Arithmetic Statistics of Algebraic Curves

By

Soumya Sankar

A dissertation submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

(Mathematics)

at the

UNIVERSITY OF WISCONSIN – MADISON

2020

Date of final oral examination: July 13, 2020

The dissertation is approved by the following members of the Final Oral Committee:

Professor Jordan S. Ellenberg, Professor, Mathematics

Professor Daniel M. Erman, Associate Professor, Mathematics

Professor Tonghai Yang, Professor, Mathematics

Professor Simon L. Marshall, Assistant Professor, Mathematics

To Dadu, for teaching me how to live.

Abstract

This thesis explores two kinds of statistical questions about rational points on certain moduli spaces of curves. The first question is, what is the probability that a curve over a finite field is ordinary? Here, a curve $\mathcal{C}$ over a field of characteristic p is said to be ordinary if its Jacobian has largest possible p-torsion. We answer this question for two kinds of curves: Artin-Schreier curves in arbitrary characteristic and superelliptic curves of prime degree in characteristic 2. The second question is, how many elliptic curves over $\mathbb{Q}$ have a cyclic rational N-isogeny? This question can be rephrased in terms of counting rational points on the moduli stacks $\mathcal{X}_0(N)$. We answer this question for $N \in \{2, 3, 4, 5, 6, 8, 9, 12, 16, 18\}$.

Acknowledgements

I would like to start by thanking my advisor, Jordan Ellenberg, although I know that nothing I say can fully express my gratitude for the years of advice and support I have received from him. I would also like to thank the rest of my thesis committee, Tonghai Yang, Daniel Erman and Simon Marshall, for their helpful comments and feedback. The second half of this thesis is based on a project with Brandon Boggess, from whom I have learned a lot, and who has been a very supportive collaborator. I would like to thank Rachel Pries and Jeffrey Achter, whose work inspired my interest in moduli spaces of curves and with whom I have had many helpful and productive conversations, particularly about the first half of this thesis. I would also like to thank David Zureick-Brown for numerous insightful conversations, and for always being ready to answer my questions, however naive. In working on the second half of this thesis, I learned a lot from John Voight and Jeremy Rouse. I would also like to thank Melanie Wood, not only for helpful comments, but also for being a constant source of inspiration.

Now to people without whom I would not have survived graduate school. Some of these people have been helpful mathematically as well, but I cannot overstate the value of their support at a personal level. I would like to thank Daniel Erman, for being incredibly kind and always willing to help; Jordan Ellenberg, for his energy and optimism and for believing in me even when I didn't; numerous people at UW Madison, including Tullia Dymarz, Kathie Brohaugh, Sara Nagreen, Betsy Stovall and Autumn Kent. Several of my friends at Madison formed my support system here, including Wanlin Li, Solly Parenti, Micky Steinberg, Michel Alexis, Iván Ongay Valverde, Polly Yu and Alisha Zachariah; Sun Woo Park, who entertained my crazy choreography ideas. I would like to thank Libby Taylor, for being a patient and inspiring collaborator. I would like to thank the Alberts and Parenti families for giving me homes away from home, and for always welcoming me with open arms. I would also like to thank my aunt and uncle, Mili and Pradeep Khowash, whose support was invaluable

and without whom moving to a different country would have been significantly more challenging.

The writing of this thesis was completed during a time of great turmoil in the world: the COVID-19 pandemic. I would like to thank the people at Dane County Community Defense for creating an incredible network of mutual aid, for being among the kindest people I have ever worked with and for helping me get back on my feet.

Finally, there are the people who have seen me through the best and the worst of times - people who have provided support and strength and without whom I couldn't imagine surviving the world: Brandon Alberts and Supurna Dasgupta, who have always been there for me. And my parents, Sonja and Sankar Datta, who have been incredibly supportive of my journey through graduate school.

Notation and Symbols

$\mathcal{C}$: smooth projective curve

$\mathrm{Jac}(\mathcal{C})$: Jacobian of the curve $\mathcal{C}$

$\mathbb{F}_q$: finite field with q elements

$\mathscr{A}_g$: moduli space of principally polarized abelian varieties of dimension g

$\mathscr{M}_g$: moduli space of smooth genus g curves

$\mathscr{AS}_g$: moduli space of Artin-Schreier curves of genus g

$\mathscr{H}_g$: moduli space of hyperelliptic curves of genus g

$\mathcal{X}_0(N)$: modular curve parametrizing pairs $(E, C \cong \mathbb{Z}/N\mathbb{Z})$

$\mathcal{X}_1(N)$: modular curve parametrizing pairs (E, P) with $\langle P \rangle \cong \mathbb{Z}/N\mathbb{Z}$

$\Gamma_0(N)$: subgroup of $\mathrm{SL}_2(\mathbb{Z})$ consisting of matrices congruent to $\begin{pmatrix} * & * \\ 0 & * \end{pmatrix}$ modulo N

$\Gamma_1(N)$: subgroup of $\mathrm{SL}_2(\mathbb{Z})$ consisting of matrices congruent to $\begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix}$ modulo N

$f(X) = O(g(X))$: there is a constant C such that for large enough X, $|f(x)| \leq C|g(X)|$

$f(x) \asymp g(X)$: there are positive constants K_1 and K_2 such that $K_1 g(X) \leq f(X) \leq K_2 g(X)$

$|Q|$: for a polynomial $Q \in \mathbb{F}_q[x]$, $|Q| = q^{\deg Q}$

$\zeta(s)$: Zeta function of $\mathbb{A}^1_{\mathbb{F}_q}$, given by $\frac{1}{1-q^{1-s}}$

Contents

Chapter 1

Introduction

Arithmetic statistics is the statistical study of objects arising in arithmetic and algebraic geometry. This thesis deals with two different questions, both of the nature:

Question 1.0.1. How many curves of a certain description are there?

The answer to both questions is classically known to be infinite. But how infinite is infinite? For instance, there are infinitely many even numbers as well as squares. However, there are roughly $\frac{1}{2}X$ even integers among all positive integers up to X, while there roughly $X^{1/2}$ squares in the same range. So although both sets are countable, one clearly seems larger than the other. In particular, if we order integers by size, we get a notion of how many of a certain kind there really are and how that number compares with all the integers in the range. In this report, we do the same, but for curves in certain moduli spaces.

1.1 Overview

We explore two directions in this report.

The first is understanding curves over finite fields in certain p-rank strata. A curve $\mathcal{C}$ over a field of characteristic p is said to be ordinary if its Jacobian has largest possible p-torsion. Fixing a finite field $\mathbb{F}_q$, one can ask what is the probability that a curve over $\mathbb{F}_q$ is ordinary? The notion of probability here is defined by counting curves up to a certain *genus* and asking how many such curves are ordinary. Chapter 2 of this report answers this question for certain commonly studied families of curves, namely Artin-Schreier and superelliptic curves. In 2012,

Cais, Ellenberg and Zureick-Brown ([6]) came up with heuristics that showed that if Frobenius behaved randomly in a certain sense, then abelian varieties of arbitrarily large dimension over $\mathbb{F}_q$ had a high probability of being ordinary (made precise later). The randomness hypothesis on the Frobenius matrix is a philosophy that is widely believed to be true. It is often termed as the 'Katz-Sarnak' philosophy and comes in many forms. Some versions of it are known to be true. For example, one such result was used by Achter in proving a large q-limit version of the Cohen-Lenstra heuristics over function fields ([1]). However, the statement required for the large g-limit in [6] is still not known to be true. And while we are quite far from knowing the truth of such a statement, what this report proves, at the very least, is that in some special families of curves, Frobenius does not behave randomly.

The second direction that this report explores is a version of the Batyrev-Manin conjecture for stacks via the classical problem of counting elliptic curves with a rational N-isogeny. It is known that for certain N, there are infinitely many such curves. However, a more precise asymptotic is only known for $N = 1, 2, 3$ and 4 ([17], [28], [29]), ordering elliptic curves by *naive height*. This report provides an asymptotic for certain higher N. Let V be a Fano variety over a number field K (one may think instead, of a scheme with lots of K-rational points) and let X be a real number. The K-rational points on V can be ordered by an invariant called the *height*, coming from an ample line bundle on V. The Batyrev-Manin conjecture predicts that the number of K-rational points on V with height bounded by X is asymptotic to $cX^a \log(X)^b$ for some constants a, b and c. Now, many spaces that parametrize objects of interest are not schemes, but stacks (e.g. moduli spaces of curves, moduli spaces of elliptic curves with an N-isogeny, to name a few), that is, spaces whose points have automorphisms. Not only is counting points on stacks harder, but until recently there was neither a well established theory of heights on stacks, nor a version of the Batyrev-Manin conjecture for them. In [13], the authors establish such a theory and make a similar conjecture in the case when V is a stack. This report shows that the conjecture has the right form when V is the classical moduli stack $\mathcal{X}_0(N)$.

This thesis has four chapters, including the present one. Chapter 2 talks about the first direction, and Chapter 3 about the second. Chapter 4 talks about future work in both directions.

1.2 Main results

1.2.1 Ordinary curves over finite fields

Let $\mathcal{C}$ be a smooth projective curve of genus g over a finite field $\mathbb{F}_q$ of characteristic $p > 0$. Its Jacobian $\mathrm{Jac}(\mathcal{C})$ is an abelian variety of dimension g. For each $n \in \mathbb{Z}_{>0}$, the n-torsion group scheme $\mathrm{Jac}[n]$ is a finite flat group scheme. When $(n, p) = 1$, this group scheme is étale, and as an abelian group, is isomorphic to $(\mathbb{Z}/n\mathbb{Z})^{2g}$ over $\overline{F}_q$. When n is not invertible in $\mathbb{F}_q$, this group scheme is never étale and its isomorphism class over $\overline{\mathbb{F}}_q$ depends significantly on the curve. In particular, there is an integer s with $0 \leq s \leq g$ such that

$$\mathrm{Jac}(\mathcal{C})[p](\overline{\mathbb{F}}_q) \cong (\mathbb{Z}/p\mathbb{Z})^s.$$

We call the curve $\mathcal{C}$ *ordinary* if $s = g$. Let $\mathscr{F}$ denote a set of curves over $\mathbb{F}_q$ of arbitrarily high genus. Define

- $N(\mathscr{F}, X) = \#\{\mathcal{C} \in \mathscr{F} \mid q^g < X\}$
- $N(\mathscr{F}, g, X) = \#\{\mathcal{C} \in \mathscr{F} \mid q^g < X, \mathcal{C} \text{ ordinary}\}$.

Consider the limit

$$P(\mathscr{F}, g) := \lim_{X \to \infty} \frac{N(\mathscr{F}, g, X)}{N(\mathscr{F}, X)}, \tag{1.1}$$

which calculates the probability that a curve $\mathcal{C} \in \mathscr{F}$ is ordinary. We study two kinds of families of curves:

1. *Artin-Schreier curves*: Such curves can be given by an equation of the form

$$y^p - y = f(x),$$

where $f(x) \in \mathbb{F}_q(x)$ and p is the characteristic of the field.

2. *Superelliptic curves*: Such curves can be given by an equation of the form

$$y^n = f(x),$$

where $f(x) \in \mathbb{F}_q[x]$ and $\gcd(n, p) = 1$. We specifically consider superelliptic curves with n an odd prime, over a field of characteristic 2.

In Chapter 2, we prove the following results:

Theorem 1.2.1 (Corollary 2.5.6)**.** *Let p be a prime and q a power of p. The probability that an Artin-Schreier curve over $\mathbb{F}_q$ is ordinary is non-zero for $p = 2$ and zero for all odd primes.*

For the family of superelliptic curves, we prove:

Theorem 1.2.2 (Theorem 2.5.18)**.** *The probability that a superelliptic curve of prime degree over a large enough finite field of characteristic 2 is ordinary, is zero.*

1.2.2 Elliptic curves with a rational N-isogeny

The contents of Chapter 3 are based on joint work with Brandon Boggess. Let E be an elliptic curve over $\mathbb{Q}$. An isogeny $\phi : E \to E^{'}$ between two elliptic curves is said to be a *cyclic N-isogeny* if $\mathrm{Ker}(\phi)(\bar{\mathbb{Q}}) \cong \mathbb{Z}/N\mathbb{Z}$. Further, it is said to be *rational* if $\mathrm{Ker}(\phi)$ is stable under the action of the absolute Galois group, $G_{\mathbb{Q}}$. Henceforth, we will omit the adjective 'cyclic', since these are the only types of isogenies we will consider.

Question 1.2.3. How many elliptic curves over $\mathbb{Q}$ have a rational cyclic N-isogeny?

It is classically known that for $N \leq 10$ and $N = 12, 13, 16, 18, 25$, there are infinitely many such elliptic curves. An elliptic curve E over $\mathbb{Q}$ has a unique minimal Weierstrass equation $y^2 = x^3 + Ax + B$ where $A, B \in \mathbb{Z}$ and $\gcd(A^3, B^2)$ is not divisible by any 12th power. Define the *naive height* of E to be $\mathrm{ht}(E) = \max\{|A|^3, |B|^2\}$. We will order our elliptic curves by this height. In Chapter 3, we show that naive height does indeed come from a line bundle on the stack of elliptic curves.

Notation 1.2.4. For two functions $f, g : \mathbb{R} \to \mathbb{R}$, we say that $f(X) \asymp g(X)$ if there exist

positive constants K_1 and K_2 such that $K_1 g(X) \leq f(X) \leq K_2 g(X)$. For a real number X and positive integer N, define

$$\mathcal{N}(N, X) = \#\{E/\mathbb{Q} \mid \mathrm{ht}(E) < X, E \text{ has a rational } N\text{-isogeny}\}.$$

Precise version of Question 1.2.3: Can we find a function $h_N(X)$ such that $\mathcal{N}(N, X) \asymp h_N(X)$?

Theorem 1.2.5. *Maintaining the notation above, we have the following values of* $h_N(X)$:

N	$h_N(X)$	N	$h_N(X)$
2	$X^{1/2}$	8	$X^{1/6}\log(X)$
3	$X^{1/2}$	9	$X^{1/6}\log(X)$
4	$X^{1/3}$	12	$X^{1/6}$
5	$X^{1/6}(\log(X))^2$	16	$X^{1/6}$
6	$X^{1/6}\log(X)$	18	$X^{1/6}$

Table 1: Values of $h_N(X)$, ordered by naive height

Chapter 2

Proportion of ordinary curves in some families

2.1 Introduction

Let $\mathcal{C}$ be a smooth curve of genus g over a field k of characteristic $p > 0$, and let $\mathrm{Jac}(\mathcal{C})$ denote its Jacobian. Let G be a finite flat group scheme over k killed by p.

Definition 2.1.1. We define the a-number of G as

$$a(G) = \dim_k \mathrm{Hom}(\alpha_p, G)$$

where α_p is the affine group scheme $\mathrm{Spec}(k[x]/x^p)$, and the Hom is in the category of k-group schemes.

Definition 2.1.2. The p-rank of G is defined as $r(G)$ where

$$G(\bar{k}) \cong (\mathbb{Z}/p\mathbb{Z})^{r(G)}$$

as abelian groups.

For the purpose of this thesis, we will only be interested in $G = \mathrm{Jac}(\mathcal{C})[p]$. In this case, it is well known that $0 \leq r(G) \leq g(\mathcal{C})$ and $0 \leq a(G) + r(G) \leq g$. The Jacobian is called *ordinary* if $r(G) = g$ or equivalently, when $a(G) = 0$ ([3]). By abuse of notation, we will denote the $a(\mathcal{C})$ and $r(\mathcal{C})$ to be the corresponding invariants of $\mathrm{Jac}(\mathcal{C})[p]$.

Fix a family $\mathscr{F}$ of curves over $\mathbb{F}_q$ of arbitrary genus. Note that by a family, we mean a set of curves satisfying a particular property, which is not necessarily a family in any geometric sense. A typical example of a family is $\cup_{g\geq 0}\mathscr{M}_g(\mathbb{F}_q)$. Let $\mathscr{F}_s = \{\mathcal{C} \in \mathscr{F} \mid r(\mathcal{C}) = s\}$. The main question that we want to study in this chapter is, what is the probability that a randomly chosen $\mathcal{C} \in \mathscr{F}$ lies in $\mathscr{F}_g$. In other words, what proportion of curves in the family $\mathscr{F}$ is ordinary? Recall the quantity defined in Equation (1.1):

$$P(\mathscr{F}, g) := \lim_{X\to\infty} \frac{N(\mathscr{F}, g, X)}{N(\mathscr{F}, X)}$$

The goal for this chapter is to prove Theorems 1.2.1 and 1.2.2, that is, to calculate $P(\mathscr{F}, g)$ for the Artin-Schreier and superelliptic families.

Notation

Throughout this chapter, k will denote the finite field $\mathbb{F}_q$ of characteristic $p > 0$, unless mentioned otherwise. Most of the definitions involved make sense over any perfect field of characteristic p, but the counting results only make sense over a finite field. All abelian varieties will be assumed to be principally polarized.

2.2 History

The goal of understanding $P(\mathscr{F}, g)$ falls into the larger context of understanding p-divisible groups of abelian varieties in the large g-limit. Let A be an abelian variety defined over k and let $A[p^n]$ denote its p^n-torsion subgroup scheme. Then its p-divisible group is defined as

$$A[p^\infty] := \varinjlim_n A[p^n].$$

This p-divisible group has height $2g$ (see [38] for the definition of the height of a p-divisible group) where g is the dimension of the abelian variety.

Example 2.2.1. If A is an ordinary abelian variety, $A[p^n] \cong (\mathbb{Z}/p^n\mathbb{Z})^g \times (\mu_{p^n})^g$ as a group

scheme. Here $\mu_{p^n} = \operatorname{Spec} k[x]/(x^{p^n} - 1)$ is the kernel of Frobenius on $\mathbb{G}_m$. Thus:

$$A[p^\infty] \cong (\mathbb{Q}_p/\mathbb{Z}_p)^g \times (\mu_{p^\infty})^g.$$

Being a variety in positive characteristic, A carries an action of the Frobenius endomorphism which in turn induces an action on $A[p^\infty]$. We will call this latter operator F. Its dual, Verschiebung, will be denoted by V. Further, $A[p^\infty]$ comes equipped with a skew-symmertric bilinear pairing called the Cartier pairing (see [27]) which realizes the duality between F and V, made more explicit below in Section 2.2.1. Thus any attempt at modelling the behavior of $A[p^\infty]$ must incorporate the actions of F and V, as well as the Cartier pairing. Such an attempt was made by the authors of [6] via Dieudonné modules.

2.2.1 Random Dieudonné Modules and heuristics for p-divisible groups

The Dieudonné functor

Let $W(k)$ denote the ring of Witt vectors over k.

Definition 2.2.2. The Dieudonné Ring over k is defined as $\mathbb{E} := W(k)[F, V]/\sim$, where F and V are two generators subject to the relations $\sim$ given by:

$$\begin{aligned} FV &= VF = p \\ F\lambda &= \lambda^\sigma F \\ V\lambda &= (\lambda)^{\sigma^{-1}} V. \end{aligned}$$

Here $\lambda \in W(k)$ and σ is a lift of the Frobenius map on k to $W(k)$.

There exists a functor, namely the Dieudonne functor $\mathbb{D}$, from the category of group schemes killed by p (resp. the category of p-divisible groups) to the category of $\mathbb{E}$-modules of finite $W(k)$ length (resp. free over $W(k)$) that satisfies the following properties.

1. The rank of the group scheme (resp. height of the p-divisible group) is the length (resp. rank) of the Dieudonne module.

