hyperelliptic curves over residue characteristic 2, which was introduced by Coleman in [4], involves expressing polynomials $f_{\alpha,\beta}(x_{\alpha,\beta})$ as the sum of the square of a polynomial $q_{\alpha,\beta}(x_{\alpha,\beta})$ plus a "remainder" polynomial $\rho_{\alpha,\beta}(x_{\alpha,\beta})$, where, for chosen elements $\alpha \in \bar{K}$ and $\beta \in \bar{K}^\times$ we define the coordinate $x_{\alpha,\beta} = \beta^{-1}(x-\alpha)$ and $f_{\alpha,\beta}$ to be the "translated and scaled" polynomial such that $f_{\alpha,\beta}(x_{\alpha,\beta}) = f(x)$. Such a decomposition of $f_{\alpha,\beta}$ allows us to replace y with $\gamma^{1/2}y + q_{\alpha,\beta}(x_{\alpha,\beta})$ and write the standard equation (1) for our hyperelliptic curve Y as

$$y^2 + 2\gamma^{-1/2}q_{\alpha,\beta}y = \gamma^{-1}\rho_{\alpha,\beta}, \tag{9}$$

for some choice of scalar $\gamma \in \bar{K}^\times$ of appropriate valuation. If the element $\alpha \in \bar{K}$ and the valuation of the element $\beta \in \bar{K}^\times$, as well as the polynomials $q_{\alpha,\beta}, \rho_{\alpha,\beta} \in K(\alpha,\beta)[x_{\alpha,\beta}]$ are chosen appropriately, the coefficients in (9) are all integral and the model $\mathcal{Y}_{D_{\alpha,v(\beta)}}/R$ defined by the equation in (9) is the normalization of $\mathcal{X}_{D_{\alpha,v(\beta)}}$ in $K(Y)$; moreover, these parameters may be chosen such that $D_{\alpha,v(\beta)}$ is a valid disc, so that $\mathcal{Y}_{D_{\alpha,v(\beta)}}$ is a separable cover of $\mathcal{X}_{D_{\alpha,v(\beta)}}$ and gives rise to a component of $(\mathcal{Y}^{\mathrm{rst}})_s$. This approach inspires the term *part-square decomposition* defined below.

Definition 4.1. *Given a nonzero polynomial $h(x) \in \bar{K}[z]$, a* part-square decomposition *of h is a way of writing $h = q^2 + \rho$ for some $q(x), \rho(x) \in \bar{K}[x]$, with $\deg(q) \leq \lceil \deg(h)/2 \rceil$.*

Given a part-square decomposition $h = q^2 + \rho$, we define the rational number $t_{q,\rho} := v(\rho) - v(h) \in \mathbb{Q} \cup \{+\infty\}$.

We moreover define the following properties of a part-square decomposition $h = q^2 + \rho$.

(a) The decomposition is said to be good *either if we have $t_{q,\rho} \geq 2v(2)$ or if we have $t_{q,\rho} < 2v(2)$ and there is no decomposition $h = \tilde{q}^2 + \tilde{\rho}$ such that $t_{\tilde{q},\tilde{\rho}} > t_{q,\rho}$.*

(b) The decomposition is said to be totally odd *if ρ only consists of odd-degree terms.*

Remark 4.2. We make the following observations about part-square decompositions.

(a) Definition 4.1 forces $\deg(\rho) \leq \deg(h)$ when h has even degree and $\deg(\rho) \leq \deg(h)+1$ when h has odd degree. The definition allows q to be equal to zero.

(b) The trivial part-square decomposition $h = 0^2 + h$ has $t_{0,h} = 0$; this immediately implies that all good decompositions $h = q^2 + \rho$ satisfy $t_{q,\rho} \geq 0$.

(c) If $h = q^2 + \rho = (q')^2 + \rho'$ are two good part-square decompositions for the same nonzero polynomial h, then we have $\min\{t_{q,\rho}, 2v(2)\} = \min\{t_{q',\rho'}, 2v(2)\}$ directly from Definition 4.1.

A fairly elementary algebraic argument shows that a totally odd part-square decomposition is always good. Moreover, it turns out that a totally odd part-square decomposition of a polynomial h always exists (so that a good part-square decomposition always exists), according to the following proposition.