2. Let $G^\vee$ denote the Cartier dual of G. A group scheme is called principally quasi-polarized if there is an isomorphism $\lambda : G \to G^\vee$, such that the following diagram

$$\begin{array}{ccc} G^{\vee\vee} & \xrightarrow{\quad -\lambda^\vee \quad} & G^\vee \\ & {\scriptstyle \imath} \searrow \quad \nearrow {\scriptstyle \lambda} & \\ & G & \end{array}$$

commutes, where $\imath : G^{\vee\vee} \cong G$ is the canonical identification of the double dual of G with G. The polarization of a principally polarized abelian variety induces a principal quasi-polarization on its p^n-torsion. If G is principally quasi-polarized, then $\mathbb{D}(G)$ has a perfect, symplectic pairing.

3. If G is a p-divisible group, then $\mathbb{D}(G[p]) \cong \mathbb{D}/p\mathbb{D}$.

4. The action of F and V on G induces an action on $\mathbb{D}(G)$.

For a more detailed exposition and construction of the Dieudonne module, we refer the reader to [27].

Examples

1. Let $G = E[p^\infty]$, where E is an elliptic curve over k. If E is ordinary, then

$$L := \mathbb{D}(G) \cong \mathbb{E}/(F, 1-V) \oplus \mathbb{E}/(F-1, V).$$

If E is supersingular, then since $a_q \equiv 0 \mod p$, we have that $I := D(G) \cong \mathbb{E}/(F+V)$.

2. Let A be a principally polarized abelian variety with p-rank f and a-number $g-f$. Then $\mathbb{D}(A[p]) \cong L^f \oplus I^{g-f}$.

In [6], the authors define the notion of a random Dieudonne module, which models the behavior of the p-divisible group of an abelian variety. A *random Dieudonné module* of height $2g$ is a tuple $(\mathbb{D}, F, V, \omega)$ where $\mathbb{D}$ is a $W(k)$-module of rank $2g$, F and V are operators subject to relations in Definition 2.2.2, and ω is a symplectic pairing on $\mathbb{D}$ such that $\omega(Fx, y) = \omega(x, Vy)^\sigma$. The authors show that if we fix one choice of Frobenius, F_0 on $W(k)^{2g}$, then all other choices of Frobenius must come from the double coset $\mathrm{Sp}_{2g}(W(k))F_0\mathrm{Sp}_{2g}(W(k))$. A *random* choice of the

tuple $(\mathbb{D}, F, V, \omega)$ amounts to picking F from $\mathrm{Sp}_{2g}(W(k))F_0\mathrm{Sp}_{2g}(W(k))$ uniformly with respect to the product of the Haar measures on each $\mathrm{Sp}_{2g}(W(k))$. Define the p-rank and a-number of a Dieudonné module $\mathbb{D}$ are defined as those of $\mathbb{D}/p\mathbb{D}$. Using the model described above, the authors show that the probability that such a module is ordinary is

$$\prod_{i=1}^{\infty}(1+q^{-i})^{-1}. \tag{2.1}$$

Further, they ask if the Dieudonné module associated to the Jacobian of a curve behaves like a randomly chosen one, i.e. whether the limit in Equation (1.1) equals the quantity (2.1) when $\mathscr{F}$ is the set of all smooth curves. They find, via numerical experiments, that hyperelliptic curves in small odd characteristic do not appear to obey their heuristics, while plane curves do. The families considered in this report are the first known cases whose behavior provably diverges significantly from the heuristics of [6]. While the results in this chapter are motivated by the heuristics in [6], the approach used is quite different. For instance, we do not prove the randomness of Frobenius in any sense. What we use instead is a combinatorial criterion for ordinariness that we deduce from work of Pries and Zhu in [30], and Elkin in [12].

2.2.2 Large g-limits and large q-limits

Arithmetic statistics of curves and abelian varieties over finite fields fall into two broad categories: taking limits as $q \to \infty$ with g fixed or as $g \to \infty$ with q fixed. The results in this report are examples of the latter, as are those in the motivating paper, [6].

Comparing the two behaviors

Large q-limit behavior can often be thought of as the geometric behavior of a family of curves with a fixed genus. For instance, one might ask, what is the codimension of the ordinary locus inside $\mathscr{M}_g$? To study this question, one can change the base field to $\overline{\mathbb{F}}_q$ without loss of generality. The large q-limit behavior in this sense is usually incomparable to the large g-limit behavior. However, studying the former can provide some insight into the latter. To illustrate our point, we list some results here that show how different the geometry of the Artin-Schreier locus is

from that of some other families of curves. It is known that the locus of ordinary curves is a non-empty Zariski open subset of $\mathscr{M}_g$ ([26]). Thus for a fixed genus g, 'most' curves of genus g tend to be ordinary. Let $\mathscr{V}_{g,r}$ denote the sublocus of $\bar{\mathscr{M}}_g$ of curves of p-rank at most r. In [14] Faber and van der Geer prove that $\mathscr{V}_{g,r}$ has codimension $g-r$. A result of Glass and Pries [15] states that $\mathscr{V}_{g,r}$ intersects the hyperelliptic locus, $\mathscr{H}_g$, inside $\bar{\mathscr{M}}_g$ in a set of dimension $g-1+r$. Since $\mathscr{H}_g$ has dimension $2g-1$, this implies that the ordinary locus is dense in $\mathscr{H}_g$. We compare this to results about $\mathscr{AS}_g$, the Artin-Schreier locus inside $\mathscr{M}_g$. In [30], Pries and Zhu prove that for $p \geq 3$, the codimension of $\mathscr{V}_{g,r} \cap \mathscr{AS}_g$ inside $\mathscr{AS}_g$ is less than $g-r$. This indicates that for $p \geq 3$, the image under the Torelli morphism of $\mathscr{AS}_g$ in $\mathscr{A}_g$ (the moduli space of principally polarized abelian varieties of genus g) is not in general position with respect to the p-rank stratification. Further, from results in [30] which we state in the next section (Theorem 2.3.2), it follows that the ordinary locus intersects only one irreducible component of $\mathscr{AS}_g$. As $g \to \infty$, the number of components of $\mathscr{AS}_g$ increases except when $p = 2$ (in which case $\mathscr{AS}_g$ is $\mathscr{H}_g$). This gives a heuristic reason for why one might expect a statement like Theorem 1.2.1. A similar heuristic explains Theorem 1.2.2 as well, as we elaborate in Remark 2.5.19.

Equidistribution results

Another context in which the large g-limit versus large q-limit dichotomy arises, is in equidistribution results. As mentioned in Chapter 1, there is a philosophy that governs statistical questions about varieties over finite fields, often called the Katz-Sarnak philosophy. This is a Chebotarev density theorem-like claim about the Frobenius endomorphism. Let G be the arithmetic monodromy group of a family of curves of a fixed genus over a finite field $\mathbb{F}_q$. Let W be a conjugacy class in G. Roughly speaking, equidistribution results about Frobenius state that the probability that Frobenius belongs to W is equal to $|W|/|G|$. Of course, this equality does not hold strictly, but is true up to an error term that is $O(1/\sqrt{q})$. This error term makes such equidistribution results amenable for proving large q-limit results (see for example, [1] or [2]). The dependence of the error term on g, however, is much more complicated and harder to bound.

2.3 Background: Artin-Schreier Curves

We now recall some facts about Artin-Schreier curves and covers. An Artin-Schreier curve $\mathcal{C}$ over k is a smooth $\mathbb{Z}/p\mathbb{Z}$ cover of $\mathbb{P}^1_k$. Such a curve has an affine model

$$y^p - y = f(x) \tag{2.2}$$

where $f(x) \in k(x)$, and is equipped with a $\mathbb{Z}/p\mathbb{Z}$ action generated by $y \mapsto y + 1$. An *Artin-Schreier cover* is an Artin-Schreier curve along with a choice of map $\iota : \mathbb{Z}/p\mathbb{Z} \hookrightarrow \mathrm{Aut}(\mathcal{C})$ and a choice of isomorphism $\mathcal{C}/(\iota(\mathbb{Z}/p\mathbb{Z})) \cong \mathbb{P}^1$. This amounts to picking a model of the form 2.2.

Let $B \subset \mathbb{P}^1(\bar{k})$ be the set of poles of f. Then, the cover above is ramified precisely at the points in B ([36]). For $\alpha \in B$, let

$$x_\alpha = \begin{cases} \frac{1}{x-\alpha} & \alpha \neq \infty \\ x & \alpha = \infty. \end{cases}$$

Then, using a partial fraction decomposition one can write

$$f(x) = \sum_{\alpha \in B} f_\alpha(x_\alpha) \tag{2.3}$$

where $f_\alpha \in \bar{k}[x]$ is a polynomial of degree d_α.

Remark 2.3.1. We now make a few helpful observations about the partial fraction decomposition above.

1. We may, and do assume that for $\alpha \neq \infty$, f_α has no constant term.
2. By a transformation of the form $y \mapsto y + z$, one can assume that in $f_\alpha(x)$, the coefficient of x^{ip} is zero for any $0 \leq i \leq \lfloor d_\alpha/p \rfloor$. In particular, we can take $d_\alpha \neq 0 \mod p$.
3. If $\alpha, \beta \in B$ are Galois conjugate, then $d_\alpha = d_\beta$.

4. Let Q be an irreducible polynomial in $k[x]$ whose zeroes are ramified in the Artin-Schreier cover under consideration. Then we will denote d_Q as the degree of any f_α, α a zero of Q. This is well defined by the above remark.

By the Riemann-Hurwitz theorem for wildly ramified covers, we know that the genus of such a curve is given by:

$$g = \left(\frac{p-1}{2}\right)\left(-2 + \sum_{\alpha \in B}(d_\alpha + 1)\right) = \left(\frac{p-1}{2}\right)\left(-2 + \sum_{\substack{Q \text{ irred.} \\ \text{ramified}}} \deg(Q)(d_Q + 1) + (d_\infty + 1)\right). \tag{2.4}$$

2.3.1 The moduli space of Artin-Schreier curves

In [30], the authors give a complete description of the irreducible components of the p-rank strata of of $\mathscr{AS}_g$. Let $\mathscr{AS}_{g,s} \subset \mathscr{AS}_g$ denote the space of Artin-Schreier curves of genus g with p-rank s. It follows from the Deuring-Shafarevich formula (see for instance, [8, Corollary 1.8]) that s is divisible by $p-1$.

Theorem 2.3.2 ([30], Theorem 1.1)**.** *Let $g = d(p-1)/2$ with $d \geq 1$ and $s = r(p-1)$ with $r \geq 0$. Then:*

1. *The set of irreducible components of $\mathscr{AS}_{g,s}$ is in bijection with partitions $\{e_1, e_2 \dots e_{r+1}\}$ of $d+2$ into $r+1$ positive integers such that each $e_i \not\equiv 1 \mod p$.*
2. *The irreducible component of $\mathscr{AS}_{g,s}$ corresponding to the partition $\{e_1, e_2 \dots e_{r+1}\}$ has dimension:*
$$d - 1 - \sum_{i=1}^{r+1} \lfloor (e_i - 1)/p \rfloor.$$

This theorem implies, in particular, that the closure of the ordinary locus has dimension $d-1$.

The following criterion for the ordinarity of an Artin-Schreier curve follows from the above description of the moduli space, but was known earlier as well ([8], [37]):

Corollary 2.3.3. *The Artin-Schreier cover $y^p - y = f(x)$ is ordinary if and only if f has only simple poles.*

This is equivalent to the condition that $d_\alpha = 1$ for each α in the partial fraction decomposition (2.3).

Let $\mathcal{S}$ be the set of rational functions $f(x) \in k(x)$ such that the partial fraction decomposition of f satisfies the conditions (1-3) from Remark 2.3.1. For simplicity, we will assume that $\infty \notin B$. This assumption is harmless, as we explain in Remark 2.5.7 and makes the computations in §2.5 much cleaner. We now restrict our attention to $k = \mathbb{F}_q$ and define the families for this section as follows:

- $\mathscr{F}$ = Set of Artin-Schreier covers $y^p - y = f(x)$, where $f(x) \in \mathcal{S}$ has no poles over $\infty \in \mathbb{P}^1$.
- $\mathscr{F}_g$ = Set of all ordinary Artin-Schreier covers $y^p - y = f(x)$ with $f(x) \in \mathcal{S}$, unramified over $\infty \in \mathbb{P}^1$.

2.3.2 Aside on counting curves versus counting covers

In our proof of the main theorem in Section 2.5, we calculate the probability $P(\mathscr{F}, g)$ by counting polynomials in the set $\mathcal{S}$ defined above. We must however, make the distinction between counting covers versus counting curves. One wishes to count rational points on $\mathscr{AS}_g \hookrightarrow \mathscr{M}_g$, that is to count isomorphism classes of Artin-Schreier *curves*. However, what we actually do in this report, is count isomorphism classes of Artin-Schreier *covers*. That is, we count models for the curves instead of curves themselves. For $p \geq 7$, this does not change the proportion of ordinarity, as we explain later in Remark 2.5.7. For $p = 3, 5$ such a conclusion is beyond reach right now, while for $p = 2$, it is simply not true.

There is a map

$$\mathcal{S} \to \bigcup_{g \geq 0} \mathscr{AS}_g(\mathbb{F}_q) \tag{2.5}$$

sending $f(x) \in \mathcal{S}$ to the curve with model $y^p - y = f(x)$. Remark 2.3.1 shows that this map is surjective. We will now bound the size of the fibers. For an Artin-Schreier curve $\mathcal{C}$, a choice of

model $\mathcal{C}_f$ amounts to a choice of homomorphism $\imath : \mathbb{Z}/p\mathbb{Z} \hookrightarrow \mathrm{Aut}(\mathcal{C})$ and a choice of isomorphism $\mathcal{C}/\imath(\mathbb{Z}/p\mathbb{Z}) \cong \mathbb{P}^1$. For $g \geq 2$, Stichtentoth proved (see [34], [35]) that $|\mathrm{Aut}(\mathcal{C})| \leq 16g(\mathcal{C})^4$. We claim that the number of choices of isomorphism $\mathcal{C}/\imath(\mathbb{Z}/p\mathbb{Z}) \cong \mathbb{P}^1$ is bounded uniformly.

Proposition 2.3.4. *Let $\mathcal{C}_f$ and $\mathcal{C}_g$ be two Artin-Schreier covers with $\phi : \mathcal{C}_f \cong \mathcal{C}_g$ such that there is a commutative diagram*

$$\begin{array}{ccc} \mathcal{C}_f & \xrightarrow{\phi} & \mathcal{C}_g \\ \downarrow & & \downarrow \\ \mathbb{P}^1_{\mathbb{F}_q} & \xrightarrow{\tilde{\phi}} & \mathbb{P}^1_{\mathbb{F}_q} \end{array}$$

where the vertical maps are quotients by the $\mathbb{Z}/p\mathbb{Z}$ actions. Then $f(x) = ug(\gamma x)$ for some $u \in \mathbb{Z}/p\mathbb{Z}^\times$ and $\gamma \in \mathrm{PGL}_2(\mathbb{F}_q)$.

Proof. The induced map $\tilde{\phi}$ is induced by some $\gamma \in \mathrm{PGL}_2(\mathbb{F}_q)$. Let D_f and D_g denote the ramification divisors of $\mathcal{C}_f$ and $\mathcal{C}_g$ respectively. By Artin-Schreier theory, these are determined by the poles of f and g respectively. Note that since the curves are defined over $\mathbb{F}_q$, so are their ramification divisors. Since ϕ must preserve the ramification invariants (namely, the number of ramified points and the ramification groups at each of these points), we must have that $\tilde{\phi}^*(D_g) = D_f$. Thus $\mathcal{C}_{f\circ\gamma}$ and $\mathcal{C}_g$ are isomorphic curves with the same ramification divisor.

Now, two Artin-Schreier covers,

$$y^p - y = f_1(x) \quad \text{and} \quad y^p - y = f_2(x)$$

with the same genus and ramification divisor are isomorphic if and only if $f_1(x) = uf_2(x) + \delta^p - \delta$ (see, for example [30], Remark 3.9) with $u \in \mathbb{Z}/p\mathbb{Z}^\times$ and $\delta \in \mathbb{F}_q(x)$. Since we have imposed the condition that $f(x), g(x) \in \mathcal{S}$, the proposition follows. □

Thus we have that for $g \geq 2$, the map on the genus g part in Equation 2.5 has fibers of size bounded by $C(q)g^4$, where $C(q)$ is a constant. For notational convenience, let $\mathscr{G} = \cup_{g\geq 0}\mathscr{A}\mathscr{S}_g(\mathbb{F}_q)$.

Since $|\{f \in \mathcal{S} : g(\mathcal{C}_f) = g\}| \leq C(q)g^4 |\{\mathcal{C} \in \mathscr{A}\mathscr{S}_g(\mathbb{F}_q)\}|$, therefore

$$\begin{aligned} P(\mathscr{G}, g) &= \lim_{X\to\infty} \frac{|\ \{\mathcal{C} \in \mathscr{A}\mathscr{S}_g(\mathbb{F}_q) \mid q^g < X, s(\mathcal{C}) = g\}\ |}{|\ \{\mathcal{C} \in \mathscr{A}\mathscr{S}_g(\mathbb{F}_q) \mid q^g < X\}\ |} \\ &\leq \lim_{X\to\infty} C(q)\log_q(X)^4 \frac{|\ \{f \in \mathcal{S} \mid q^{g(\mathcal{C}_f)} < X, s(\mathcal{C}_f) = g\}\ |}{|\ \{f \in \mathcal{S} \mid q^{g(\mathcal{C}_f)} < X\}\ |}. \end{aligned}$$

The counting arguments in Section 2.5 will show that for $p \geq 7$, the quantities $P(\mathscr{G}, g)$ and $P(\mathscr{F}, g)$ are the same.

2.4 Background: Superelliptic curves

A *superelliptic curve* over a field k is a curve defined by the affine equation

$$y^n = f(x),$$

where $f(x) \in k[x]$ and n is coprime to the characteristic of k. This curve has an action of $\boldsymbol{\mu}_n$ (n-th roots of unity) on it, namely the map

$$(x, y) \mapsto (x, \zeta_n y),$$

where ζ_n is a primitive n-th root of unity. One can make a transformation to write

$$f(x) = \prod_{i=1}^{n-1} (f_i(x))^i \tag{2.6}$$

where each $f_i(x)$ is a squarefree polynomial, and $f_i(x)$ and $f_j(x)$ are coprime if $i \neq j$. The quotient $\mathcal{C}/\boldsymbol{\mu}_n$ gives a map to $\mathbb{P}^1$, sending $(x, y) \mapsto x$. We let $N := \sum_{i=1}^{n-1} i \deg(f_i)$. Then the curve $\mathcal{C}$ is unramified over $\infty \in \mathbb{P}^1$ if and only if $N \equiv 0 \mod n$. In the case that $N \not\equiv 0 \mod n$, we let n_∞ be the smallest positive integer such that $N + n_\infty \equiv 0 \mod n$.

The map $\mathcal{C} \to \mathbb{P}^1$ is ramified at the zeros of f and possibly at ∞. The ramification indices at the ramified points $\alpha \in \mathbb{P}^1(\bar{k})$ are given by (see [21])

$$e(\alpha) = \begin{cases} \frac{n}{(n,i)} & if\ f_i(\alpha) = 0 \\ \frac{n}{(n,n_\infty)} & if\ \alpha = \infty \end{cases}$$

The genus of this curve is given by

$$g = -n + 1 + \frac{1}{2}\sum_{i=1}^{n-1} \deg(f_i)(n - (n,i)) + \frac{1}{2}\epsilon(n - (n, n_\infty)) \tag{2.7}$$

where ϵ is 0 if the map $\mathcal{C} \to \mathbb{P}^1$ is unramified over ∞ and 1 otherwise.

Remark 2.4.1. Since the techniques of this chapter are based on counting polynomials, it is necessary to separate the case when the map is ramified over $\infty \in \mathbb{P}^1$, even though that seems unnatural.

We now specialize to the case where n is an odd prime. Let $B \subset \mathbb{P}^1(\bar{k})$ be the set of points ramified in the cover $y^n = f(x)$. Let $\mathbf{m} = \mid B \mid$. If $\epsilon = 0$, then $\mathbf{m} = \sum_{i=1}^{n-1} \deg(f_i)$ and if $\epsilon = 1$, then $\mathbf{m} = \sum_{i=1}^{n-1} \deg(f_i) + 1$.

In either case, we have,

$$g = \frac{1}{2}(n-1)(\mathbf{m} - 2). \tag{2.8}$$

Thus with regard to superelliptic curves, we will be interested in the family $\mathscr{F}$ of covers $y^n = f(x)$, where

- n is prime,
- the curve is defined over $\mathbb{F}_q$, where q is a power of 2,
- $f(x) \in \mathbb{F}_q[x]$ is n-th power-free.

2.4.1 a-numbers of superelliptic curves in characteristic 2

We now give a combinatorial criterion for the ordinarity for superelliptic curves in characteristic 2. The discussion in this section is based on a paper by Elkin [12]. Let $\mathcal{C}$ be a smooth proper superelliptic curve over $\mathbb{F}_q$, q a power of 2, with affine model $y^n = f(x)$, where n is an odd prime. We maintain the same notation as before. The space $H^0(\mathcal{C}, \Omega^1_{\mathcal{C}})$ inherits the action of $\boldsymbol{\mu}_n$ and decomposes into eigenspaces as follows:

$$H^0(\mathcal{C}, \Omega^1_{\mathcal{C}}) = \oplus_{i=1}^{n-1} \mathcal{D}_i.$$

A key player in Elkin's work is the Cartier operator, $\mathscr{C}$. This is a Frob^{-1}-linear operator on $H^0(\mathcal{C}, \Omega^1_{\mathcal{C}})$, which annihilates exact differentials and preserves logarithmic differentials. It can be thought of as capturing the action of Verschiebung. It is well known that the a-number, $a(\mathcal{C})$, equals $g(\mathcal{C}) - \mathrm{rank}(\mathscr{C})$. To state the result in Elkin's paper, we first describe some notation. Let $d_i = \dim(\mathcal{D}_i)$. Let σ be the permutation of $\{1, 2, \ldots n-1\}$ defined by

$$p\sigma(i) \equiv i \mod n. \tag{2.9}$$

By bounding the rank of the Cartier operator, Elkin proves the following proposition.

Proposition 2.4.2 ([12], Corollary 1.4)**.** *Let $\mathcal{C}$ be as above. Then,*

$$g(\mathcal{C}) - a(\mathcal{C}) = \sum_{i=1}^{n-1} \min(d_i, d_{\sigma(i)})$$

where the $d_i = \dim(\mathcal{D}_i)$ can be computed explicitly from the ramification invariants of the curve and σ is the permutation of the set $\{1, 2, \cdots n-1\}$ defined by the congruence (2.9)*.*

For any rational number r, let $\langle r \rangle = r - \lfloor r \rfloor$. Elkin proves that the d_i's are given by the formula:

$$d_i = \sum_{j=1}^{n-1} \deg(f_j) \left\langle \frac{ij}{n} \right\rangle + \left\langle \frac{in_\infty}{n} \right\rangle - 1. \tag{2.10}$$

Recall that the ordinarity of an abelian variety is equivalent to the condition that its a-number is 0. Proposition 2.4.2 tells us that if $a(\mathcal{C}) = 0$, then $g(\mathcal{C}) = \sum_{i=1}^{n-1} \min(d_i, d_{\sigma(i)})$. We now give a condition for ordinarity in terms of the degrees of f_i. For better exposition, we will treat the case $n = 3$ separately from the case of a general odd prime.

The case $n = 3$

In this subsection, we consider curves of the form $\mathcal{C} : y^3 = f(x)$. The equation for the genus simplifies to

$$g = \mathbf{m} - 2.$$

Proposition 2.4.3. *A curve of the form $y^3 = f_1 f_2^2$, with f_1, f_2 squarefree is ordinary if and only if one of the following is true:*

1. *$n_\infty = 0$ and $\deg(f_1) = \deg(f_2)$, or*
2. *$n_\infty = i$ for some $i \in \{1, 2\}$ and $\deg(f_i) + 1 = \deg(f_{3-i})$.*

Proof. Since $\sigma = (1\ 2)$, therefore $g = 2\min(d_1, d_2)$. This in turn implies $g = 2d_1$ or $g = 2d_2$. We prove case (1) here. The other case follows by a similar calculation.

In this case,

$$d_1 = \frac{1}{3}\deg(f_1) + \frac{2}{3}\deg(f_2) - 1$$

and

$$d_2 = \frac{2}{3}\deg(f_1) + \frac{1}{3}\deg(f_2) - 1.$$

Therefore, $\deg(f_1) = \deg(f_2)$. For case (2), we just replace $\deg(f_i)$ by $\deg(f_i) + 1$ in the expression for each d_j. □

The case of a general odd prime

Proposition 2.4.4. *A curve defined by* $y^n = \prod_{i=1}^{n-1}(f_i(x))^i$ *as in section 2.4 with* n *an odd prime, is ordinary if and only if one of the following is true:*

1. $n_\infty = 0$ *and* $\deg(f_i) = \deg(f_{n-i})$*, or*
2. $n_\infty = i$ *for some* $i \in \{1, 2 \dots n-1\}$*, and* $\deg(f_i) + 1 = \deg(f_{n-i})$*, and for all* $j \neq i, n-i$*,* $\deg(f_j) = \deg(f_{n-j})$.

Proof. As before, we only prove case (1) and the other case follows from a modified, but similar calculation. The condition for ordinarity gives: $\sum_i d_i = \sum_i \min(d_i, d_{\sigma(i)})$. This automatically implies that $d_i = d_j$ for all $1 \leq i, j \leq n$. Since we are considering the case where $n_\infty = 0$,

$$d_i = \sum_{j=1}^{n-1} \deg(f_j) \left\langle \frac{ij}{n} \right\rangle - 1.$$

Define the matrix A, with $A_{ij} = \left\langle \frac{ij}{n} \right\rangle$. Thus the degrees of f_i's must be solutions to the linear system

$$A \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \end{pmatrix} = \begin{pmatrix} d+1 \\ d+1 \\ \vdots \\ d+1 \end{pmatrix} \tag{2.11}$$

for some $d \geq 0$. Let V denote the space of $n-1 \times 1$ vectors whose coordinates are all equal. We are interested in (the integral points of) the space of $x = (x_1, x_2 \dots x_{n-1})^T$ such that $Ax \in V$.

Lemma 2.4.5. *The space* $\{x \in \mathbb{Z}^{n-1} \mid Ax \in V\}$ *consists of vectors* x *for which* $x_k = x_{n-k}$ *for all* $k = 1, 2 \dots n-1$.

Proof of Lemma. We prove this lemma by constructing an explicit basis for the kernel of A, $\mathrm{Ker}(A)$. Let $x^{(k)}$ denote the $n-1 \times 1$ vector which has 1's in the kth and $n-k$th positions and -1's in the $\frac{n-1}{2}$th and $\frac{n+1}{2}$th positions. We claim that $\{x^{(k)} \mid k = 1, 2, \dots \frac{n-3}{2}\}$ is a basis for $\mathrm{Ker}(A)$.

$$\begin{aligned}(Ax^{(k)})_i &= \left(\frac{ik}{n} - \left\lfloor \frac{ik}{n} \right\rfloor\right) + \left(\frac{i(n-k)}{n} - \left\lfloor \frac{i(n-k)}{n} \right\rfloor\right) - \left(\frac{i(n-1)}{2n} - \left\lfloor \frac{i(n-1)}{2n} \right\rfloor\right) \\ &\quad - \left(\frac{i(n+1)}{2n} - \left\lfloor \frac{i(n+1)}{2n} \right\rfloor\right) \\ &= \left\lfloor \frac{i(n-1)}{2n} \right\rfloor + \left\lfloor \frac{i(n+1)}{2n} \right\rfloor - \left\lfloor \frac{ik}{n} \right\rfloor - \left\lfloor \frac{i(n-k)}{n} \right\rfloor \\ &= 0\end{aligned}$$

Thus, it only remains to prove that A has rank at least $\frac{n+1}{2}$. Now, nA can be row reduced such that the top left $\frac{n+1}{2} \times \frac{n+1}{2}$ submatrix looks like

$$\begin{pmatrix} 1 & 2 & 3 & \dots & \frac{n-1}{2} & \frac{n+1}{2} \\ 0 & 0 & 0 & \dots & 0 & * \\ 0 & 0 & 0 & \dots & * & * \\ \vdots & & & & & \\ 0 & 0 & * & \dots & * & * \\ 0 & * & * & \dots & * & * \end{pmatrix}$$

where each of the entries immediately below the anti-diagonal is necessarily non zero. Such a matrix has non-zero determinant. Thus, any element in $\mathrm{Ker}(A)$ is of the form

$$(x_1, x_2, \dots - \sum_{i=1}^{\frac{n-3}{2}} x_i, - \sum_{i=1}^{\frac{n-3}{2}} x_i, \dots x_2, x_1)^T$$

This proves the lemma and hence the proposition.

□

Remark 2.4.6. Perhaps a more natural way to interpret Propositions 2.4.3 and 2.4.4 is to say that for a curve $y^n = f(x)$ (n prime) has ordinary Jacobian if and only if the same number of points are ramified to degree i and $n - i$ for any $i \in \{1, 2 \dots n-1\}$. Here we say that a point

P is 'ramified to degree i' if the curve locally looks like $y^n = ux_P^i$, where x_P is a uniformizer at P and u is unit.

2.4.2 Aside on counting curves versus counting covers

One might wonder, as in the Artin-Schreier case in §2.3, what the difference is between counting superelliptic curves and covers of the form $y^n = f(x)$. We choose to restrict our attention to *covers*, i.e. to *equations* of the form $y^n = f(x)$ with $f(x) \in \mathbb{F}_q[x]$ n-th power-free, and make the claim that this does not significantly affect our results.

We first introduce some notation for this section alone. For any $u \in (\mathbb{Z}/n\mathbb{Z})^\times$, let $[u]$ be the map that takes $\prod_i (f_i(x))^i$ to $\prod_i (f_i(x))^{(ui \mod n)}$. By a straightforward sequence of transformations, one can see that if f_i is squarefree for each i the two curves given by

$$y^n = \prod_i (f_i(x))^i \qquad \text{and} \qquad y^n = \prod_i (f_i(x))^{(ui \mod n)}$$

are indeed isomorphic. By abuse of notation, we also call this isomorphism of curves $[u]$. We claim that up to an automorphism of $\mathbb{P}^1_{\mathbb{F}_q}$, the only isomorphisms between superelliptic covers are of the form $[u]$, with $u \in (\mathbb{Z}/n\mathbb{Z})^\times$. This is a standard Kummer theory argument, whose proof we recall here.

Proposition 2.4.7. *For n an odd prime, let $f(x) = \prod_{i=1}^{n-1} (f_i(x))^i$ and $g(x) = \prod_{i=1}^{n-1} (g_i(x))^i$ be two monic n-th power-free polynomials in $\mathbb{F}_q[x]$ such that:*

- *For each i, $f_i(x)$ and $g_i(x)$ are squarefree,*
- $\mathrm{div}_0(f) = \mathrm{div}_0(g)$.

Suppose that $\mathcal{C}_f : y^n = f(x)$ and $\mathcal{C}_g : y^n = g(x)$ are isomorphic as curves via an isomorphism ϕ. Let ζ_n be an n-th root of unity that acts as an automorphism of the curve sending $(x, y) \mapsto (x, \zeta_n y)$. Then there is a $u \in (\mathbb{Z}/n\mathbb{Z})^\times$ such that $\phi = \zeta_n \circ [u]$.

Proof. Let $K = \mathbb{F}_q(\mathbb{P}^1)$ and $L = \mathbb{F}_q(\mathcal{C}_f) \cong \mathbb{F}_q(\mathcal{C}_g)$. Note $L(\zeta_n)/K(\zeta_n)$ is a Galois extension. Let

$\varphi : \mathrm{Gal}(\overline{K(\zeta_n)}/K(\zeta_n)) \to \boldsymbol{\mu}_n$ be the homomorphism corresponding to $L(\zeta_n)$. Any other field $L^{'}$ that is isomorphic to $L(\zeta_n)$ corresponds to the homomorphism φ^u for some $u \in (\mathbb{Z}/n\mathbb{Z})^\times$. Therefore, if $[\alpha] \in K(\zeta_n)^\times/(K(\zeta_n)^\times)^n$ is the class corresponding to φ via the Kummer map, then there is a $u \in (\mathbb{Z}/n\mathbb{Z})^\times$ such that the isomorphism $\mathbb{F}_q(\zeta_n)(\mathcal{C}_f) \cong \mathbb{F}_q(\zeta_n)(\mathcal{C}_g)$ corresponds to the class $[\alpha^u]$. This proves the claim. □

For n an odd prime, let $\mathcal{T}_n$ denote the set of n-th power free polynomials in $\mathbb{F}_q[x]$. Let $\mathscr{SE}_{g,n}(\mathbb{F}_q)$ denote the set of superelliptic curves of degree n and genus g over $\mathbb{F}_q$. Then the above claim shows that the fibers of the map:

$$\begin{aligned} \mathcal{T}_n &\to \bigcup_{g \geq 0} \mathscr{SE}_{g,n}(\mathbb{F}_q) \\ f(x) &\mapsto (y^n = f(x)) \end{aligned}$$

have size bounded by $n \mid \mathbb{Z}/n\mathbb{Z}^\times \mid\mid \mathrm{PGL}_2(\mathbb{F}_q) \mid$. As in §2.3, this proves that understanding the proportion of ordinarity in $\mathscr{F}$ is the same as understanding it for the family of superelliptic curves of a fixed degree over $\mathbb{F}_q$.

2.5 Proofs of Main Results

In this section we describe the main results obtained from counting each of the families described above. Our main tool will be the following Tauberian theorem.

Theorem 2.5.1 (See [7], Appendix A)**.** *Let $\{\lambda_n\}_{n \in \mathbb{Z}_{>0}}$ be strictly increasing sequence of positive integers. Let f be the Dirichlet series:*

$$f(s) = \sum_{n=1}^{\infty} c_n \lambda_n^{-s}$$

Further, assume the following:

1. *$f(s)$ converges for $Re(s) > a > 0$.*
2. *f admits a meromorphic continuation to $Re(s) > a - \delta_0 > 0$ for some $\delta_0 > 0$.*

3. *The right-most pole of f is at $s = a$, with multiplicity $b \in \mathbb{N}$. Let $\Theta = \lim_{s\to a} f(s)(s-a)^b$.*

4. *(Technical assumption) There exists a $\kappa > 0$ such that for $Re(s) > a - \delta_0$,*

$$\left|\frac{f(s)(s-a)^b}{s^b}\right| = O((1 + Im(s))^\kappa).$$

Then there exists a (monic) polynomial P of degree $b-1$ such that for any $\delta < \delta_0$, we have,

$$\sum_{\lambda_n < X} c_n = \frac{\Theta}{a(b-1)!} X^a P(\log(X)) + O(X^{a-\delta}).$$

Notation 2.5.2. We will henceforth use the notation $\mid Q \mid$ to denote $q^{\deg(Q)}$, where Q is an irreducible polynomial over $\mathbb{F}_q$. We will denote by $\zeta(s)$, the zeta function of $\mathbb{A}^1_{\mathbb{F}_q}$. Thus $\zeta(s) = \prod_Q (1 - \mid Q \mid^{-s})^{-1}$, where the product is over monic irreducible polynomials over $\mathbb{F}_q$.

2.5.1 Artin-Schreier curves

To recall, the family $\mathscr{F}$ that we are interested in in this section is that of covers $y^p - y = f(x)$, with $f(x) \in \mathcal{S}$, such that the corresponding map $\mathcal{C} \to \mathbb{P}^1$ is unramified over ∞.

We first set up some notation in order to calculate $N(\mathscr{F}, X)$ and $N(\mathscr{F}, g, X)$.

- Define a new invariant: $m = \frac{2g}{p-1} + 2$. By equation (2.4), this is an integer and is equal to

$$\sum_Q \deg(Q)(d_Q + 1).$$

- For any $m \geq 2$, let $a(m)$ be the number of Artin-Schreier covers $\mathcal{C}$ with the above invariant equal to m. Let $b(m)$ be the number of such covers with p-rank g.

- Define

$$N^*(\mathscr{F}, X) = \sum_{q^m < X} a(m) \qquad \text{and} \qquad N^*(\mathscr{F}, g, X) = \sum_{q^m < X} b(m)$$

We will calculate these as an intermediate step towards finding

$$N(\mathscr{F}, X) = N^*(\mathscr{F}, q^2X^{2/(p-1)}) \text{ and } N(\mathscr{F}, g, X) = N^*(\mathscr{F}, g, q^2X^{2/(p-1)}).$$

For this section, define the zeta function:

$$Z(s) = \sum_{\mathcal{C}\in\mathscr{F}} q^{-m(\mathcal{C})s} = \sum_m a(m)q^{-ms}$$

Lemma 2.5.3. *$Z(s)$ converges for $Re(s) > 1$ and has a pole of order $p-1$ at $s = 1$.*

Proof. For a monic irreducible polynomial $Q \in \mathbb{F}_q[x]$, let d_Q be the ramification invariant defined in Section 2.3 if $\mathcal{C} \to \mathbb{P}^1$ is ramified over $\operatorname{div} Q$ and -1 otherwise. Thus

$$m = \sum_Q \deg(Q)(d_Q + 1)$$

where the sum is over all monic irreducible polynomials in $\mathbb{F}_q[x]$. Since this is a sum of local factors, we factor $Z(s)$ as a product of local functions, i.e. $Z(s) = \prod_Q Z_Q(s)$, where Q varies over monic irreducible polynomials in $\mathbb{F}_q[x]$. We can write $Z_Q(s) = \sum_{k\geq 0} c(k)\,|\,Q\,|^{-ks}$. Recall from §2.3 that if $\alpha \in B$ and $Q(\alpha) = 0$, then $d_Q = \deg(f_\alpha)$ as in the partial fraction decomposition of $f(x)$ (2.3). Further, in each f_α, the coefficient of x^{ip} is 0 for each $0 \leq i \leq \lfloor d_Q/p \rfloor$. Since $k = d_Q + 1$,

$$c(k) = \#\{f_\alpha \in \mathbb{F}_{|Q|}[x] \mid \deg(f_\alpha) = k-1, \text{ coefficient of } x^{ip} = 0\}$$

where $k \not\equiv 1 \mod p$ (since $d_Q \not\equiv 0 \mod p$). We write $d_Q = np + i$, with $1 \leq i \leq p-1$. The above discussion gives us that for $k = np + i + 1$,

$$c(k) = (|\,Q\,| - 1)\,|\,Q\,|^{i-1}|\,Q\,|^{n(p-1)}$$

For convenience, we distinguish the cases where $p = 2$ and $p \geq 3$.