Proposition 4.3. *Given a nonzero polynomial $h(z) \in \bar{K}[z]$, there always exists a totally odd part-square decomposition $h = q^2 + \rho$ with $q(z), \rho(z) \in \bar{K}[z]$.*

Proof. Let the polynomial h_e be the sum of the even-degree terms of h, whose roots are easily seen to come in pairs $\pm\sqrt{\alpha_i}$ for some elements $\alpha_i \in \bar{K}$. Writing $c_0z^m + c_1z^{m-1} + \cdots + c_m = \sqrt{c}\prod_i(z + \sqrt{\alpha_i})$, where $\sqrt{c}$ is a square root of the leading coefficient c of h_e, one verifies straightforwardly that setting q to be the polynomial whose ith coefficient is c_i for even i and $\sqrt{-1}c_i$ for odd i (given some fixed square root of -1) produces a totally odd decomposition of h. $\square$

Now given any polynomial h and elements $\alpha \in \bar{K}$ and $\beta \in \bar{K}^{\times}$, write $h_{\alpha,\beta}$ for the "translated and scaled" polynomial such that $h_{\alpha,\beta}(x_{\alpha,\beta}) = h(x)$. It is almost immediate to see that if $h = q^2 + \rho$ is a part-square decomposition, then $h_{\alpha,\beta} = q_{\alpha,\beta}^2 + \rho_{\alpha,\beta}$ is also a part-square decomposition, and that in the case that $\alpha = 0$ (i.e. if the coordinate is only scaled and not translated), if one of these decompositions is totally odd, then so is the other. It is moreover elementary to verify that the valuation $v(h_{\alpha,\beta})$ does not change if we replace α by another center of the disc $D_{\alpha,v(\beta)}$ or replace β by another element of the same valuation; in other words, $v(h_{\alpha,\beta})$ depends only on the choice of disc $D_{\alpha,v(\beta)}$. In light of this and of Remark 4.2(c), the following definition makes sense.

Definition 4.4. *Given a polynomial h and a disc $D \subset \bar{K}$, define $\underline{v}_h(D)$ to be $v(h_{\alpha,\beta})$ for some choice of α, β such that $D = D_{\alpha,v(\beta)}$.*

Recalling the number $t_{q,\rho} \in \mathbb{Q} \cup \{+\infty\}$ defined above, given a finite subset $\mathfrak{s} \subset \bar{K}$ and a disc $D \subset \bar{K}$, let

$$\mathfrak{t}^{\mathfrak{s}}(D) = \underline{v}_\rho(D) - \underline{v}_h(D),$$

where h is a polynomial whose set of roots coincides with $\mathfrak{s}$ and $h = q^2 + \rho$ is a good part-square decomposition.

Our main use of the objects defined above is in the $p = 2$ setting when $\mathfrak{s} \subset \mathcal{R}$ is a cluster, and we choose an element $\alpha \in \mathfrak{s}$ and compute the values $\mathfrak{t}^{\mathcal{R}}(D_{\alpha,b})$ for rational numbers $b \leq d(\mathfrak{s})$. It is not difficult to see that the function $b \mapsto \mathfrak{t}^{\mathcal{R}}(D_{\alpha,b})$ is a piecewise linear function with rational slopes.

Our main result regarding the identification of valid discs containing a given cluster $\mathfrak{s}$ is as follows.

Theorem 4.5. *In the above situation, a disc D satisfying $D \cap \mathcal{R} = \mathfrak{s}$ is valid if and only if we have $D = D_{\alpha,b}$, where b is either the greatest or the least (rational) number in the interval $[d(\mathfrak{s}'), d(\mathfrak{s})]$ satisfying $\mathfrak{t}^{\mathcal{R}}(D_{\alpha,b}) = 2v(2)$ (where $\mathfrak{s}'$ as before is the smallest cluster properly containing $\mathfrak{s}$).*

The above theorem visibly shows that Theorem 3.4(b) holds and that computing the number $B_{f,\mathfrak{s}}$ provided by Theorem 3.4(c) boils down to describing the piecewise linear function $b \mapsto \mathfrak{t}^{\mathcal{R}}(D_{\alpha,b})$ (which in turn can always be computed using a totally odd decomposition of f).

Proposition 4.6. *In the above situation, we have*

$$\mathfrak{t}^{\mathcal{R}}(D_{\alpha,b}) = \min\{\mathfrak{t}^{\mathfrak{s}}(D_{\alpha,b}), \mathfrak{t}^{\mathcal{R} \smallsetminus \mathfrak{s}}(D_{\alpha,b})\}$$

for $d(\mathfrak{s}) \leq b \leq d(\mathfrak{s}')$, where $\mathfrak{s}'$ as before is the smallest cluster properly containing $\mathfrak{s}$.

The above proposition reduces the computation of $b \mapsto \mathfrak{t}^{\mathcal{R}}(D_{\alpha,b})$ to computations which are simpler, as a polynomial with whose roots are $\mathfrak{s}$ or $\mathcal{R} \smallsetminus \mathfrak{s}$ is of lower degree and easier to work with than f. In [8, §6], the authors have also developed further methods of simplifying the computation of these piecewise linear functions via lower-degree polynomials that apply to many special cases.

5. Finding centers of valid discs in the $p = 2$ setting

Given the hyperelliptic curve $y^2 = f(x)$, with $f(x) \in K[x]$ of odd degree $2g + 1$, Proposition 4.3 allows us to produce (for instance by using the procedure explained in the proof) a totally odd decomposition of the translated polynomial $f_{T,1}(z) := f(z + T)$, in which T remains generic rather than being assigned to be particular center $\alpha \in \bar{K}$. Such a decomposition will have the form

$$f_{T,1} = q_{T,1}^2 + \rho_{T,1},$$

with

$$q_{T,1}(z) = Q_0(T) + Q_1(T)z + \ldots + Q_g(T)z^g \qquad \text{and}$$
$$\rho_{T,1}(z) = R_1(T)z + R_3(T)z^3 + \ldots + R_{2g+1}(T)z^{2g+1},$$

where $Q_i(T)$ and $R_i(T)$ are elements of $\overline{K(T)}$, i.e. algebraic functions of the variable T.

Definition 5.1. *Let $L \subset \overline{K(T)}$ be the smallest Galois extension of $K(T)$ to which $R_1(T)$ belongs. We define $F(T) \in K(T)$ to be the norm of $R_1(T)$ with respect to the extension $L/K(T)$. (It can be shown that $F(T)$ is in fact a polynomial in T.)*

Remark 5.2. In the cases of $g \in \{1, 2\}$, assuming for simplicity that f is monic, we may easily find polynomials $q_{T,1}$ and $\rho_{T,1}$ and compute $F(T)$ as the norm of $R_1(T)$. For $0 \leq i \leq 2g + 1$, let $P_i(T) \in K[T]$ be the z^i-coefficient of $f(z + T) \in K[T][z]$. Then for $g = 1$, we have the formula

$$F = P_1^2 - 4P_2P_0, \tag{10}$$

and for $g = 2$, we have the formula

$$F = (P_1^2 - 4P_2P_0)^2 - 64P_4P_0^3. \tag{11}$$

Our motivation for defining the polynomial F is that we are able to obtain the following theorem, which is essentially a generalization of [10, Theorem 5.1] (which treats only the geometrically equidistant case); the underlying strategy of its proof is inspired by that of Lehr and Matignon.

Theorem 5.3. *Each root of F is the center of a valid disc; conversely, if D is a valid disc which contains no cluster, then D contains a root of F. Equivalently, each valid disc contains a root of f or a root of F.*

Given a root α of F which is a center of some valid disc D such that $D \cap \mathcal{R} = \varnothing$, it is possible to show D that is the only valid disc containing α and no root of f and to find the radius of D using minor variations of the methods discussed in §4: defining $\mathfrak{t}^{\mathcal{R}}(D)$ for any disc $D \subset \bar{K}$ as above, there is exactly 1 rational number b in the interval $(d(\mathfrak{s}), \infty)$ such that $\mathfrak{t}^{\mathcal{R}}(D_{\alpha,b}) = 2v(2)$, where $\mathfrak{s}$ is the smallest cluster such that $\mathfrak{s} \subset D_{\alpha,d(\mathfrak{s})}$, and the valid disc we are looking for is $D = D_{\alpha,b}$.

6. The toric rank of a semistable model of a hyperelliptic curve

Let us begin by defining viable clusters, as they directly determine the toric rank of $(\mathcal{Y}^{\mathrm{rst}})_s$.

Definition 6.1. *We say that a cluster $\mathfrak{s}$ is* viable *if the following are satisfied:*

- *(a) $\mathfrak{s}$ has even cardinality; and*
- *(b) there exist 2 distinct valid discs D satisfying that either $D \cap \mathcal{R} = \mathfrak{s}$ or D the minimal disc such that $D \cap \mathcal{R} \supsetneq \mathfrak{s}$.*

Remark 6.2. In the above definition, Theorem 3.2 shows that in the $p \neq 2$ setting, (a) implies (b), while Theorem 3.4(a) shows that in the $p = 2$ setting, (b) implies (a).