For $p = 2$,

$$\begin{aligned} Z_Q(s) &= 1 + \sum_{n=0}^{\infty} (|Q| - 1)\, |Q|^n |Q|^{-s(2n+2)} \\ &= \frac{1 - |Q|^{-2s}}{1 - |Q|^{1-2s}}. \end{aligned}$$

For $p \geq 3$,

$$\begin{aligned} Z_Q(s) &= 1 + \sum_{i=1}^{p-1} \sum_{n=0}^{\infty} (|Q| - 1)\, |Q|^{i-1} |Q|^{n(p-1)} |Q|^{-s(np+i+1)} \\ &= 1 + \left(\frac{(|Q| - 1)\, |Q|^{-2s}}{1 - |Q|^{p-1-ps}} \right) \sum_{i=1}^{p-1} |Q|^{i(1-s)} \\ &= \frac{1 + \sum_{i=0}^{p-3} |Q|^{(i+1)-(i+2)s} - \sum_{i=0}^{p-2} |Q|^{i-(i+2)s}}{1 - |Q|^{p-1-ps}}. \end{aligned}$$

For $p \geq 3$, let

$$\psi_{p,Q}(s) = \left(1 + \sum_{i=0}^{p-3} |Q|^{(i+1)-(i+2)s} - \sum_{i=0}^{p-2} |Q|^{i-(i+2)s} \right) \prod_{i=0}^{p-3} (1 - |Q|^{(i+1)-(i+2)s}).$$

Define

$$\psi_p(s) := \begin{cases} \zeta(2s)^{-1} & if\ p = 2 \\ \prod_Q \psi_{p,Q}(s) & if\ p \geq 3. \end{cases} \tag{2.12}$$

Therefore we have that

$$\prod_Q Z_Q(s) = \psi_p(s) \prod_{i=0}^{p-2} \zeta(s(i+2) - (i+1)).$$

We now claim that there is a constant δ_p (depending only on p) such that $\psi_p(s)$ converges for $Re(s) > \delta_p$. For $p = 2$, this is known classically, since $\zeta(s) = \frac{1}{1-q^{1-s}}$. Thus $\delta_2 = \frac{1}{2}$. For $p \geq 3$, we introduce some shorthand notation for convenience.

Let $a_i = \mid Q \mid^{(i+1)-(i+2)s}$ and $b_i = \mid Q \mid^{i-(i+2)s}$. For $k \in \mathbb{Z}_{\geq 0}$ and $l \in \mathbb{Z}_{>0}$, the $\mid Q \mid^{k-ls}$ will be called *good* if $k + 1 < l$. Observe that:

- The product $\prod_Q (1 - \mid Q \mid^{k-ls})$ converges for $Re(s) > \frac{k+1}{l}$. For a good term, the location of the pole is to the left of $s = 1$.
- If $k_1 + 1 < l_1$ and $k_2 + 1 < l_2$, then $k_1 + k_2 + 1 < l_1 + l_2$. Therefore a product of two good terms is good.
- For any $0 \leq i \leq p-2$, b_i is good. Further, a product of two or more a_i's is a good term.
- For any i, j, the term $a_i b_j = \mid Q \mid^{(i+1)+j-(i+2)s-(j+2)s}$ is good.

Let $p \geq 3$. Then

$$\begin{aligned}
\psi_{p,Q}(s) &= \left(1 + \sum_{i=0}^{p-3} a_i - \sum_{j=0}^{p-2} b_j\right) \prod_{i=0}^{p-3} (1 - a_i) \\
&= \left(1 + \sum_{i=0}^{p-3} a_i - \sum_{j=0}^{p-2} b_j\right) \left(1 - \sum_{i=0}^{p-3} a_i + \text{ good terms }\right) \\
&= \left(1 - \sum_{i=0}^{p-3} a_i + \sum_{i=0}^{p-3} a_i + \text{ good terms}\right) = (1 - \text{ good terms}).
\end{aligned}$$

If, for a moment, we consider $a_0, a_1 \ldots a_{p-3}$ and $b_0, b_1 \ldots b_{p-2}$ as variables, then we see that the set of monomials appearing in the expression for $\psi_{p,Q}(s)$ is finite and independent of Q. Let δ_p be the maximum of the $\frac{k+1}{l}$ such that $\mid Q \mid^{k-ls}$ appears in the simplified expression for $\psi_{p,Q}(s)$.

Then $\delta_p < 1$, and $\psi_p(s)$ converges for $Re(s) > \delta_p$.

Therefore $\prod_Q Z_Q(s) = \psi_p(s) \prod_{i=0}^{p-2} \zeta(s(i+2) - (i+1))$ converges for $Re(s) > 1$ and has a pole of order $p-1$ at $s = 1$. Further, the residue at $s = 1$ is given by

$$\lim_{s \to 1} Z(s)(s-1)^{p-1} = \frac{\psi_p(1)}{\log(q)^{p-1}}.$$

□

To count the number of ordinary curves, we define

$$Z_0(s) = \sum_{\mathcal{C} \in \mathscr{F}_g} q^{-m(\mathcal{C})s} = \sum_m b(m) q^{-ms}.$$

Recall that for such curves, $d_\alpha = 1$ for all α. Therefore, $Z_0(s) = \prod_Q Z_{0,Q}(s)$, where the local factors are given by:

$$Z_{0,Q}(s) = 1 + (|Q| - 1) \, |Q|^{-2s} \, .$$

Lemma 2.5.4. *$Z_0(s)$ converges for $Re(s) > 1$ and has a simple pole at $s = 1$.*

Proof. Note that

$$(1 + |Q|^{1-2s} - |Q|^{-2s})(1 - |Q|^{1-2s}) = 1 - |Q|^{-2s} - |Q|^{2-4s} + |Q|^{1-4s}$$

and

$$\phi(s) := \prod_Q (1 - |Q|^{-2s} - |Q|^{2-4s} + |Q|^{1-4s})$$

converges for $Re(s) > 3/4$. Therefore $Z_0(s) = \phi(s)\zeta(2s-1)$ converges for $Re(s) > 1$ and has a simple pole at $s = 1$. Further, the residue at $s = 1$ is

$$\lim_{s \to 1} Z_0(s)(s-1) = \frac{\phi(1)}{\log(q)}.$$

□

Proposition 2.5.5. *For any $\delta > 0$,*

$$N^*(\mathscr{F}, X) = \frac{\psi_p(1)}{\log(q)} X(\log_q(X))^{p-2} + O(X^{1-\delta}),$$
$$N^*(\mathscr{F}, g, X) = \frac{\phi(1)}{\log(q)} X + O(X^{1-\delta}).$$

Proof. This follows from Theorem 2.5.1 applied to the results of Lemmas 2.5.3 and 2.5.4, since $\zeta(s)$ has a meromorphic continuation to the entire complex plane. □

Corollary 2.5.6. *For any $\delta > 0$,*

$$N(\mathscr{F}, X) = \frac{\psi_p(1)}{\log(q)} q^2 X^{2/(p-1)} (\log_q(X^{2/(p-1)}))^{p-2} + O(X^{\frac{2}{p-1}-\delta}),$$
$$N(\mathscr{F}, g, X) = \frac{\phi(1)}{\log(q)} q^2 X^{2/(p-1)} + O(X^{\frac{2}{p-1}-\delta}).$$

In particular, the probability that an Artin-Schreier cover unramified over ∞ is ordinary is

$$\begin{aligned} \phi(1)\zeta(2) \qquad & if\ p = 2, \\ 0 \qquad & if\ p \geq 3. \end{aligned}$$

Proof. $N(\mathscr{F}, X) = N^*(\mathscr{F}, q^2 X^{2/(p-1)})$.

□

Remark 2.5.7. We now make some concluding remarks about counting Artin-Schreier curves. Data associated to this subsection can be found in Section 2.6.

1. If we modify $\mathscr{F}$ to include the covers ramified over ∞, we must modify the partial fraction decomposition in (2.3) to:
$$f(x) = \sum_{\substack{\alpha \in B \\ \alpha \neq \infty}} f_\alpha(x_\alpha) + g(x).$$

Here $g(x) \in \mathbb{F}_q[x]$ is a polynomial that, like the other f_α's, has degree coprime to p and for each $0 \leq i \leq \lfloor \deg(g)/p \rfloor$, the coefficient of x^{ip} in $g(x)$ is 0. This manifests as a change in the zeta functions $Z(s)$ and $Z_0(s)$ defined in the above discussion by factors that we will call $Z_\infty(s)$ and $Z_{0,\infty}(s)$ respectively. That is, we write $Z(s) = Z_\infty(s) \prod_Q Z_Q(s)$ and $Z_0(s) = Z_{0,\infty}(s) \prod_Q Z_{0,Q}(s)$. Both these factors only affect the residues of $Z(s)$ and $Z_0(s)$ and not the order of growth, which means that for $p \geq 3$, the probability of ordinarity for the modified family is still 0. For $p = 2$,

$$Z_\infty(s) = 1 + q^{-1} \qquad \text{and} \qquad Z_{0,\infty}(s) = 1 - q^{-1} + q^{-2}.$$

Therefore the probability of ordinarity in the modified family is

$$\left(\frac{1 - q^{-1} + q^{-2}}{1 + q^{-1}} \right) \phi(1)\zeta(2) = 1 - 3q^{-1} + 6q^{-2} + O(q^{-3}). \tag{2.13}$$

2. Recall from Section 2.3.2, that the probability that an Artin-Schreier *curve* is ordinary, is bounded above by the quantity,

$$\lim_{X \to \infty} C(q) \log_q(X)^4 \frac{N(\mathscr{F}, g, X)}{N(\mathscr{F}, X)}.$$

Since the order of growth of $N(\mathscr{F}, g, X)$ is $X^{2/(p-1)}$, and the that of $N(\mathscr{F}, X)$ is $X^{2/(p-1)} \log(X)^{p-2}$, this quantity is 0 whenever $p \geq 7$. The geometric description of the Artin-Schreier locus leads us to believe that the same might be true for $p = 3, 5$, as explained in part (4) of this remark. However a proof for these cases requires more work.

3. Recall that if the Jacobian of a curve behaves randomly in the sense of [6], the heuristics predict that the probability of that a curve is ordinary is

$$\prod_{i=1}^{\infty} (1 + q^{-i})^{-1}.$$

Corollary 2.5.6 and the previous remark prove that the Jacobian of an Artin-Schreier

curve does not behave randomly in the sense of [6]. For $p \geq 3$ this is clear. For $p = 2$, elementary calculations show that the constants are not equal. In fact,

$$\prod_{i=1}^{\infty}(1+q^{-i})^{-1} = 1 - q^{-1} - q^{-3} + q^{-4} + O(q^{-5}).$$

One must remember, however, that since we are counting covers instead of curves, as observed in Section 2.3, that this does not disprove the heuristic for isomorphism classes of Artin-Schreier *curves* in characteristic 2.

4. Theorem 2.3.2 implies that the ordinary locus intersects exactly one irreducible component of $\mathscr{AS}_g$, namely the one corresponding to the partition $\{2, 2, \ldots 2\}$ of $d+2 = \frac{2g}{p-1}+2$. On the other hand, from work in [23], we know that the Artin-Schreier locus is equidimensional, each component having dimension $d-1$. In particular, this implies that for $p \geq 3$, the proportion of components intersecting the ordinary locus goes to 0 as $g \to \infty$. Indeed, for $p \geq 3$, let $A = \{2, 3, \ldots p\}$ and let $p_A(n)$ denote the number of partitions of an integer n into integers from the set A. Then the number of components of dimension $d-1$ is $p_A(d+2)$. As $n \to \infty$, $p_A(n) \sim Kn^{p-2}$ for some constant K. This might be a somewhat satisfying geometric explanation, especially for those taken aback by the fact that counting squarefree rational functions in this order gives a proportion of 0.

2.5.2 Superelliptic curves in characteristic 2

For this section, we use the notation of Section 2.4. We are interested in counting covers in the family $\mathscr{F}$ of covers $y^n = f(x)$ over a field $\mathbb{F}_q$ of characteristic 2, where

- n is prime,
- $f(x) \in \mathbb{F}_q[x]$ is n-th power free.

For convenience, we count by $q^{\mathbf{m}}$ instead of q^g where

$$\mathbf{m} := \frac{2g}{n-1} + 2$$

is the number of points in $\mathbb{P}^1(\bar{k})$ over which the curve given by $y^n = f(x)$ is ramified. Since n is fixed in the entire discussion, this will not change the order of counting significantly. Define $N^*(\mathscr{F}, X)$ as the set of curves in $\mathscr{F}$ with $q^{\mathbf{m}} < X$ and $N^*(\mathscr{F}, g, X)$ similarly. From the definition of $\mathbf{m}$, it follows that

$$N(\mathscr{F}, X) = N^*(\mathscr{F}, q^2 X^{2/(n-1)}) \quad \text{and} \quad N(\mathscr{F}, g, X) = N^*(\mathscr{F}, g, q^2 X^{2/(n-1)}). \tag{2.14}$$

Define

$$\mathcal{F}_{e_1,e_2\cdots e_r} = \{F_1 F_2^2 \cdots F_r^r \mid F_i \in \mathbb{F}_q[x] \text{ monic, squarefree and mutually coprime, } \deg(F_i) = e_i\}.$$

When we write $\mathbf{m} = \sum_{i=1}^{n-1} e_i$, we will be interested in the case when there are e_i points ramifying to degree i. This is the same as the notion defined in Remark 2.4.6. To express this concretely in terms of polynomials, it is best to use an example. For instance, for a curve given by $y^3 = F_1(x)(F_2(x))^2$, where $F_1(x)(F_2(x))^2 \in \mathcal{F}_{2,4}$, there are 2 points that occur with degree 1 and 4 that occur with degree 2. If on the other hand, the curve is given by $y^3 = F_1(x)(F_2(x))^2$, where $F_1(x)(F_2(x))^2 \in \mathcal{F}_{3,2}$, there are 3 points that occur with degree 1 and 3 that occur with degree 2 (since $n_\infty = 2$, the curve is ramified over $\infty \in \mathbb{P}^1$ to degree 2).

Proposition 2.5.8. *Consider the set $S_{\mathbf{m}}$ of superelliptic curves with the number of ramified points $\mathbf{m} = \sum_{i=1}^{n-1} e_i$, such that there are e_i points that ramify to degree i. Then the size of $S_{\mathbf{m}}$ is:*

$$|\,\mathcal{F}_{e_1,e_2\cdots e_{n-1}}\,| + \sum_{i=1}^{n-1} |\,\mathcal{F}_{e_1,\cdots e_i-1,\cdots e_{n-1}}\,|$$

Proof. Let $\mathcal{C} \in \mathscr{F}$, such that $\mathcal{C} \to \mathbb{P}^1$ is ramified over $\mathbf{m}$ points in $\mathbb{P}^1(\overline{\mathbb{F}}_q)$. If the map is not ramified over ∞, then $\mathcal{C} \in \mathcal{F}_{e_1,e_2\ldots e_{n-1}}$. If it is ramified over ∞ and $n_\infty = i$, then $\mathcal{C} \in \mathcal{F}_{e_1,\ldots e_i-1,\ldots e_{n-1}}$. □

In the above proposition, imposing the condition $\mathbf{m} = \sum_{i=1}^{n-1} e_i$, with e_i points occuring with

degree i, implies that $\sum_{i=1}^{n-1} ie_i \equiv 0 \mod n$. Therefore we are interested in the following quantity:

$$\sum_{q^{\mathbf{m}}<X} \left(|\mathcal{F}_{e_1,e_2\cdots e_{n-1}}| + \sum_{i=1}^{n-1} |\mathcal{F}_{e_1,\cdots e_i-1,\cdots e_{n-1}}| \right) \tag{2.15}$$

where the sum is over tuples $(e_1, e_2 \dots e_{n-1})$ such that $\sum_{i=1}^{n-1} ie_i \equiv 0 \mod n$. Further, observe that for a fixed $1 \leq i \leq n-1$,

$$\sum_{\substack{(e_j), q^{e_1+e_2\cdots e_{n-1}}<X \\ \sum je_j \equiv 0 \mod n}} |\mathcal{F}_{e_1,\cdots e_i-1,\cdots e_{n-1}}| = \sum_{\substack{(d_j), q^{d_1+d_2\cdots d_{n-1}}<X/q \\ \sum jd_j \equiv i \mod n}} |\mathcal{F}_{d_1,\cdots d_i,\cdots d_{n-1}}| .$$

Therefore Equation (2.15) can be rewritten as

$$\left(\sum_{q^{e_1+e_2\cdots e_{n-1}}<X/q} |\mathcal{F}_{e_1,e_2\dots e_{n-1}}| \right) + \sum_{\substack{X/q<q^{e_1+\cdots e_{n-1}}<X \\ \sum ie_i \equiv 0 \mod n}} |\mathcal{F}_{e_1,e_2\dots e_{n-1}}|$$

where the first sum is over all tuples $(e_1, e_2 \dots e_{n-1})$ with $q^{\sum e_i} < X/q$. Thus the number of superelliptic curves with $q^{\mathbf{m}} < X$ is bounded below by the quantity

$$\left(\sum_{q^{e_1+e_2\cdots e_{n-1}}<X/q} |\mathcal{F}_{e_1,e_2\dots e_{n-1}}| \right).$$

For our proof, we only need a lower bound for the total number of superelliptic curves. Therefore, it suffices to estimate this quantity.

Notation 2.5.9. From now on, a sum of the form

$$\sum_{q^{e_1+e_2\cdots e_r}<X}$$

will denote a sum over all tuples of non-negative integers $(e_1, e_2 \dots e_r)$, with $q^{\sum e_i} < X$.

For any non-negative integer m, let $a(m) = \sum_{e_1+e_2\ldots+e_{n-1}=m} |\mathcal{F}_{e_1,e_2\ldots e_{n-1}}|$. Define:

$$Z(s) = \sum_{m\geq 0} a(m)q^{-ms}$$

One way to think about an element of $\mathcal{F}_{e_1,e_2\ldots e_{n-1}}$ is to say that we are considering a polynomial

$$f(x) = \prod_{i=1}^{n-1}(F_i(x))^i$$

such that $H := \prod_{i=1}^{n-1} F_i(x)$ is squarefree. We will use this characterization to calculate $Z(s)$.

Consider a squarefree polynomial H. Let $H = \prod h_j$ be its factorization into irreducible polynomials. We want to count the number of ways in which H can be written as a product of squarefree polynomials $\prod_{i=1}^{n-1} F_i$. For each factor h_j, there are $n-1$ choices of squarefree polynomial that it could divide. Therefore, the number of factorizations $H = \prod_{i=1}^{n-1} F_i$ is

$$(n-1)^{\omega(H)}$$

where $\omega(H)$ = the number of distinct irreducible factors of H. Therefore,

$$Z(s) = \sum_{H \text{ sq. free}} (n-1)^{\omega(H)} |H|^{-s} = \prod_{Q}(1 + (n-1)|Q|^{-s}).$$

Note 2.5.10. Let $\Phi_k(s) = \prod_Q(1 + k|Q|^{-s})$. Then $\Phi_k(s)\zeta(s)^{-k}$ is a function that converges for $Re(s) > 1/2$. We will denote $\Phi_k(s)\zeta(s)^{-k}$ by $\phi_k(s)$.