Definition 6.3. *An cluster $\mathfrak{s}$ is said to be* übereven *if it is viable and if all of its children clusters are also viable.*

Remark 6.4. In the $p \neq 2$ setting, every even-cardinality cluster is viable, and so an übereven cluster is just a cluster whose children are all even; this is the definition of "übereven" used in [6].

Proposition 6.5. *Let $\mathfrak{s}$ be a viable cluster.*

- *(a) The 2 valid discs $D_1, D_2 \subset K$ containing $\mathfrak{s}$ correspond to 2 components of $(\mathcal{X}^{(\mathrm{rst})})_s$ (corresponding to the models $\mathcal{X}_{D_1}, \mathcal{X}_{D_2}$ of X) which meet at a node whose inverse image in $(\mathcal{Y}^{\mathrm{rst}})_s$ consists of 2 nodes. Each of these 2 nodes has thickness equal to $(\delta(\mathfrak{s}) - B_{f,\mathfrak{s}})/v(\pi)$ where $B_{f,\mathfrak{s}}$ is the "threshold depth" given by Theorem 3.4(c) (resp. $B_{f,\mathfrak{s}} = 0$) in the $p = 2$ (resp. $p \neq 2$) setting (here π is a uniformizer of K').*
- *(b) The cluster $\mathfrak{s}$ is übereven if and only if the minimal disc $D_{\mathfrak{s}}$ containing $\mathfrak{s}$ is valid and the special fiber $(\mathcal{Y}_{D_{\mathfrak{s}}})_s$ (as defined above in §3.1 for $p \neq 2$ and §4 for $p = 2$) consists of 2 disjoint lines.*

It is not difficult, by applying a combinatorial argument to both parts of the above proposition, to provide as a corollary a simple formula for the toric rank of the special fiber of the relatively stable (and thus any semistable) model of Y.

Corollary 6.6. *The toric rank of the special fiber $(\mathcal{Y}^{\mathrm{rst}})_s$ is given by the number of non-übereven viable clusters.*

7. Computations for hyperelliptic curves in genus 2

We now investigate the structure of $(\mathcal{Y}^{\mathrm{rst}})_s$ where Y is a genus-2 hyperelliptic curve; let $Y : y^2 = f(x)$ be the equation of Y, where the polynomial f has degree 5. We moreover make the simplifying assumptions that f is monic, that the depth of the set of roots $\mathcal{R}$ is 0, and that we have $0 \in \mathcal{R}$. The denote the roots of f by $a_1 := 0, a_2, \ldots, a_5$. Clearly there may be 0, 1, 2, or 3 even-cardinality clusters among the cluster data associated to f; except for in the last case of 3 even-cardinality clusters, there may be a single cardinality-3 cluster as well.

The below theorem describes our results on the possible structures of $(\mathcal{Y}^{\mathrm{rst}})_s$ depending on various arithmetic conditions, under the assumption that there exists at most one even-cardinality cluster. Actually, the theorem only addresses the case in which the even-cardinality cluster, if it exists, has cardinality 2 and its parent cluster coincides with $\mathcal{R}$, but it may be adapted to any other cluster picture having at most one even-cardinality cluster; see Remark 7.2(a) below for more details. To treat the case of more than one even-cardinality cluster, instead see Remark 7.2(b),(c).

Theorem 7.1. *In the above situation with the above assumptions on the genus-2 hyperelliptic curve $Y : y^2 = f(x)$, suppose that there are no cardinality-4 clusters and there is at most one cardinality-2 cluster $\mathfrak{s} \subset \mathcal{R}$; if this cluster exists, we denote its relative depth by $m := \delta(\mathfrak{s})$, whereas if there is no even-cardinality cluster, we set $m = 0$. It is clear that $\mathcal{R}$ can contain at most one cardinality-3 cluster $\mathfrak{s}'$; if it exists, we denote its relative depth by $m' := \delta(\mathfrak{s}')$, whereas if there is no cardinality-3 cluster, we set $m' = 0$. We assume that, when both m and m' are > 0, we have $\mathfrak{s} \cap \mathfrak{s}' = \varnothing$.*

We label the roots $a_1, \ldots a_5$ of f in such a way that, when $m > 0$, we have $\mathfrak{s} = \{a_1 = 0, a_2\}$, and when $m' > 0$, we have $\mathfrak{s}' = \{a_3, a_4, a_5\}$. Under the assumption that $m > 0$, we write

$$(1 - a_3^{-1}z)(1 - a_4^{-1}z)(1 - a_5^{-1}z) = 1 + M_1 z + M_2 z^2 + M_3 z^3, \tag{12}$$

and let $w = v(M_1 - 2\sqrt{M_2}) \geq 0$ for some choice of square root of M_2; when $m' > 0$, we have $w = 0$.

Define the polynomial

$$F(T) = (P_1^2(T) - 4P_2(T)P_0(T))^2 - 64P_4(T)P_0^3(T) \in K[T],$$

where $P_i(T)$ is the z^i-coefficient of $f(z+T)$ for $0 \leq i \leq 5$, which we have seen in Remark 5.2 is the polynomial F defined in §5. For any root $\alpha \in \bar{K}$ of F, let $f = q^2 + \rho$ be a part-square decomposition such that the "translated and scaled" part-square decomposition $f_{\alpha,1} = q_{\alpha,1}^2 + \rho_{\alpha,1}$ is totally odd such that $\rho_{\alpha,1}$ has no linear term, and let $\kappa(\alpha)$ be the valuation of the cubic term of $\rho_{\alpha,1}$.

In the language of Theorem 3.4, when $m > 0$, we have $B_{f,\mathfrak{s}} = \max\{4v(2) - w, \frac{8}{3}v(2)\}$. The set of valid discs and the structure of $(\mathcal{Y}^{\mathrm{rst}})_s$ are fully described more precisely as follows. All elements α_i mentioned in parts (b), (c), and (d) below may be chosen to be roots of F, so that in particular $\kappa(\alpha_i)$ is always defined.

- (a) *Suppose that $m > \frac{8}{3}v(2)$ and $w \geq \frac{4}{3}v(2)$. Then there are exactly 2 valid discs $D_- := D_{0,\frac{2}{3}v(2)}$ and $D_+ := D_{0,m-2v(2)}$ such that we have $D_\pm \cap \mathcal{R} = \mathfrak{s}$. The special fiber $(\mathcal{Y}^{\mathrm{rst}})_s$ consists of 2 components corresponding to the discs D_- and D_+ which intersect at 2 nodes and have abelian ranks 1 and 0 respectively.*
- (b) *Suppose that $m > 0$ and $4v(2) - m < w < \frac{4}{3}v(2)$. Then there are two valid discs $D_+ := D_{0,m-2v(2)}$ and $D_- := D_{0,2v(2)-w}$ such that we have $D_\pm \cap \mathcal{R} = \mathfrak{s}$; their corresponding components of $(\mathcal{Y}^{\mathrm{rst}})_s$ each have abelian rank 0 and intersect each other at 2 points. There is moreover another valid disc D_{α_1,b_1}, which does not contain a root of f; we have $v(\alpha_1 - a_i) =$*

$\frac{1}{2}w$ for $i = 1, 2$, $v(\alpha_1 - a_i) = m'$ for $i = 3, 4, 5$ and $b_1 = m' + \frac{1}{3}(w - \kappa(\alpha_1) + 2v(2))$. The corresponding component of $(\mathcal{Y}^{\mathrm{rst}})_s$ has abelian rank 1 and intersects the component corresponding to D_- at 1 node.

(c) Suppose that we have $m > 0$, $w < \frac{1}{2}m$, and $w \leq 4v(2) - m$. Then there are valid discs $D_1 := D_{\alpha_1,b_1}$ and $D_2 := D_{\alpha_2,b_2}$ with $v(\alpha_1 - a_i) = \frac{1}{2}w$ for $i = 1, 2$, $v(\alpha_1 - a_i) = m'$ for $i = 3, 4, 5$, $v(\alpha_2 - a_i) = \frac{1}{2}(m - w)$ for $i = 1, 2$, and $v(\alpha_2 - a_i) = 0$ for $i = 3, 4, 5$, $b_1 = m' + \frac{1}{3}(w - \kappa(\alpha_1) + 2v(2))$, and $b_2 = \frac{1}{3}(m - w - \kappa(\alpha_2) + 2v(2))$. The discs D_i each do not contain a root of f if $w < 4v(2) - m$; when $w = 4v(2) - m$, the disc D_1 does not, but the disc D_2 is the unique valid disc satisfying $D_2 \cap \mathcal{R} = \mathfrak{s}$ and coincides with $D_{0,m-2v(2)}$. The special fiber $(\mathcal{Y}^{\mathrm{rst}})_s$ consists of 2 components corresponding to the discs D_1 and D_2, each of abelian rank 1, which intersect at 1 node.