Proposition 2.5.11. *As $X \to \infty$,*

$$N^*(\mathscr{F}, X) \geq \frac{\phi_{n-1}(1)}{q\log(q)(n-2)!}X(\log_q(X))^{n-2} + O(X(\log(X))^{n-3}).$$

Proof. Note that

$$\begin{aligned} Z(s) &= \Phi_{n-1}(s) \\ &= \zeta(s)^{n-1}\phi_{n-1}(s). \end{aligned}$$

This function has a pole of order $n-1$ at $s=1$. Thus, the Tauberian theorem implies that:

$$\sum_{m<X/q} a(m) = \frac{\phi_{n-1}(1)}{q\log(q)(n-2)!}X(\log_q(X))^{n-2} + O(X(\log(X))^{n-3}).$$

This provides a lower bound for $N^*(\mathscr{F}, X)$. □

Corollary 2.5.12. *The number of superelliptic covers with invariant* $\mathbf{m}$ *such that* $q^{\mathbf{m}} < X$ *is bounded below by*

$$\kappa_n(q)X^{2/(n-1)}\log_q(X^{2/(n-1)})^{n-2} + O(X^{2/(n-1)}\log(X^{2/(n-1)})^{n-3})$$

where

$$\kappa_n(q) = \frac{q\phi_{n-1}(1)}{\log(q))(n-2)!}.$$

Proof. This follows from the fact that $N(\mathscr{F}, X) = N^*(\mathscr{F}, q^2X^{2/(n-1)})$. □

Upper bounds for $N^*(\mathscr{F}, g, X)$

In this subsection, we find an upper bound for the quantity $N^*(\mathscr{F}, g, X)$, as defined in Equation (2.14). We will maintain the notation of Section 2.4.

Suppose we consider covers with $\mathbf{m} = \sum_{i=1}^{n-1} e_i$ ramification points, e_i points occurring with degree i. Using the criterion for ordinarity in Proposition 2.4.4, we can derive the following conditions on $\mathcal{F}_{e_1,e_2\ldots e_{n-1}}$:

1. If $n_\infty = 0$, $\deg(f_i) = \deg(f_{n-i})$. Note that in this case, the cover belongs to $\mathcal{F}_{e_1,e_2\ldots e_{n-1}}$ with $\deg(f_i) = e_i$. Therefore the condition for ordinarity implies that there are

$$|\,\mathcal{F}_{e_1,\ldots e_{\frac{n-1}{2}},e_{\frac{n-1}{2}}\ldots e_1}\,|$$

curves of such kind over $\mathbb{F}_q$.

2. If $n_\infty = i$, then $\mathcal{C} \in \mathcal{F}_{e_1,\ldots e_i-1\ldots e_{n-1}}$ with $\deg(f_j) = e_j$ for $j \neq i$ and $\deg(f_i) = e_i - 1$. Further, the condition for ordinarity gives: for $j \neq i, n-i$, $\deg(f_j) = \deg(f_{n-j})$ and therefore $e_j = e_{n-j}$. Also, $\deg(f_i) + 1 = \deg(f_{n-i})$ implies $e_i = e_{n-i}$. Therefore, the number of such curves is

$$|\,\mathcal{F}_{e_1,e_2\ldots,e_i-1,\ldots e_{\frac{n-1}{2}},e_{\frac{n-1}{2}}\ldots,e_i\ldots e_2,e_1}\,|$$

if $i \leq \frac{n-1}{2}$, and

$$|\,\mathcal{F}_{e_1,e_2\ldots,e_i,\ldots e_{\frac{n-1}{2}},e_{\frac{n-1}{2}}\ldots,e_i-1\ldots e_2,e_1}\,|$$

if $i > \frac{n-1}{2}$.

We are interested in the size

$$N^*(\mathscr{F}, g, X) = \sum_{q^{\mathbf{m}}<X}^{*} \left(|\,\mathcal{F}_{e_1,e_2\cdots e_{n-1}}\,| + \sum_{i=1}^{n-1} |\,\mathcal{F}_{e_1,\cdots e_i-1,\cdots e_{n-1}}\,| \right) \tag{2.16}$$

where the sum is now over tuples $(e_1, e_2 \ldots e_{n-1})$ that satisfy the ordinarity criterion $e_i = e_{n-i}$. Note that for such a tuple, the condition $\sum_{i=1}^{n-1} ie_i \equiv 0 \mod n$ is satisfied automatically. We now proceed to find an upper bound on this quantity, using a result of Bucur et. al. in [5] that we will recall below. Let

$$L_{n-2} = \prod_{j=1}^{n-2} \prod_{Q} \left(1 - \frac{j}{(|\,Q\,|+1)(|\,Q\,|+j)}\right)$$

where the product is over monic irreducible polynomials $Q \in \mathbb{F}_q[x]$.

Theorem 2.5.13 ([5], Prop 4.3)**.** *Fix a tuple of positive integers (e_1, e_2). Then, for any $\epsilon > 0$ and as q gets large,*

$$|\mathcal{F}_{e_1,e_2}| = \frac{L_1 q^{e_1+e_2}}{\zeta(2)^2}\left(1 + O(q^{-e_2(1-\epsilon)} + q^{-e_1/2})\right)$$

Remark 2.5.14. The number of monic polynomials of degree d in $\mathbb{F}_q[x]$ is q^d and the proportion of these that are squarefree is $(1 - 1/q)$. One might expect, similarly, that the proportion of pairs of monic polynomials of degrees (e_1, e_2) that are squarefree and coprime, also form a positive proportion of the total number of pairs of monic polynomials, $q^{e_1+e_2}$. The above theorem shows that this is indeed the case. The next proposition shows that the same is true for $(e_1, e_2 \dots e_{n-1})$ for any odd prime n, although with a weaker error term.

For the following proposition, we refer the reader to [5], Corollary 7.2.

Proposition 2.5.15. *Fix a tuple of positive integers $(e_1, e_2 \dots e_{n-1})$. Fix an $\epsilon > 0$. Then, as q gets large,*

$$|\mathcal{F}_{(e_1,e_2\dots e_{n-1})}| = \frac{L_{n-2} q^{e_1+e_2\dots e_{n-1}}}{\zeta(2)^{n-1}}(1+O(q^{\epsilon(e_2+\dots e_{n-1}+q)+(1-\epsilon)q}(q^{-e_2}+\cdots q^{-e_{n-1}})+q^{-(e_1-3q)/2}))$$

Proof. Consider the expression given in [5], Corollary 7.2. Summing the expression over all possible partitions $m = k_1 + k_2 \dots + k_{n-1}$ gives:

$$\frac{L_{n-2} q^{e_1+e_2\dots e_{n-1}}}{\zeta(2)^{n-1}}\left(\frac{n-1}{q+n-1}\right)^m\left(\frac{q}{(q+n-1)(q-1)}\right)^{q-m}$$
$$\times (1 + O(q^{\epsilon(e_2+\dots e_{n-1}+q)+(1-\epsilon)m}(q^{-e_2} + q^{-e_3}\cdots q^{-e_{n-1}}) + q^{-(e_1-m)/2+q})).$$

Summing over all possibilities of m now gives the result. □

Parsing these propositions tells us that for large enough q (depending only on n),

$$| \mathcal{F}_{(e_1,e_2\ldots e_{n-1})} | \leq K_1 q^{e_1+e_2\ldots e_{n-1}} + K_2 q^{e_1/2+e_2\ldots e_{n-1}} + \sum_{i=2}^{n-1} K_{3,i} q^{e_1+\ldots \epsilon e_i + \ldots e_{n-1}}$$

where K_1, K_2 and the $K_{3,i}$'s depend on ϵ, q and n, but are independent of the e_j's. Since for $\epsilon < 1$ the first term in the above expression is the largest, we let $K = \max(K_1, K_2, K_{3,2} \ldots K_{3,n-1})$ and so for large enough q:

$$| \mathcal{F}_{(e_1,e_2\ldots e_{n-1})} | \leq K q^{e_1+e_2\ldots e_{n-1}}.$$

Thus Equation (2.16) implies that

$$N^*(\mathscr{F}, g, X) \leq K \left(\frac{q+n-1}{q} \right) \left(\sum_{q^{2(e_1+e_2\cdots e_{(n-1)/2})} < X} q^{2(e_1+e_2\ldots e_{(n-1)/2})} \right). \tag{2.17}$$

The following lemma will be used to find an upper bound for the expression above.

Lemma 2.5.16. *As X gets large,*

$$\sum_{q^{e_1+e_2\cdots e_r} < X} q^{e_1+e_2\ldots e_r} = O(X \log(X)^{r-1}).$$

Here, the implied constants depend on q and r.

Proof. Consider the expression:

$$\left(\frac{1}{1-qT} \right)^r$$

The coefficient of T^m in this expression is $\sum_{e_1+e_2\ldots e_r = m} q^{e_1+e_2\cdots+e_r}$. On the other hand, by the

binomial theorem, the coefficient of T^m in $(1-qT)^{-r}$ is: $\binom{r+m-1}{r-1} q^m$. Further,

$$\binom{r+m-1}{r-1} \leq \frac{(m+r)^{r-1}}{(r-1)!}.$$

Therefore, we have that

$$\begin{aligned}
\sum_{q^{e_1+e_2\cdots e_r}<X} q^{e_1+e_2\ldots e_r} &= \sum_{q^m<X} \sum_{e_1+e_2\ldots e_r=m} q^{e_1+e_2\ldots e_r} \\
&= \sum_{q^m<X} \binom{r+m-1}{r-1} q^m \\
&\leq \frac{(2r)^{r-1}}{(r-1)!} \sum_{\substack{m<r \\ q^m<X}} q^m + \sum_{\substack{m\geq r \\ q^m<X}} \frac{(2m)^{r-1}}{(r-1)!} q^m \\
&= D_r X \log(X)^{r-1} + O(X \log(X)^{r-2}),
\end{aligned}$$

where the last step follows by Euler Summation. This proves the lemma. □

Proposition 2.5.17. *For large enough q,*

$$N^*(\mathscr{F}, g, X) = O(X \log(X)^{\frac{n-3}{2}}).$$

Hence, $N(\mathscr{F}, g, X) = O(X^{2/(n-1)} \log(X)^{\frac{n-3}{2}})$, where the implied constants depend on q and n.

Proof. To obtain the first statement, we use Equation (2.17):

$$N^*(\mathscr{F}, g, X) \leq K \left(\frac{q+n-1}{q}\right) \left(\sum_{q^{2(e_1+e_2\ldots e_{(n-1)/2})}<X} q^{2(e_1+e_2\ldots e_{(n-1)/2})} \right)$$

and Lemma 2.5.16, with q replaced by q^2. The second part of the statement follows from the

fact that $N(\mathscr{F}, g, X) = N^*(\mathscr{F}, g, q^2 X^{2/(n-1)})$.

□

We remind the reader here that for the quantity that we are interested in, namely the probability that a superelliptic curve is ordinary, q and n are fixed. Therefore, the fact that the implied constants above depend on q and n will make no difference to the theorem below.

Theorem 2.5.18. *The probability that a superelliptic curve $y^n = f(x)$ over $\mathbb{F}_{2^r}$ with n prime and r large enough depending only on n, is ordinary, is zero. That is,*

$$\lim_{X\to\infty} \frac{N(\mathscr{F}, g, X)}{N(\mathscr{F}, X)} = 0$$

Proof. By Proposition 2.5.17, the numerator $N(\mathscr{F}, g, X)$ is bounded above by the quantity

$$X^{2/(n-1)} \log(X)^{\frac{n-3}{2}}.$$

By Corollary 2.5.12, the denominator grows faster than $X^{2/(n-1)} \log(X)^{n-2}$. This proves the theorem.

□

Remark 2.5.19. It is interesting to note that for a given g, the space of superelliptic curves of degree n and genus g decomposes over $\overline{\mathbb{F}_q}$ into irreducible components that correspond to partitions of $\mathbf{m} = \sum_{i=1}^{n-1} e_i$ such that $\sum_{i=1}^{n-1} i e_i \equiv 0 \mod n$. The ordinary locus intersects a small proportion of these components. For $n = 3$, for instance, it only intersects one component. A similar thing was true for the Artin-Schreier locus $\mathscr{AS}_g$. For fixed p-rank s, one can obtain a combinatorial description of the components contained in the stratum $\mathscr{AS}_{g,s}$. One can ask if a similar result holds for superelliptic curves in odd characteristic.

2.6 Numerical Data on Artin-Schreier curves

We conclude this chapter by listing some values of constants computed in the previous section. Recall that

$$P(\mathscr{AS}, g) := \left(\frac{1 - q^{-1} + q^{-2}}{1 + q^{-1}}\right) \phi(1)\zeta(2)$$

is the probability that an Artin-Schreier curve in characteristic 2 is ordinary (Corollary 2.5.6). For brevity, we let $\varphi(q) = \prod_{i=1}^{\infty}(1 + q^{-i})^{-1}$, the constant predicted in [6].

p	q	$\phi(1)$	$P(\mathscr{AS}, g)$	$\varphi(q)$
2	2	0.314148	0.314148	0.419422
2	4	0.593976	0.514777	0.737512
2	8	0.776577	0.702617	0.873264
2	16	0.882162	0.833730	0.937270
2	32	0.939367	0.911820	0.968720

Table 2: Proportion of ordinary Artin-Schreier curves in characteristic 2

Chapter 3

Counting elliptic curves with a rational N-isogeny

3.1 Introduction

We quickly recall some of the notation from Chapter 1. Throughout this chapter, E will denote an elliptic curve over $\mathbb{Q}$. We say that E has a rational N-isogeny if there is an isogeny $\phi : E \to E'$ such that $\mathrm{Ker}(\phi)(\bar{\mathbb{Q}}) \cong \mathbb{Z}/N\mathbb{Z}$ and if $\mathrm{Ker}(\phi)$ is stable under the action of the absolute Galois group of $\mathbb{Q}$. An elliptic curve E over $\mathbb{Q}$ has a unique minimal Weierstrass equation $y^2 = x^3 + Ax + B$ where $A, B \in \mathbb{Z}$ and $\gcd(A^3, B^2)$ is not divisible by any 12th power. Define the *naive height* of E to be $\mathrm{ht}(E) = \max\{|A|^3, |B|^2\}$.

Notation 3.1.1. For two functions $f, g : \mathbb{R} \to \mathbb{R}$, we say that $f(X) \asymp g(X)$ if there exist positive constants K_1 and K_2 such that $K_1 g(X) \leq f(X) \leq K_2 g(X)$. For a real number X and positive integer N, define

$$\mathcal{N}(N, X) = \#\{E/\mathbb{Q} \mid \mathrm{ht}(E) < X, E \text{ has a rational } N\text{-isogeny}\},$$

where we count elliptic curves up to isomorphism over $\mathbb{Q}$.

Recall that we want to find a function $h_N(X)$ such that $\mathcal{N}(N, X) \asymp h_N(X)$. We will often call this the *asymptotic growth rate* of the the function $\mathcal{N}(N, X)$. Our goal is to prove Theorem 1.2.5.

An elliptic curve over $\mathbb{Q}$ equipped with a rational N-isogeny gives rise to a $\mathbb{Q}$-rational point on the modular curve $\mathcal{X}_0(N)$ (defined in Section 3.1.1). Thus, up to a constant equal to the degree

of the forgetful map $\mathcal{X}_0(N) \to \mathcal{X}_0(1)$, the counting function $\mathcal{N}(N, X)$ is the number of points of bounded naive height on the modular curve $\mathcal{X}_0(N)$. In [24], Mazur gave the list of prime N for which there exists an elliptic curve over $\mathbb{Q}$ with a rational N isogeny. Various cases of composite N were proved by several people, and a complete list of N for which there exists an elliptic curve over $\mathbb{Q}$ with a rational N-isogeny can be found in work of Kenku ([20]). For the modular curves that we are interested in, the existence of such a non-cuspidal rational point is not in question. For $N \leq 10$ and $N = 12, 13, 16, 18, 25$, the coarse space of $\mathcal{X}_0(N)$ is isomorphic to $\mathbb{P}^1$ (see for e.g. the genus computation in [11]). In particular, there are infinitely many j-invariants j such that there is an elliptic curve E with a rational N-isogeny and $j(E) = j$. Further, the fact every point of $\mathcal{X}_0(N)$ has automorphism group of size at least two implies that there are infinitely many elliptic curves of a given j invariant that have a rational N-isogeny. This makes counting them significantly more challenging.

3.1.1 Modular curves

Let N be a positive integer. Let $\mathcal{Y}_0(N)$ denote the modular curve such that for a $\mathbb{Z}[\frac{1}{N}]$ scheme S,

$$\mathcal{Y}_0(N)(S) = \{(E/S, C/S) \mid C \cong_S \mathbb{Z}/N\mathbb{Z}\}$$

where E/S is an elliptic curve over S, C is a sub-group scheme of E defined over S, and the pair is taken up to isomorphism. Let $\mathcal{X}_0(N)$ denote the compactification of $\mathcal{Y}_0(N)$ in the sense of [10]. Every point of this moduli space possesses the extra automorphism $[-1]$, since the automorphism of E sending $P \mapsto -P$ also induces an automorphism of C. Let $B\mu_n$ over a scheme S be the stacky quotient $[S/\mu_n]$ parametrizing principal μ_n-bundles. Thus $\mathcal{X}_0(N)$ is a stack over $\mathbb{Z}$ with generic inertia stack $B\mu_2$.

Let $\mathcal{Y}_1(N)$ denote the curve whose points are given by:

$$\mathcal{Y}_1(N)(S) = \{(E/S, P/S) \mid N \cdot P = 0\}$$

where E/S is an elliptic curve over S, $P \in E(S)$ is a point of order N, and the pair is taken

up to isomorphism. Let $\mathcal{X}_1(N)$ denote the compactification of $\mathcal{Y}_1(N)$. For $N \geq 5$, $\mathcal{X}_1(N)$ is a scheme. There is a natural map $\Phi_N : \mathcal{X}_1(N) \to \mathcal{X}_0(N)$ which sends (E, P) to $(E, \langle P \rangle)$, where $\langle P \rangle$ denotes the subgroup of E generated by P. We remark here that the cusps of modular curves also have a moduli interpretation. They paramterize 'generalized' elliptic curves with $\Gamma_0(N)$ or $\Gamma_1(N)$ structures. For a more detailed exposition on these, we refer the reader to [10] or [40]. A short description is given in the appendix.

Definition 3.1.2. Let $\mathcal{M}$ denote any modular curve. For any point $S \to \mathcal{M}$, let $p : E \to S$ denote the corresponding elliptic curve. The *Hodge bundle* $\lambda_{\mathcal{M}}$ is the line bundle on $\mathcal{M}$ such that $(\lambda_{\mathcal{M}})_S = p_* \omega_{E/S}$. For ease of notation, we will omit the $\mathcal{M}$ in $\lambda_{\mathcal{M}}$ whenever the underlying modular curve is clear from context. Further, it follows from the definition of the Hodge bundle that if $\mathcal{M}$ is a modular curve parametrizing elliptic curves with a certain level structure, and $\psi : \mathcal{M} \to \mathcal{X}_0(1)$ is the forgetful map, then $\lambda_{\mathcal{M}} = \psi^* \lambda_{\mathcal{X}_0(1)}$.

Modular forms of weight k and level N are sections of the k-th power of the Hodge bundle on $\mathcal{X}_0(N)$. The coefficients A and B in the Weierstass equation $y^2 = x^3 + Ax + B$ are, up to a scalar, the Eisenstein series E_4 and E_6 on $\mathcal{X}_0(1)$ respectively (see for e.g. [32]). Thus A^3 and B^2 are sections of the 12th power of the Hodge bundle on any modular curve. Thus counting elliptic curves of bounded naive height is the same as counting elliptic curves of bounded height with respect to $\lambda^{\otimes 12}$ on any modular curve that is a scheme. We will see later, that the same is true for moduli stacks.

Definition 3.1.3. Let M denote the coarse space of a modular curve $\mathcal{M}$. When $M \cong \mathbb{P}^1$, its function field is freely generated by a single element; this element is called a *hauptmodul*. These hauptmoduln parametrize elliptic curves with a given level structure, and can be used to write equations for modular curves.