(d) Finally, suppose that we have $m = 0$, or $0 < m \leq \min\{2w, \frac{8}{3}v(2)\}$. Then there is a valid disc $D_1 := D_{\alpha_1,b_1}$ with $v(\alpha_1 - a_i) = \frac{1}{4}m$ for $i = 1, 2$, and $v(\alpha_1 - a_i) = 0$ for $i = 3, 4, 5$, and $b_1 \geq v(\alpha_1)$. We have the following subcases.

- *(i) Suppose that $\kappa(\alpha_1) < \frac{2}{5}(\frac{1}{2}m + 2v(2))$. Then there is a second valid disc $D_2 := D_{\alpha_2,b_2}$ where α_2 satisfies $v(\alpha_2 - a_i) = \frac{1}{4}m$ for $i = 1, 2$ and $v(\alpha_2 - a_i) = m'$ for $i = 3, 4, 5$, and we have $b_1 = \frac{1}{3}(\frac{1}{2}m - \kappa(\alpha_1) + 2v(2))$ and $b_2 = b_1 + m'$. Neither of the discs D_i contains a root of f. The special fiber $(\mathcal{Y}^{\mathrm{rst}})_s$ consists of 2 components corresponding to the discs D_i, each of abelian rank 1, which intersect at 1 node.*
- *(ii) Suppose that $\kappa(\alpha_1) \geq \frac{2}{5}(\frac{1}{2}m + 2v(2))$. Then the only valid disc is D_1; it is (the unique valid disc) satisfying $D_1 \cap \mathcal{R} = \mathfrak{s}$ if $m = \frac{8}{3}v(2)$ but otherwise does not contain a root of f. Its depth is $b_1 = \frac{1}{5}(\frac{1}{2}m + 2v(2))$. The special fiber $(\mathcal{Y}^{\mathrm{rst}})_s$ thus has exactly 1 component, which has abelian rank 2 (so Y attains good reduction in this case).*

Remark 7.2. The theorem only treats the situation where there are no cardinality-4 clusters and at most one cardinality-2 cluster which is not contained in a cardinality-3 cluster; here we briefly explain how to treat cases where this hypothesis does not hold.

(a) If we consider a situation where the only even-cardinality cluster $\mathfrak{s}$ has relative depth m and cardinality 4 (instead of 2), then on applying an appropriate fractional linear transforation, we obtain cluster data for the resulting isomorphic hyperelliptic curve (now defined using a different polynomial) in which there is a cardinality-2 cluster (and possibly a cardinality-3 cluster disjoint from it). From general formulas for such a fractional linear transformation, one may relate the associated objects defined in the above theorem to those associated to our original polynomial f and derive analogous statements to everything in the above theorem. A similar trick may be used in the case that we begin with a cluster picture such that there is a cardinality-3 cluster $\mathfrak{s}'$ containing 0.

(b) Suppose that there are exactly 2 even-cardinality clusters $\mathfrak{s}_1$ and $\mathfrak{s}_2$ containing roots a_1 and a_2 respectively. Then it is possible to show that we have $B_{f,\mathfrak{s}_1} = B_{f,\mathfrak{s}_2} = 4v(2)$ and to compute each of the valid discs containing $\mathfrak{s}_1$ or $\mathfrak{s}_2$. In fact, this can be derived from the theorem by observing that the quantity w, as defined in its statement, is equal to 0 when there are exactly 2 even-cardinality clusters.

(c) In the case that there are 3 even-cardinality clusters, the computation of valid discs is in general much less straightforward, but the authors have classified the outcomes in the case that each of these clusters has relative depth $\geq 2v(2)$; in particular, if this inequality is strict for all of them, then all of the 3 clusters are viable; the cardinality-4 cluster is übereven; and Corollary 6.6 says that the toric rank is 2.

Remark 7.3. Let $\alpha \in \bar{K}$ be a root of F. One can show that the rational number $\kappa(\alpha)$ is well defined in all contexts of the statement in which its precise value is relevant (more precisely, one can show that it does not depend on the choice of totally odd decomposition with no linear term as long as it is $< 2v(2)$, which is guaranteed to be the case outside of parts (a) and (d)(ii)). It can be computed as

$$\kappa(\alpha) = v\Big(P_3(\alpha) - 2\sqrt{P_4(\alpha)}\sqrt{P_2(\alpha) - 2\sqrt{P_4(\alpha)P_0(\alpha)}}\Big) \tag{13}$$

only for particular choices of the square roots in the above formula.

Example 7.4. Let Y be the hyperelliptic curve of genus 2 over $\mathbb{Z}_2^{\mathrm{unr}}$ given by

$$y^2 = x(x-16)(x-1)(x^2+x-1),$$

so that we have a cardinality-2 cluster $\mathfrak{s} = \{0, 16\}$ of relative (and absolute) depth $m = 4v(2)$. It is straightforward to compute that the polynomial in (12) equals $1 - 2z + z^3$ and so we have $w = v(-2 - 2\sqrt{0}) = v(2)$. The hypothesis of Theorem 7.1(b) clearly holds, and so we have valid discs $D_1 := D_- = D_{0,v(2)}$ and $D_2 := D_+ = D_{0,2v(2)}$ satisfying $D_\pm \cap \mathcal{R} = \mathfrak{s}$; the cluster $\mathfrak{s}$ is therefore viable. In fact, the changes in coordinate corresponding to these discs may be written as

$$x = 2x_1 = 4x_2, \quad y = 4y_1 + 2x_1 = 8y_2 + 4x_2. \tag{14}$$

We now get equations for the corresponding models $\mathcal{Y}_1$ and $\mathcal{Y}_2$ as

$$\begin{aligned} \mathcal{Y}_1 &: y_1^2 + x_1y_1 = 2x_1^5 - 2^4x_1^4 - x_1^3 + 2^3x_1^2 - 2x_1 \\ \mathcal{Y}_2 &: y_2^2 + x_2y_2 = 2^4x_2^5 - 2^6x_2^4 - 2x_2^3 + 2^3x_2^2 - x_2 \end{aligned} \tag{15}$$

The special fibers of these models are the $\bar{\mathbb{F}}_2$-curves given by

$$(\mathcal{Y}_1)_s : y_1^2 + x_1y_1 = x_1^3, \qquad y_2^2 + x_2y_2 = x_2. \tag{16}$$

The desingularizations of $(\mathcal{Y}_1)_s$ and $(\mathcal{Y}_2)_s$ are each smooth curves of genus 0 and give rise to 2 of the components of $(\mathcal{Y}^{\mathrm{rst}})_s$.

However, these are not all of the components of $(\mathcal{Y}^{\mathrm{rst}})_s$, as Theorem 7.1(b) asserts the existence of another valid disc $D_3 := D_{\alpha_1,b_1}$ for some root α_1 of F with $v(\alpha_1) = \frac{1}{2}v(2)$ and $b_1 = 1 - \frac{1}{3}\kappa(\alpha_1)$. Now through tedious but straightforward calculations, one can show that $v(P_3(\alpha_1)) = v(2)$ and $v(P_4(\alpha_1)) = \frac{1}{2}v(2)$ (where P_3 and P_4 are defined as in the statement of Theorem 7.1). From this one may compute by writing down a general formula for a totally odd decomposition of a degree-5 polynomial that we have $\kappa(\alpha_1) = v(2)$ and so $b_1 = \frac{2}{3}v(2)$.

For an appropriate part-square decomposition $f = q^2 + \rho$ such that the decomposition $f_{\alpha_1,1} = q_{\alpha_1,1}^2 + \rho_{\alpha_1,1}$ is totally odd, the change in coordinates corresponding to D_3 can be written as

$$x = 2^{2/3}x_3 + \alpha_1, \quad y = 2^{3/2}y_3 + q_{\alpha_1,1}(2^{2/3}x_3)y_3.$$

We now get an equation for the model $\mathcal{Y}_3$ corresponding to D_3 as

$$\mathcal{Y}_3 : y_3^2 + 2^{-1/2}q_{\alpha_1,1}(2^{2/3}x_3)y_3 = 2^{-3}\rho_{\alpha_1,1}(2^{2/3}x_3). \tag{17}$$

Through further computations of valuations of polynomials appearing in (17), one can now readily verify that the special fiber of $\mathcal{Y}_3$ is the $\bar{\mathbb{F}}_2$-curve given by

$$y_3^2 + c_1y_3 = c_2x_3^3, \tag{18}$$

for some $c_1, c_2 \in k^\times$, and its desingularization is a smooth curve of genus 1 which gives rise to the remaining component of $(\mathcal{Y}^{\mathrm{rst}})_s$. The configuration of the components of $(\mathcal{Y}^{\mathrm{rst}})_s$ is seen in Figure 1.