3.1.2 Two approaches to counting points on a stack

We take two approaches to counting rational points on $\mathcal{X}_0(N)$. The first is based on the work of Harron and Snowden in their paper [17], and uses the fact $\mathcal{X}_0(N)$ has an open substack $\mathcal{Y}$ such that every point on $\mathcal{Y}$ has an automorphism group of size exactly 2. The second uses

the theory of heights on stacks developed in [13] to show that the naive height comes from geometry. We outline the approaches below.

The work of Harron and Snowden and counting points on stacky modular curves

In [17], Harron and Snowden ask, for a given group G from Mazur's list in [24, Theorem 2] , how many elliptic curves have $E(\mathbb{Q})_{tors} \cong G$? They compute the asymptotic growth rate of this quantity for each G. Part of their framework involves counting elliptic curves in families. Let M be a modular curve parametrizing elliptic curves with prescribed level structure. Suppose further that there exists a universal family $\mathcal{M}$ over M. That is there is a map $\mathcal{M} \to M$ such that any family of elliptic curves E/S corresponding to a point $S \to M$ is pulled back from $\mathcal{M}$. In particular, if $M \cong \mathbb{P}^1$, then one can find an equation:

$$y^2 = x^3 + f(t)x + g(t)$$

such that every elliptic curve over $\mathbb{Q}$ with prescribed level structure is isomorphic to one of this form for some $t \in \mathbb{Q}$. To count such elliptic curves with bounded height therefore, one must count the set of pairs $(A, B) \in \mathbb{Z}^2$ such that

- $4A^3 + 27B^2 \neq 0$,
- $\max\{|A|^3, |B|^2\} < X$,
- $\gcd\{A^3, B^2\}$ is not divisible by any 12th powers, and
- $\exists u, t \in \mathbb{Q}$ such that $u^4 f(t) = A$ and $u^6 g(t) = B$.

In [17], the authors give the asymptotic growth rate for such pairs (A, B) as an explicit power of X in the special case that f and g are coprime, and

$$\max\left\{\frac{\deg(f)}{4}, \frac{\deg(g)}{6}\right\} =: \frac{n}{m}$$

in lowest terms with either m or $n = 1$. As an example, the curves $\mathcal{X}_1(N)$ for $N = 3, 4, 5, 6, 7, 8, 9, 10, 12$ all admit such a universal family.

Now of course, $\mathcal{X}_0(N)$ does not admit such a universal family (Proposition 3.2.4). However, for certain N, one can construct a double cover of $\mathcal{X}_0(N)$ that does. This double cover satisfies the property that any elliptic curve with a rational N-isogeny is a quadratic twist of one arising as a rational point on the cover. In particular, this leads to a similar counting problem as above, just with the last condition replaced by:

$$\exists u, t \in \mathbb{Q} \text{ such that } u^2 f(t) = A \text{ and } u^3 g(t) = B.$$

This puts us back in the framework of [17, Theorem 4.1], with some extra conditions on the degree, as made precise in Section 3.4. We use this to compute $h_N(X)$ for $N \in \{3, 4, 6, 8, 9, 12, 16, 18\}$.

Embedding curves into a weighted projective space

Let V be a projective variety defined over $\mathbb{Q}$. Let $\mathcal{L}$ be an ample line bundle on it. Then one can embed $V \overset{f}{\hookrightarrow} \mathbb{P}^M$ via $\mathcal{L}^{\otimes n}$ for some n, i.e. $\mathcal{O}_{\mathbb{P}^M}(1) \cong \mathcal{L}^{\otimes n}$. Let $x \in V(\mathbb{Q})$. Then $f(x) \in \mathbb{P}^M(\mathbb{Q})$ can be extended to an integral point $\overline{f(x)} \in \mathbb{P}^M(\mathbb{Z})$. Writing $\overline{f(x)} = [y_0 : y_1 \ldots : y_N]$ with the y_i's mutually coprime, we define

$$\mathrm{Ht}_{\mathcal{L}}(x) = \prod_{\nu} \max\{|y_0|_\nu, |y_1|_\nu \ldots |y_N|_\nu\}^{1/n},$$

where the product is over all the places of $\mathbb{Q}$. For stacks, there is neither an embedding into projective space, nor can every rational point be extended to an integral point. These problems are solved in forthcoming work, [13], where the authors give a theory of heights on stacks. As a particular case of their results, if a stack $\mathcal{X}$ is equipped with a line bundle $\mathcal{L}$ such that a power of $\mathcal{L}$ can be used to embed $\mathcal{X}$ into a *weighted projective space* instead, then the height of a rational point can be computed in a manner very similar to the scheme case (Proposition 3.5.2). Now, the map

$$(y^2 = x^3 + Ax + B \text{ in minimal form }) \mapsto [A : B]$$

is a map from the modular curve $\mathcal{X}(1)$ to the weighted projective space $\mathbb{P}(4,6)$. Thus with the definition of height from [13], the naive height on $\mathcal{X}_0(N)$ is a constant times the height with respect to the 12th power of the Hodge bundle. Using the fact that naive height comes from geometry allows us to use different sections that globally generate the twelfth power of the Hodge bundle. For a general stack, it is not always easy to find sections that globally generate a line bundle, but in our case, we are in luck, since the ring of modular forms of $\mathcal{X}_0(N)$ is classically well understood (see for e.g. [18]).

Remark 3.1.4. We believe that forthcoming work of Bruin and Najman also proves the asymptotic growth rate for $\mathcal{X}_0(4)$ by using the fact that it is isomorphic over $\mathbb{Z}[1/6]$ to the weighted projective stack $\mathbb{P}(2,2)$.

3.2 Preliminaries

3.2.1 Rationally defined subgroups

In this subsection, we describe a degree two cover of $\mathcal{X}_0(N)$ that we will use to set up our counting problem. To this end, let $N \geq 3$ and let $G = (\mathbb{Z}/N\mathbb{Z})^\times$. Then $\Phi_N : \mathcal{X}_1(N) \to \mathcal{X}_0(N)$ is a branched G-cover of $\mathcal{X}_0(N)$, with branch locus supported possibly at cusps and points with $j = 0, 1728$. Away from the branch locus, G acts freely and transitively on the fibers of Φ_N, by sending $(E,P) \mapsto (E,aP)$. Let H be an index two subgroup of G. We denote by $\mathcal{Y}_{1/2}(N)$ the quotient $\mathcal{Y}_1(N)/H$. One can make sense of the quotient by the action of H at the cusps by using the modular interpretation of cusps. This is explained in Appendix A.1. We denote by $\mathcal{X}_{1/2}(N)$ the quotient $\mathcal{X}_1(N)/H$.

Remark 3.2.1. We begin with some comments about the curve $\mathcal{X}_{1/2}(N)$.

1. The curve $\mathcal{X}_{1/2}(N)$ is not a novel construction. It can be understood classically as the quotient of the upper half plane by an index 2 subgroup of $\Gamma_0(N)$. Further, we do not claim that $\mathcal{X}_1(N)/H$ is a scheme. In fact it is a stack in many cases (see Section 3.4).

2. The notation $\mathcal{X}_{1/2}(N)$ might be misleading, since there is not always a unique index two subgroup of $(\mathbb{Z}/N\mathbb{Z})^\times$. However, in our case we will only consider the H for which G/H

is represented by $\{+H, -H\}$. As an example, $(\mathbb{Z}/8\mathbb{Z})^\times \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$. We will write this set as $\{1, 3, 5, 7\}$. This has three index two subgroups: $H_1 = \{1, 3\}$, $H_2 = \{1, 5\}$ and $H_3 = \{1, 7\}$. The two cosets of H_1 are therefore $H_1 = \{1, 3\}$ and $-H_1 = \{5, 7\}$. The case H_2 is similar. However, the two cosets of H_3 are $H_3 = \{1, 7\}$ and $3H_3 = \{3, 5\}$. We will make it a point *not* to pick H_3. The choice between H_1 or H_2 will not affect our final result.

3. In the context of the remark above, we note that there are some values of N (namely $N = 5, 10, 13, 25$) for which there is no choice of index 2 subgroup such that $G/H = \{\pm H\}$. For these N, while the construction of $\mathcal{X}_{1/2}(N)$ still makes sense, it does not have the nice properties that we want (see Lemma 3.2.2 and Proposition 3.4.1). Another way to rephrase the condition that $G/H = \{\pm H\}$ is in terms of the subgroup $\Gamma_0(N) \subset \mathrm{SL}_2(\mathbb{Z})$. Consider the short exact sequence:

$$1 \to \{\pm 1\} \to \Gamma_0(N) \to \mathbb{P}\Gamma_0(N) \to 1. \tag{3.1}$$

If -1 is a square modulo N, then $\Gamma_0(N)$ contains a primitive fourth root of unity. The converse is also true. Consider the homomorphism $g : \mathrm{SL}_2(\mathbb{Z}) \to \mu_{12}$ given by,

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto e^{\frac{2i\pi}{12}((1-c^2)(bd+3(c-1)d+c+3)+c(a+d+3))}$$

and its restriction to $\Gamma_0(N)$. If $\Gamma_0(N)$ doesn't contain a fourth root of unity, then the image of the restricted map is contained in $\mu_6 \cong \mu_2 \times \mu_3$. One can check that the map g, composed with the map $\mu_6 \to \mu_2$ provides a splitting of the sequence 3.1. Thus for $N \in \{5, 10, 13, 25\}$, this sequence is non-split, while for $N \in \{3, 4, 6, 7, 8, 9, 12, 16, 18\}$, it does split. This splitting allows us to identify $\mathbb{P}\Gamma_0(N)$ with a subgroup of $\Gamma_0(N)$ and hence construct a degree two cover of $\mathcal{X}_0(N)$ without generic inertia.

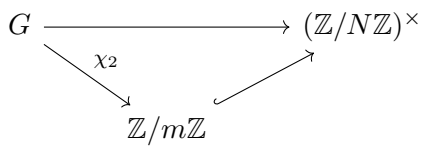
We now explain the significance of the curves $\mathcal{X}_{1/2}(N)$. Most of what follows is well known (e.g., see [31], [16]) but we recall them here for completeness. Let E be an elliptic curve

over $\mathbb{Q}$ with a rational N-isogeny. For notational convenience, we fix a Weierstrass form for $E : y^2 = x^3 + Ax + B$, with $A, B \in \mathbb{Z}$. Fix an isomorphism of the kernel of the rational N-isogeny with $\mathbb{Z}/N\mathbb{Z}$. The Galois action of $G_{\mathbb{Q}}$ on the kernel defines a homomorphism:

$$\chi : G_{\mathbb{Q}} \to (\mathbb{Z}/N\mathbb{Z})^{\times}.$$

For each $N \in \{3, 4, 6, 7, 8, 9, 12, 16, 18\}$, we can write $f : (\mathbb{Z}/N\mathbb{Z})^{\times} \cong (\mathbb{Z}/2\mathbb{Z}) \times (\mathbb{Z}/m\mathbb{Z})$ for some $m \in \mathbb{Z}$. This allows us to factor χ into two characters $\chi_1 : G_{\mathbb{Q}} \to \mathbb{Z}/2\mathbb{Z}$ and $\chi_2 : G_{\mathbb{Q}} \to \mathbb{Z}/m\mathbb{Z}$. That is, we may write $\chi = \chi_1\chi_2$ using the isomorphism f. Now, since χ_1 is a quadratic character, it factors through a quadratic extension $K = \mathbb{Q}(\sqrt{d})$, with d a squarefree integer. Let $E^{\chi_1} : dy^2 = x^3 + Ax + B$ denote the quadratic twist of E over K.

Lemma 3.2.2. *Maintaining the above notation, E^{χ_1} has a rational N-torsion subgroup on which $G_{\mathbb{Q}}$ acts via χ_2. That is, the Galois action on this N-torsion subgroup factors as:*



Proof. Let C denote the kernel of the rational N isogeny of E. Let $\phi : E \to E^{\chi_1}$ denote the isomorphism of elliptic curves defined over $\mathbb{Q}(\sqrt{d})$. For $P \in C$ and $\sigma \in G_{\mathbb{Q}}$, $P^{\sigma} = \chi(\sigma)P$ by assumption. Further, by the definition of a twist,

$$\phi(P)^{\sigma} = \chi_1(\overline{\sigma})\phi(P^{\sigma})$$

where $\overline{\sigma}$ is the image of σ in $\mathbb{Z}/2\mathbb{Z}$. Since χ_1 is quadratic, $\chi_1\chi = \chi_2$. It follows that $G_{\mathbb{Q}}$ acts on $\phi(C)$ via χ_2. □

We have thus proved the following for $N \in \{3, 4, 6, 7, 8, 9, 12, 16, 18\}$.

Proposition 3.2.3. *Fix an appropriate index 2 subgroup $H \subset (\mathbb{Z}/N\mathbb{Z})^{\times}$ and consider the corresponding curve $\mathcal{X}_{1/2}(N)$. Let $(E, C) \in \mathcal{X}_0(N)(\mathbb{Q})$. Then there exists a unique $d \in \mathbb{Z}$*

squarefree, such that the corresponding twist $(E^{\chi_1}, \phi(C))$ *satisfies:*

1. $(\phi(C))^\times$ *has an index two subgroup* H_C *defined over* $\mathbb{Q}$*, and therefore*
2. $(E, H_C), (E, -H_C) \in \mathcal{X}_{1/2}(N)(\mathbb{Q})$.

Proof. This follows from combining the interpretation of $\mathcal{X}_{1/2}(N)$ as a fiberwise quotient of $\mathcal{X}_1(N)$ with Lemma 3.2.2. □

A nice example of Proposition 3.2.3 is in the cases $N = 3, 4, 6$, where $(\mathbb{Z}/N\mathbb{Z})^\times \cong \mathbb{Z}/2\mathbb{Z}$. In these cases $\mathcal{X}_{1/2}(N) = \mathcal{X}_1(N)$. For these values of N, Proposition 3.2.3 says that if E has a rational N-isogeny then there exists a quadratic twist of E that has a rational N torsion point.

3.2.2 Automorphisms and universal families

In this section, we briefly recall the relation between automorphisms and the existence of universal families. For more details, we refer the reader to [19], Chapter 4 and Appendix A.4. Let F be a functor on the category Ell_R of elliptic curves over a ring R. Let $\tilde{F}$ denote the corresponding functor on the category of R-schemes sending an R-scheme S to isomorphism classes of pairs $(E/S, \alpha)$, where E is an elliptic curve over S and $\alpha \in F(E/S)$ is an 'F-level structure'. The functor F (resp. $\tilde{F}$) is *representable* if there exists a universal elliptic curve $\mathcal{E}$ over a scheme $\mathcal{M}$ (resp. a scheme $\mathcal{M}$) such that $F(E/S) = \mathrm{Hom}(E/S, \mathcal{E}/\mathcal{M})$ (resp. $\tilde{F}(S) = \mathrm{Hom}(S, \mathcal{M})$). Note that the representability of F guarantees the existence of $\mathcal{M}$, and therefore implies the representability of $\tilde{F}$. The functor F is said to *rigid* if for any $E/S \in \mathrm{Ell}_R$, and any $\alpha \in F(E/S)$, the pair $(E/S, \alpha)$ has no non-trivial automorphisms. In general, if F is representable, then F is rigid. This is because a non-trivial automorphism of $(E/S, \alpha)$ would induce a non-trivial automorphism of an element of $\mathrm{Hom}(E/S, \mathcal{E}/\mathcal{M})$, which is not allowed for Hom sets in the category of schemes. The following proposition tells us when the converse is true.

Proposition 3.2.4 ([19], 4.7.0)**.** *Suppose that for every elliptic curve* E/S*, the functor* F_E

on the category of schemes over S defined by the map

$$T \mapsto F(E_T/T)$$

is representable by a scheme $F_{E/S}$. Suppose further that F is affine over Ell_R*, that is, the morphism $F_{E/S} \to S$ is affine. Then F is representable if and only F is rigid.*

In this paper, we will be interested in the functors of points corresponding to $\mathcal{X}_0(N)$, $\mathcal{X}_1(N)$ and the intermediate quotient $\mathcal{X}_{1/2}(N)$. To see that in these cases, the two hypotheses of Proposition 3.2.4 are satisfied, we refer the reader to [19]. Thus we may move freely between the existence of universal families and rigidity.

3.2.3 Counting lattice points in a region

In this section, we state a theorem of Davenport on a Lipschitz principle ([9]). Let $\mathcal{R}$ be a closed and bounded region in $\mathbb{R}^n$. Suppose $\mathcal{R}$ satifies the following two conditions:

1. Any line parallel to one of the coordinate axes intersects $\mathcal{R}$ in a set that is a union of at most h intervals.

2. The same is true (with n replaced by m) for any of the m-dimensional regions obtained by projecting $\mathcal{R}$ down to an m-dimensional coordinate axis ($1 \leq m \leq n-1$).

Let $V(\mathcal{R})$ be the volume of the region $\mathcal{R}$ and $N(\mathcal{R})$ the number of lattice points in it. Then, the following theorem holds.

Theorem 3.2.5 (Davenport, [9])**.** *For $\mathcal{R}$ satisfying 1 and 2,*

$$|N(\mathcal{R}) - V(\mathcal{R})| \leq \sum_{m=0}^{n-1} h^{n-m} V_m,$$

where V_m is the sum of the (m-dimensional) volumes of the m-dimensional projections of $\mathcal{R}$ and $V_0 = 1$.

3.3 Main counting results

From Section 3.2.1, we see that in order to count elliptic curves in $\mathcal{X}_0(N)(\mathbb{Q})$ with respect to naive height, we must count elliptic curves for which there exists a quadratic twist that gives a rational point on $\mathcal{X}_{1/2}(N)(\mathbb{Q})$. In this section we state and prove the counting results that will enable us to do so.

Proposition 3.3.1 ([17], Theorem 4.1)**.** *Let $f, g \in \mathbb{Q}[t]$ be coprime polynomials of degrees r and s respectively. Let $\max\{r, s\} > 0$ and let m and n be coprime integers such that*

$$\max\left\{\frac{r}{2}, \frac{s}{3}\right\} = \frac{n}{m}.$$

Assume that either $n = 1$ or $m = 1$. Let $S(X)$ be the set of pairs $(A, B) \in \mathbb{Z}^2$ such that

- $4A^3 + 27B^2 \neq 0$,
- $\gcd(A^3, B^2)$ *not divisible by a 12th power,*
- $|A| < X^{1/3}$ *and* $|B| < X^{1/2}$,
- $\exists u, t \in \mathbb{Q}$ *such that* $A = u^2 f(t)$ *and* $B = u^3 g(t)$.

Define

$$k(x) = \begin{cases} X^{(m+1)/6n} & m+1 > n \\ X^{1/6}\log(X) & m+1 = n \\ X^{1/6} & m+1 < n \end{cases}$$

Then,

$$S(X) \asymp k(x).$$

As we will see in section 4, this theorem is not enough for all the cases that we are interested in. For $N = 3$, the condition: 'either $n = 1$ or $m = 1$' is not satisfied. We will thus prove a generalization of this proposition.

Remark 3.3.2. We note here that we do not prove the most general version of Theorem 3.3.3 possible, since we do not need it. It might be an interesting exercise in analytic number theory to prove such a version, independent of the interpretation of counting points on a moduli space.

Theorem 3.3.3. *Let $f, g \in \mathbb{Q}[t]$ be coprime polynomials of degrees r and s respectively. Let $\max\{r, s\} > 0$ and let m and n be coprime integers such that:*

$$\max\left\{\frac{r}{2}, \frac{s}{3}\right\} = \frac{n}{m}.$$

Suppose that $n, m \neq 1$. Define

$$h = \left\lfloor \frac{n(m-1)}{m} \right\rfloor$$

and $w = \max\{\frac{3h}{s}, \frac{2h}{r}\}$. Suppose further that,

- $m + 1 > n$, *and*
- $\frac{m+1}{n} - (w+1) = -1$
- $\min\{3rm - 6h, 2sm - 6h\} \leq 6$.

Let $S(X)$ be the set of pairs $(A, B) \in \mathbb{Z}^2$ such that

- $4A^3 + 27B^2 \neq 0$,
- $\gcd(A^3, B^2)$ *not divisible by a 12th power,*
- $|A| < X^{1/3}$ *and* $|B| < X^{1/2}$,
- $\exists u, t \in \mathbb{Q}$ *such that* $A = u^2 f(t)$ *and* $B = u^3 g(t)$.

Then,

$$S(X) \asymp X^{(m+1)/6n} \log(X).$$

Remark 3.3.4. Note that the hypotheses on m, n, r and s make it so that there aren't many choices of these variables that satisfy all the hypotheses together. The degree conditions for

$\mathcal{X}_0(3)$ are perhaps the only moduli problem of interest that that satisfy these. However, stating the theorem in this manner instead of using numbers makes the method less opaque and more amenable to generalization.

Proof of Theorem 3.3.3

The proof of this theorem closely follows that in [17]. We provide the key parts of the proof here for the sake of completeness. We prove the upper bound and the lower bound in two separate sections. For the reader's convenience, we outline each proof first.

Notation 3.3.5. For any two real valued functions $h(X)$ and $k(X)$, we say that $h(X) \lesssim k(X)$ if there is a positive constant C such that $h(X) \leq Ck(X)$.

Upper bound. Our goal is to reduce the problem of counting pairs in $S(X)$ to the problem of counting tuples of integers in a bounded region, perhaps with some divisibility conditions. Let $S_1(X)$ be the set of u, t such that $(u^2 f(t), u^3 g(t)) \in S(X)$. Counting $S_1(X)$ gives an upper bound for $S(X)$. We will express u and t as $qc^{-1}db^n$ and ab^{-m} respectively for some integers a, b, c and d and some rational number q. Lemmas 3.3.6, 3.3.7 and 3.3.8 enable us to do this. The next key observation is that there are only finitely many possibilities for q. Thus for the kind of upper bound that we are looking for, we can count 4-tuples of integers (a, b, c, d) in a particular region. Lemma 3.3.9 gives the bounds for such a region. Lemma 3.3.10 outlines what divisibility conditions these integers must satisfy, and also calculates the number of such tuples. □

Lemma 3.3.6 ([17], Lemma 2.2)**.** *For each place p of $\mathbb{Q}$, there is a constant $c_p > 0$ such that for each $t \in \mathbb{Q}$:*

$$\max(|f(t)|_p, |g(t)|_p) \geq c_p.$$

Furthermore, we can take $c_p = 1$ for all sufficiently large p.

Let $S_1(X)$ be the set of u, t such that $(u^2 f(t), u^3 g(t)) \in S(X)$.

Lemma 3.3.7. *For each prime p there is a constant C_p such that for all $(u, t) \in S_1(X)$, we*

have:

$$\mathrm{val}_p(u) = \epsilon' + \begin{cases} \lceil -\frac{n}{m} \mathrm{val}_p(t) \rceil & \mathrm{val}_p(t) < 0 \\ 0 & \mathrm{val}_p(t) \geq 0 \end{cases} \tag{3.2}$$

for some $|\epsilon'| \leq C_p$. *Moreover, we can take* $C_p = 1$ *for all* p *sufficiently large.*

Proof. The proof of this lemma closely follows that of Lemma 2.3 in [17]. Fix a prime p. Since A and B must be integral, we have that:

$$\mathrm{val}_p(A) = 2\,\mathrm{val}_p(u) + \mathrm{val}_p(f(t)) \geq 0 \tag{3.3}$$

$$\mathrm{val}_p(B) = 3\,\mathrm{val}_p(u) + \mathrm{val}_p(g(t)) \geq 0 \tag{3.4}$$

Thus,

$$\mathrm{val}_p(u) \geq \max\left(\lceil -\frac{1}{2} \mathrm{val}_p(f(t)) \rceil, \lceil -\frac{1}{3} \mathrm{val}_p(g(t)) \rceil \right) =: K. \tag{3.5}$$

Note that if $\mathrm{val}_p(u) \geq 2 + K$, then by replacing u by $p^2 u$ we see that $p^{12} \mid \gcd(|A|^3, B^2)$. Thus we must have $K \leq \mathrm{val}_p(u) \leq K + 1$. The rest of the proof goes exactly like in [17]. Suppose $\mathrm{val}_p(t) < 0$. Pick K_1 such that $|\,\mathrm{val}_p(f(t)) - r\,\mathrm{val}_p(t)| < K_1$ and $|\,\mathrm{val}_p(g(t)) - s\,\mathrm{val}_p(t)| < K_1$ for all such t. Note that K_1 can depend on p and is 0 for large enough p. Then,

$$K = \epsilon + \max\left(\lceil -\frac{r}{2} \mathrm{val}_p(t) \rceil, \lceil -\frac{s}{3} \mathrm{val}_p(t) \rceil \right) = \epsilon + \left\lceil \frac{-n}{m} \mathrm{val}_p(t) \right\rceil$$

where $|\epsilon| < K_2$ for some K_2. Thus we have:

$$\epsilon + \left\lceil \frac{-n}{m} \mathrm{val}_p(t) \right\rceil \leq \mathrm{val}_p(u) \leq \epsilon + \left\lceil \frac{-n}{m} \mathrm{val}_p(t) \right\rceil + 1$$

for $\mathrm{val}_p(t) < 0$.

Now consider the case when $\mathrm{val}_p(t) \geq 0$. By Lemma 3.3.6, there exists K_3 such that $\min(\mathrm{val}_p(f(t)), \mathrm{val}_p(g(t)) \leq K_3$. Further, $K_3 = 0$ for $p \gg 0$. Thus $-\mathrm{val}_p(u) \leq K_4$ for some constant K_4. Since $\mathrm{val}_p(t) \geq 0$, there is a K_5 such that $\mathrm{val}_p(f(t)) \geq K_5$ and $\mathrm{val}_p(g(t)) \geq K_5$ for all such t. This gives a lower bound on $-\mathrm{val}_p(u)$, appealing again to (3.5). Thus there is a constant K_7 such that $|\mathrm{val}_p(u)| \leq K_7$. We remark here to avoid confusion that all the K_i's are constant with respect to t and u, but do depend on p, f and g.

This gives us first part of the lemma. For the second part of the lemma, we need only take, as in [17], $p \gg 0$ such that: (1) the coefficients of f and g are p-integral, (2) the leading coefficients of f and g are p-units and (3) the constant c_p in lemma 3.3.6 can be taken to be 1. Since $K \leq \mathrm{val}_p(u) \leq K + 1$, we can only get $C_p = 1$ for $p \gg 0$.

□

The next step is to prove an analogue of Lemma 2.4 in [17]. This will enable us to reduce our problem to that of counting lattice points in a region. We start with some notation. Recall that $w = \max\{3h/s, 2h/r\}$. For a given pair of positive integers (a, b), we say a prime p satisfies $(*)$ if:

$$p|b \implies p^w|a$$

Lemma 3.3.8. *Suppose $(u, t) \in S_1(X)$. There is a* finite *set $Q \subset \mathbb{Q}^\times$ (independent of u and t) such that: we can write $t = ab^{-m}$ and $u = qc^{-1}db^n$, where:*

1. *$a, b \in \mathbb{Z}$, with $b > 0$,*
2. *$\gcd(a, b^m)$ is m-th power free,*
3. *d is a squarefree integer,*
4. *$q \in Q$, and*
5. *$c \in \mathbb{Z}$ such that $\mathrm{val}_p(c) \leq h$ for all p and $\mathrm{val}_p(c) > 0$ if and only if p satisfies $(*)$.*

Proof. Given $t \in \mathbb{Q}$, one can always write $t = ab^{-m}$ satisfying (1) and (2). Pick any such

representation. We now analyze ub^{-n} and show that $\mathrm{val}_p(ub^{-n})$ must satisfy the required constraints. For convenience we will fix N_0 to be an integer such that $C_p = 1$ for $p \geq N_0$. Such an N_0 exists by Lemma 3.3.7.

We divide the set of all primes into two groups: $p|b$ and $p \nmid b$. If $p \nmid b$, then $\mathrm{val}_p(t) \geq 0$ and by Lemma 3.3.7, we have $|\mathrm{val}_p(ub^{-n})| \leq C_p$. If $p|b$, then we write $-\mathrm{val}_p(t) = m\,\mathrm{val}_p(b) - k$, where $0 \leq k < m$. Therefore, by Lemma 3.3.7 again, we have:

$$\mathrm{val}_p(u) = \epsilon' + n\,\mathrm{val}_p(b) + \lceil \frac{-n}{m} k \rceil.$$

Therefore for any p, we have $-C_p - h \leq \mathrm{val}_p(ub^{-n}) \leq C_p$.

If $p \leq N_0$, we have no control over C_p, but we know that there are finitely many possibilities for the N_0-smooth part of ub^{-n}, since $|\mathrm{val}_p(ub^{-n})| \leq C_p + h$ (here, N_0-smooth means the part of the numerator or denominator that is divisible only by primes less than or equal to N_0). For $p \geq N_0$, we have:

$$|\mathrm{val}_p\,(ub^{-n})| \leq 1 \text{ if } p \nmid b$$

and

$$-1 - h \leq \mathrm{val}_p\,(ub^{-n}) \leq 1 \text{ if } p|b.$$

In the case that $p \nmid b$ and $p \geq N_0$, we see that $\mathrm{val}_p(t) \geq 0$. Further, in the proof of the previous lemma, N_0 is picked so that for $p \geq N_0$, the $\mathrm{val}_p(f(t)) \geq 0$ and $\mathrm{val}_p(g(t)) \geq 0$, with at least one of them being an equality. In particular, this implies that $\mathrm{val}_p(ub^{-n}) \geq 0$ for such p. Similarly, for $p|b$ ($\mathrm{val}_p(t) < 0$) and $p \geq N_0$, we can take $\epsilon' = 0$. Thus, $\mathrm{val}_p(ub^{-n}) = \lceil \frac{-n}{m} k \rceil$ or $\lceil \frac{-n}{m} k \rceil + 1$. Therefore $\mathrm{val}_p(ub^{-n}) \geq -h$.

We factor the $p \geq N_0$ part of ub^{-n} as $c^{-1}d$, where $\mathrm{val}_p(d) \neq 0$ iff either $p \nmid b$ or $\mathrm{val}_p(ub^{-n}) = 1$ if $p|b$. Further, in these cases, we set $\mathrm{val}_p(d) = \mathrm{val}_p(ub^{-n})$. The previous paragraph shows that

d is a squarefree integer and that $\mathrm{val}_p(c) \leq h$.

We now explain the condition $(*)$. This comes from the fact that A and B are required to be integers. For any p:

$$\begin{aligned}\mathrm{val}_p(u^2 f(t)) &= 2\,\mathrm{val}_p(qc^{-1}b^n) + \mathrm{val}_p(f(t)) \\ &= 2\,\mathrm{val}_p(q) - 2\,\mathrm{val}_p(c) + 2m(n/m - r/2)\,\mathrm{val}_p(b) + r\,\mathrm{val}_p(a) + K_1 \\ &\geq 2\,\mathrm{val}_p(q) + K_1 + r\,\mathrm{val}_p(a) - 2\,\mathrm{val}_p(c)\end{aligned}$$

where K_1 is a positive constant that can be taken to be 0 for $p \gg 0$. Similarly, for B, we get that: $\mathrm{val}_p(u^3 g(t)) = 3\,\mathrm{val}_p(q) + K_1 + s\,\mathrm{val}_p(a) - 3\,\mathrm{val}_p(c)$. Since q is N_0-smooth, for p large enough, the condition of integrality of A and B translates directly to condition $(*)$. Further, since we are only interested in an upper bound for the asymptotic growth, not imposing conditions on say, $2\,\mathrm{val}_p(q) + K_1$ for small p causes us no harm. □

Now consider $(u,t) \in S_1(X)$ and write them as in Lemma 3.3.8. The fact that $\max\{|A|^3, |B|^2\} < X$ implies bounds for a, b, c and d, which we now find.

Lemma 3.3.9. *Let $(u,t) \in S_1(X)$. Represent $u = qc^{-1}db^n$ and $t = ab^{-m}$ as in Lemma 3.3.8. Then,*

$$|a| \lesssim X^{m/6n} c^{m/n} d^{-m/n} \qquad \text{and} \qquad |b| \lesssim X^{1/6n} c^{1/n} d^{-1/n}.$$

Proof. If $A = u^2 f(t)$ and $B = u^3 g(t)$, the bound $\max(|A|^3, |B|^2) < X$ translates to:

$$|u| \max(|f(t)|^{1/2}, |g(t)|^{1/3}) < X^{1/6}.$$

Let K be the positive constant such that $\max(|f(t)|^{1/2}, |g(t)|^{1/3}) > K$ for all t. Thus: $|u| \leq K^{-1} X^{1/6}$. Let $M_2 = K^{-1}(\max_{q \in Q} |q|^{-1})$. Thus, we have that

$$|c^{-1} d b^n| < M_2 X^{1/6},$$

i.e. $|b| < M_2 X^{1/6n} c^{1/n} d^{-1/n}$.

We now turn to bounding $a(= tb^m)$. Suppose $t < 1$. Then, by the above bound for b, we have $|a| < M_2 X^{m/6n} c^{m/n} d^{-m/n}$. If $t \geq 1$, then we can find a constant $M > 0$ such that $M^2|t|^r \leq |f(t)|$ and $M^3|t|^s \leq |g(t)|$. Thus we have that,

$$X^{1/6} > |u| \max(|f(t)|^{1/2}, |g(t)|^{1/3}) > M|u| \max(|t|^{r/2}, |t|^{s/3}) = M|ut^{n/m}|.$$

Now, $|ut^{n/m}| = |qc^{-1}da^{n/m}|$ and so, $|c^{-1}da^{n/m}| < M^{-1}(\max_{q \in Q} |q|^{-1}) X^{1/6}$. Thus, we see that

$$|a| \lesssim X^{m/6n} c^{m/n} d^{-m/n}.$$

□

Lemma 3.3.10. *Under the hypotheses of Theorem 3.3.3,* $|S_1(X)| \lesssim X^{m+1/6n} \log(X)$.

Proof. Fix a $c > 1$. Let $S_1(X; c)$ denote the set of all $(a, b, d) \in \mathbb{Z}^3$ such that

1. $|ad^{m/n}| < X^{m/6n} c^{m/n}$,
2. $|bd^{1/n}| < X^{1/6n} c^{1/n}$,
3. $p|c \iff p|b, p^w|a$, and $\mathrm{val}_p(c) \leq h$ for all p.

Let $T(X; d, c) = \{(a, b) \mid (a, b, d) \in S_1(X; c)\}$. By standard analytic number theory and Theorem 3.2.5, it follows that

$$|T(X; d, c)| = \frac{1}{c^{w+1}} X^{(m+1)/6n} d^{-(m+1)/n} c^{(m+1)/n} + O\left(\frac{1}{c^{w+1}} X^{m/6n} d^{-m/n} c^{m/n}\right).$$

Thus we have,

$$|S_1(X; c)| = \sum_{d < X^{1/6}} |T(X; d, c)|$$

$$= \sum_{d<X^{1/6}} c^{(m+1)/n-(w+1)} X^{(m+1)/6n} d^{-(m+1)/n} + O\left(\sum_{d<X^{1/6}} \frac{1}{c^{w+1}} X^{m/6n} d^{-m/n} c^{m/n} \right).$$

We will only consider the case $m+1 > n$. Further, since $m \neq n$, the error term above just becomes

$$O\left(\frac{1}{c^{w+1}} X^{m/6n} c^{m/n} \right).$$

Therefore we have

$$|S_1(X;c)| \lesssim c^{(m+1)/n-(w+1)} X^{(m+1)/6n}.$$

Summing over $(h+1)$-th power-free c, with $c < X^{\alpha}$ (for any α), since $\frac{m+1}{n} - (w+1) = -1$:

$$|S_1(X)| \lesssim X^{(m+1)/6n} \log(X).$$

□

Lower Bound. The outline of the proof of the lower bound is as follows: we know that if $(u,t) \in S_1(X)$, then u and t have expressions as in Lemma 3.3.8. Instead of counting all of these, we only count ones of the form $u = c^{-1}b^n$ and $t = ab^{-m}$, where a, b and c are within appropriate bounds. Let $S_2(X)$ be the set of such triples (a,b,c). There is a map $S_2(X) \to S(X)$, and the bulk of the proof is in showing that this map has bounded fibers. We first form another intermediary set, which we call $S_3(X)$. We then describe maps $S_2(X) \to S_3(X) \to S(X)$, and bound the fibers of these maps. This will enable us to find a lower bound for $S(X)$ by finding one for $S_2(X)$ instead. □

Since we only need a lower bound, observe that by changing u to Mu for large enough M, we can assume that $f(t), g(t) \in \mathbb{Z}[t]$. For a triple $(a,b,c) \in \mathbb{Z}^3$, set $u = c^{-1}b^n$ and $t = ab^{-m}$. Let $A = u^2 f(t)$ and $B = u^3 g(t)$. Fix some constant $\kappa > 0$. Define $S_2(X)$ to be: the set of triples $(a,b,c) \in \mathbb{Z}^3$ such that:

- $c = \prod_{p|b, p^w|a} p^h$,
- $0 < b < \kappa X^{1/6n} c^{1/n}$, $|a| < \kappa X^{m/6n} c^{m/n}$, $\gcd(a, b^m)$ is m-th power free
- $4A^3 + 27B^2 \neq 0$ (where A and B are as defined above).

Note that if $(a, b) \in S_2(X)$, then for a suitable value of κ, we get $(A, B) \in S(X)$, since $|A| = |u^2 f(t)| \lesssim |c^{-2} a^r b^{2n-mr}| \lesssim X^{1/3}$, and similarly for B.

Notation: Define $S_3(X) \subset \mathbb{Z}^2$ to be the set of $(A, B) \in \mathbb{Z}^2$ coming from $S_2(X)$. We then have a map from $S_3(X) \to S(X)$ sending $(A, B) \mapsto (A/d^4, B/d^6)$ where $d^{12} \mid\mid \gcd(A^3, B^2)$. Stratify $S_2(X)$ by sets $S_2(X; c)$ of pairs (a, b) such that $\prod_{p|b, p^w|a} p^h = c$. Define $S_3(X; c)$ as the pairs (A, B) coming from $(a, b) \in S_2(X; c)$.

The following lemma will help us bound the fibers of the map $S_3(X) \to S(X)$.

Lemma 3.3.11. *There exists a non-zero integer D (depending only on f and g) with the following property: if $(a, b, c) \in S_2(X)$, then $\gcd(A^3, B^2)$ can be factored as $(M_D)\beta$ such that M_D divides D and $p|\beta \implies p|b$.*

Proof. We follow the same method of proof as Harron and Snowden. Let $(a, b) \in S_2(X; c)$ and let p be a prime. Let M_1 be a constant such that $|3 \operatorname{val}_p(f(t)) - 3r \operatorname{val}_p(t)| < M_1$ and $|2 \operatorname{val}_p(g(t)) - 2s \operatorname{val}_p(t)| < M_1$ for all $t \in \mathbb{Q}$ with $\operatorname{val}_p(t) < 0$. Let M_2 be the constant for which $\min\{3 \operatorname{val}_p(f(t)), 2 \operatorname{val}_p(g(t))\} \leq M_2$ for all $t \in \mathbb{Q}$ with $\operatorname{val}_p(t) \geq 0$. Note that $\max\{M_1, M_2\}$ is 0 for $p \gg 0$ (specifically, $p \geq N_0$, as defined in Lemma 3.3.9).