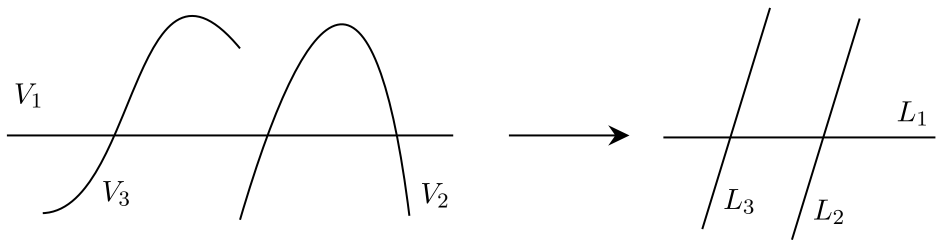

FIGURE 1. The special fiber $(\mathcal{Y}^{\mathrm{rst}})_s$, shown on the left, mapping to $(\mathcal{X}^{(\mathrm{rst})})_s$; each component V_i of $(\mathcal{Y}^{\mathrm{rst}})_s$ maps to each component $L_i := (\mathcal{X}_{D_i})_s$ of $(\mathcal{X}^{(\mathrm{rst})})_s$.

References

[1] Michael Artin and Gayn Winters. Degenerate fibres and stable reduction of curves. *Topology*, 10(4):373–383, 1971.

[2] Irene I. Bouw and Stefan Wewers. Semistable reduction of curves and computation of bad Euler factors of L-functions. ICERM course notes, 2015.

[3] Irene I Bouw and Stefan Wewers. Computing l-functions and semistable reduction of superelliptic curves. *Glasgow Mathematical Journal*, 59(1):77–108, 2017.

[4] Robert Coleman. Computing stable reductions. In *Séminaire de Théorie des Nombres, Paris 1985–86*, pages 1–18. Springer, 1987.

[5] Pierre Deligne and David Mumford. The irreducibility of the space of curves of given genus. *Publications Mathématiques de l'IHES*, 36:75–109, 1969.

[6] Tim Dokchitser, Vladimir Dokchitser, Céline Maistret, and Adam Morgan. Arithmetic of hyperelliptic curves over local fields. *Mathematische Annalen*, pages 1–110, 2022.

[7] Vladimir Dokchitser and Adam Morgan. A note on hyperelliptic curves with ordinary reduction over 2-adic fields. *Journal of Number Theory*, 244:264–278, 2023.

[8] Leonardo Fiore and Jeffrey Yelton. Clusters and semistable models of hyperelliptic curves in the wild case. *arXiv preprint arXiv:2207.12490*, 2022.

[9] Tim Gehrunger and Richard Pink. Reduction of hyperelliptic curves in characteristic $\neq 2$. *arXiv preprint arXiv:2112.05550*, 2021.

[10] Claus Lehr, Michel Matignon, et al. Wild monodromy and automorphisms of curves. *Duke Mathematical Journal*, 135(3):569–586, 2006.

[11] Qing Liu. *Algebraic geometry and arithmetic curves*, volume 6. Oxford University Press on Demand, 2002.

[12] Michel Matignon. Vers un algorithme pour la réduction stable des revêtements p-cycliques de la droite projective sur un corps p-adique. *Mathematische Annalen*, 325(2):323–354, 2003.

[13] Joseph H. Silverman. *The arithmetic of elliptic curves. Graduate Texts in Mathematics*, 106, 2009.

[14] Jeffrey Yelton. Semistable models of elliptic curves over residue characteristic 2. *Canadian Mathematical Bulletin*, 64(1):154–162, 2021.

DEPARTMENT OF MATHEMATICS "FEDERIGO ENRIQUES", THE UNIVERSITY OF MILAN
VIA CESARE SALDINI, 50, 20133 MILANO MI, ITALY
Email address: `leonardo@leonardofiore.it`

DEPARTMENT OF MATHEMATICS AND COMPUTER SCIENCE, WESLEYAN UNIVERSITY
265 CHURCH STREET, MIDDLETOWN, CT 06459-0128
Email address: `jyelton@wesleyan.edu`