Now, consider the case where $\operatorname{val}_p(t) < 0$. In particular, $p|b$. Let $\operatorname{val}_p(b) = k$ and let $\operatorname{val}_p(a) = l (< m)$. We then have:

$$\operatorname{val}_p(A^3) = \begin{cases} 6mk\left(\frac{n}{m} - \frac{r}{2}\right) + 3rl - 6h + \epsilon & p|c \\ 6mk\left(\frac{n}{m} - \frac{r}{2}\right) + 3rl + \epsilon & p \nmid c \end{cases}$$

$$\mathrm{val}_p(B^2) = \begin{cases} 6mk\left(\frac{n}{m} - \frac{s}{3}\right) + 2sl - 6h + \delta & p|c \\ 6mk\left(\frac{n}{m} - \frac{s}{3}\right) + 2sl + \delta & p \nmid c \end{cases}$$

where $|\epsilon| < M_1$ and $|\delta| < M_1$. Let $M_0 = \min\{3rm, 2sm\}$. Let

$$e_p = \begin{cases} \max\{M_1 + M_0, M_2\} & p \leq N_0 \\ M_0 & p \geq N_0 \end{cases}$$

and take $D = \prod_{p \leq N_0} p^{e_p}$. This proves the lemma. □

Remark 3.3.12. We find that D is N_0-smooth and β consists of $p|b$ for $p \geq N_0$. It is crucial that D, M_1, M_2 and M_0 do not depend on (a, b, c) in any way. They only depend on f and g.

We now use this lemma to bound the fibers of $S_3(X) \to S(X)$ in our case of interest, namely when

$$\min\{3rm - 6h, 2sm - 6h\} \leq 6.$$

We will call this assumption $(**)$.

Lemma 3.3.13. *There exists a constant N such that the size of the fibers of $S_3(X) \to S(X)$ is bounded by N.*

Proof. The fiber over a point $(A', B') \in S(X)$ is in bijection with the set $\{d \in \mathbb{Z} \mid (d^4A', d^6B') \in S_3(X)\}$. Thus for any $(A, B) \in S_3(X)$, the size of the fiber above the pair is bounded above by the number of 12th powers dividing $\gcd(|A|^3, B^2)$. We show that this is exactly the number of 12th powers dividing D from Lemma 3.3.11, i.e. no 12th powers divide β.

Consider a prime $p \geq N_0$. We claim that p^{12} cannot divide $\gcd(|A|^3, |B|^2)$. If $p|b$ and $p|c$, then this follows from assumption $(**)$. If $p \nmid b$, then since $K_2 = 0$, p doesn't divide $\gcd(|A|^3, |B|^2)$. If $p|b$ and $p \nmid c$, then by definition of c, we must have that $p^w \nmid a$. Since $h = 1$, this forces $l \leq 1$ in Lemma 3.3.11. Since assumption $(**)$ implies $\min\{3r, 2s\} < 12$, we are done. □

The rest of the proof follows by the exact argument as that of Harron and Snowden, which we recall below.

Lemma 3.3.14. *There exists a constant M such that every fiber of the map $S_2(X) \to S_3(X)$ has size bounded by M.*

Proof. Fix any $(A, B) \in S_3(X)$. An element in fiber of the map $S_2(X) \to S_3(X)$ above (A, B) is of the form $(a, b, c) \in \mathbb{Z}^3$ with $A = (c^{-1}b^n)^2 f(ab^{-m})$ and $B = (c^{-1}b^n)^3 g(ab^{-m})$, and $c = \prod_{p|b, p^w|a} p^h$. Set $x = cb^{-n}$ and $y = ab^{-m}$. Then an element (a, b, c) in the fiber satisfies the equations:

$$Ax^2 = f(y) \qquad Bx^3 = g(y).$$

These can be thought of as defining curves in $\mathbb{P}^2$ that intersect transversally, since f and g are coprime. Thus by Bezout's theorem, the maximum number of solutions is bounded above by: $M = \max(2, r)\max(3, s)$. □

It only remains to bound the size of $S_2(X)$. Now, $S_2(X) = \coprod_c S_2(X; c)$ and the size of $S_2(X; c)$ is precisely

$$\frac{1}{c^{w+1}} c^{(m+1)/n} X^{(m+1)/6n} + O\left(\frac{1}{c^{w+1}} c^{m/n} X^{m/6n}\right),$$

where the error term comes from Theorem 3.2.5. Summing over c gives us that

$$X^{(m+1)/6n} \log(X) \lesssim S_2(X).$$

3.4 Proof of Theorem 1.2.5 for $N \neq 2, 5$

The proof of Theorem 1.2.5 for the cases $N \neq 2, 5$ involves applying the appropriate theorems from §3.3. We start off with geometric descriptions of $\mathcal{X}_{1/2}(N)$ as constructed in Section 3.2.1.When we say that a curve has n *stacky points*, we are talking about n stacky geometric points ([33, Tag 04XE]). For modular curves, this can be thought of as referring to n distinct

values of the corresponding hauptmoduln (Definition 3.1.3) or cusps. We use the term 'stacky curve' as defined in [39] to mean curves that have a trivial generic inertia stack.

Proposition 3.4.1. *For any $N \in \mathbb{Z}_{>0}$, consider the curve $\mathcal{X}_{1/2}(N)$ constructed in §2. Then:*

1. *If $N = 3$, then $\mathcal{X}_{1/2}(N) = \mathcal{X}_1(3)$, which is a stacky curve with one stacky point corresponding to the elliptic curves with j-invariant 0.*
2. *If $N = 4$, then $\mathcal{X}_{1/2}(N) = \mathcal{X}_1(4)$, which is a stacky curve whose only stacky point is at the irregular cusp.*
3. *If $N = 7$, then $\mathcal{X}_{1/2}(N)$ is a stacky curve with two stacky points whose hauptmoduln are defined over $K = \mathbb{Q}(\sqrt{-3})$ and are conjugate over $\mathbb{Q}$.*
4. *If $N = 6, 8, 9, 12, 16, 18$, then $\mathcal{X}_{1/2}(N)$ is a scheme.*
5. *If $N = 5, 10, 13, 25$, then $\mathcal{X}_{1/2}(N)$ has generic inertia stack $B\mu_2$*

Proof. These claims follow from the construction of $\mathcal{X}_{1/2}(N)$, by analysing the automorphisms of its points and applying Proposition 3.2.4. For $N = 3, 4$ and 6, this is classical, as in each of these cases $\mathcal{X}_{1/2}(N) = \mathcal{X}_1(N)$. We demonstrate the cases $N = 5$, 7 and 8, and leave the rest to the reader. For readability, we do not separately talk about the cusps, but the non-stackiness in the cases of interest follows from the modular interpretation given in A.1.

Consider the map $\mathcal{X}_{1/2}(N) \to \mathcal{X}_0(N)$. Since any point in $\mathcal{X}_{1/2}(N)$ lies in some geometric fiber of this map, it is enough to analyse automorphisms of points in each fiber. For any point $(E, C) \in \mathcal{X}_0(N)$, choose an isomorphism $C \cong \mathbb{Z}/N\mathbb{Z}$ and thus $\mathrm{Aut}(C) \cong (\mathbb{Z}/N\mathbb{Z})^\times$. Let P be a generator for C.

For $N = 7$, the fiber above a point (E, C) contains the points

$$(E, \{P, 2P, 4P\}) \text{ and } (E, \{-P, -2P, -4P\}).$$

If E has non-zero j-invariant, then the only extra automorphism of the pair (E, C) is $[-1]$ and thus the points in the fiber do not have any extra automorphisms. Recall that $\mathcal{X}_0(7)$ has exactly two elliptic points, both with j-invariant 0 (see for instance, [11]). Each of these points has an automorphism group that is cyclic of order 6. Let μ be an automorphism of the pair (E, C) of order 6. Then $\mu = [-1] \circ \mu'$ for some μ' of order 3. By the definition of $\mathcal{X}_{1/2}(N)$, this automorphism must fix the points $(E, \{P, 2P, 4P\})$ and $(E, \{-P, -2P, -4P\})$. In order to find the hauptmoduln corresponding to these points, we use the tables in [25]. Since $\mathcal{X}_{1/2}(N)$ and $\mathcal{X}_0(N)$ have isomorphic coarse spaces, they have the same hauptmoduln (in fact, in this case they have the same even weight modular forms). Let t be the hauptmodul in [25] for $\mathcal{X}_0(7)$. Then the family over $\mathcal{X}_{1/2}(7)$ is given by $y^2 = x^3 + A(t)x + B(t)$ where

$$A(t) = \frac{1}{3}(t^2 + 245t + 2401)(t^2 + 13t + 49)$$
$$B(t) = \frac{1}{2}(t^4 - 10 \cdot 7^2 t^3 - 9 \cdot 7^4 t^2 - 2 \cdot 7^6 t - 7^7)(t^2 + 13t + 49),$$

and the j-invariant is given by

$$j = \frac{(t^2 + 245t + 2401)^2(t^2 + 13t + 49)}{t^7}.$$

This family is not universal over the points where $t^2 + 13t + 49 = 0$, since $\mathcal{X}_0(N)$ parametrizes only semistable curves. The roots of this equation are defined over $\mathbb{Q}(\sqrt{-3})$.

Remark 3.4.2. Another way to find the stacky points is to note that the universal family over $\mathcal{Y}_1(7)$ is

$$y^2 + (1 + v - v^2)xy + (v^2 - v^3)y = x^3 + (v^2 - v^3)x^2,$$

with torsion point $(0, 0)$ [22, Table 3]. The j-invariant of the universal family is

$$\frac{(v^6 - 11v^5 + 30v^4 - 15v^3 - 10v^2 + 5v + 1)^3(v^2 - v + 1)^3}{(v - 1)^7 v^7 (v^3 - 8v^2 + 5v + 1)}.$$

This gives exactly eight values of v producing a curve of j-invariant 0. Let α be a root of

$v^2 - v + 1 = 0$. Then the fiber over the curve

$$y^2 + 2xy + \alpha y = x^3 + \alpha x^2$$

contains the points: $\{P = (0,0), (-\alpha, \alpha), (-1, -\alpha + 1), (-1, 1), (-\alpha, 0), (0, -\alpha)\}$. The automorphism of this curve given by $(x, y) \mapsto (-\alpha x - \alpha, y + (\alpha + 1)x + \alpha)$ fixes the subgroup $\{P, 2P, 4P\}$.

If $N = 8$, then recall from Section 3.2.1 that the choice of index 2 subgroup of $(\mathbb{Z}/N\mathbb{Z})^\times$ is not unique, and we choose one that works for us. That is, write $(\mathbb{Z}/8\mathbb{Z})^\times = \{P, 3P, -3P, P\}$ and suppose we choose the subgroup $\{P, 3P\}$, so that fiber above (E, C) consists of the points $(E, \{P, 3P\})$ and $(E, \{-P, -3P\})$. Neither pair has extra automorphisms. If $N = 5$. Then $C^\times \cong \{P, 2P, -2P, -1P\}$, which has a unique index 2 subgroup: $\{P, -P\}$. Thus the fiber above (E, C) has two points: $(E, \{P, -P\})$ and $(E, \{2P, -2P\})$. Each of these points still has the automorphism $[-1]$. This proves the theorem for $N = 5$. □

What this proposition tells us is that if $N \in \{3, 4, 6, 7, 8, 9, 12, 16, 18\}$, then there is an open substack $\mathcal{U}$ of $\mathcal{X}_{1/2}(N)$ that is isomorphic to a scheme. Therefore $\mathcal{U}(\mathbb{Q})$ can be parametrized via the universal family over $\mathcal{U}$. For $N \in \{4, 6, 8, 9, 12, 16, 18\}$, the non-stacky locus contains $\mathcal{Y}_{1/2}(N)$, and thus there exist f_N and $g_N \in \mathbb{Q}[t]$ coprime such that every elliptic curve arising from a rational point on $\mathcal{Y}_{1/2}(N)$ is isomorphic to one of the form:

$$\mathcal{E}_{N,t} : y^2 = x^3 + f_N(t)x + g_N(t).$$

Thus, by Proposition 3.2.3, we have the following:

$$\begin{aligned}\mathcal{N}(N, X) &= \#\{E \mid \mathrm{ht}(E) < X, \text{ and } \exists d \in \mathbb{Z}, u, t \in \mathbb{Q}, \text{s.t. } E_d : y^2 = x^3 + u^4 f_N(t)x + u^6 g_N(t)\} \\ &= \#\{E \mid \mathrm{ht}(E) < X, \text{ and } \exists u, t \in \mathbb{Q}, \text{s.t. } E : y^2 = x^3 + u^2 f_N(t)x + u^3 g_N(t)\}.\end{aligned}$$

To find the asymptotic growth for $\mathcal{N}(N, X)$ in these cases, we use Proposition 3.3.1 to find the

value of $h_N(X)$, given in Table 3 below.

For $N = 3$, the situation is slightly different. The curve $\mathcal{X}_{1/2}(3) = \mathcal{X}_1(3)$ has one stacky point lying above the elliptic curve with j-invariant 0. Let $\Phi_3 : \mathcal{X}_{1/2}(3) \to \mathcal{X}(1)$ be the usual forgetful map. Set $Y = \mathcal{Y}_{1/2}(3) \setminus \phi_3^{-1}(\{j = 0\})$. Then, for a suitable embedding of $Y \hookrightarrow \mathbb{A}^1$, there is a universal family $\mathcal{E}_{3,t}$ over Y (e.g. see [17]) given by

$$\mathcal{E}_{3,t} : y^2 = x^3 + \left(2t - \frac{1}{3}\right) x + \left(t^2 - \frac{2}{3}t + \frac{2}{27}\right).$$

Every elliptic curve with non zero j-invariant and a rational 3-torsion point is isomorphic to one of the above form for some $t \in \mathbb{Q}$. However, this family does not extend to a universal family over $t = 1/6$. Indeed $\mathcal{E}_{3,1/6}$ is given by $y^2 = x^3 - \frac{1}{108}$ and its torsion subgroup of order 3 is generated by the rational point: $(1/3, 1/6)$. On the other hand, all curves $E^D : y^2 = x^3 + D^2$, $D \in \mathbb{Q}$ contain the rational 3 torsion point $(0, D)$ and have j-invariant 0, but none of them is isomorphic to $\mathcal{E}_{3,1/6}$ over $\mathbb{Q}$. For this reason, we separate our counting function into two pieces:

$$\mathcal{N}(3, X) = \mathcal{N}(3, X)_{j=0} + \mathcal{N}(3, X)_{j\neq 0}.$$

By Theorem 3.3.3, we have the following proposition.

Proposition 3.4.3. *Maintaining the notation as above,*

$$\mathcal{N}(3, X)_{j\neq 0} \asymp X^{1/3} \log(X).$$

In order to find the asymptotics for $\mathcal{N}(3, X)_{j=0}$, we observe the following: by lemma 3.4 in [17], we know that any elliptic curve that has j-invariant 0, a rational 3 torsion point, but is not of the form $\mathcal{E}_{3,t}$ for any $t \in \mathbb{Q}$, admits an equation of the form $y^2 = x^3 + D^2$, $D \in \mathbb{Z}$. Thus the curves missing from our count are those that are quadratic twists of these exceptional curves.

That is, they are elliptic curves of the form

$$y^2 = x^3 + u^3t^2$$

for some $u, t \in \mathbb{Q}$ with u^3t^2 integral and minimal. This is the same as counting elliptic curves $y^2 = x^3 + b$, with $b^2 < X$ and b 6th power free. This number is just a constant times $X^{1/2}$.

Remark 3.4.4. Note that our result agrees with that in [28]. In fact the argument for $\mathcal{N}(3, X)_{j=0}$ is exactly the same as in their paper, albeit stated slightly differently.

To complete the proof of the main theorem, for each N we need only calculate r, s, m and n in the notation of Proposition 3.3.1 and Theorem 3.3.3. In Table 3, we give the components required to compute r and s in each of the cases of interest. We do not give the explicit polynomials f_N and g_N here, since we do not need them, but these polynomials are given in Appendix A.2.

N	r	s	m	n	Reference	$h_N(X)$
3	1	2	3	2	3.3.3	$X^{1/2}$
4	2	3	1	1	3.3.1	$X^{1/3}$
6	4	6	1	2	3.3.1	$X^{1/6}\log(X)$
8	4	6	1	2	3.3.1	$X^{1/6}\log(X)$
9	4	6	1	2	3.3.1	$X^{1/6}\log(X)$
12	8	12	1	4	3.3.1	$X^{1/6}$
16	8	12	1	4	3.3.1	$X^{1/6}$
18	12	18	1	6	3.3.1	$X^{1/6}$

Table 3: Values of invariants

Remark 3.4.5 (Distinction between $N = 3$ and $N = 7$). One might wonder why one can find a model for an open substack $\mathcal{U}$ of $\mathcal{Y}_{1/2}(3)$ with f_3 and g_3 coprime, but not for $\mathcal{Y}_{1/2}(7)$. A priori, a model of the form $y^2 = x^3 + f(t)x + g(t)$ found for $\mathcal{U}$ might not have f and g coprime. However, since $\mathcal{Y}_1(3)$ only has one geometric stacky point, that point can be moved to $\infty \in \mathbb{P}^1$ via a transformation. On the other hand, $\mathcal{Y}_{1/2}(7)$ has two stacky points, neither of which is rational. Thus we cannot find f_7 and g_7 coprime, and therefore cannot apply our method.

3.5 Counting points of bounded height on stacks

In this section, we prove Theorem 1.2.5 for $N = 2, 3, 4, 5, 6, 8, 9$ by using results from forthcoming work of Ellenberg, Satriano and Zureick-Brown in [13]. As we have seen, one can define *some height* on $\mathcal{X}_0(N)$, namely the naive height. The question is does this height come from geometry? We know that this is true for modular curves that are schemes – the naive height is the height with respect to the twelfth power of the Hodge bundle. It follows from the work in [13] that the same is true for moduli *stacks* of elliptic curves, and we use their machinery to count the number of points of bounded height. Before we proceed, we must set up some notation.

Notation 3.5.1. Recall that we use $\mathrm{ht}(E)$ for the naive height of a point E on any modular curve. Let $\mathcal{X}$ be a stack and $\mathcal{V}$ a vector bundle on it. We will let $\mathrm{ht}_{\mathcal{V}}$ denote the logarithmic height with respect to $\mathcal{V}$ as defined in [13] and $\mathrm{Ht}_{\mathcal{V}}$ the multiplicative height corresponding to it. That is to say, $\mathrm{Ht}_{\mathcal{V}} = \exp(\mathrm{ht}_{\mathcal{V}})$.

We will not define $\mathrm{ht}_{\mathcal{V}}$ here, but we will use the fact that if $\mathcal{V} = \lambda^{\otimes 12}$ on $\mathcal{X}_0(N)$, then for an elliptic curve E corresponding to a rational point $x : \operatorname{Spec} \mathbb{Q} \to \mathcal{X}_0(N)$, $\log \mathrm{ht}(E) = \mathrm{ht}_{\mathcal{V}}(x) + O(1)$ (see Example 3.5.3 below). Thus our counting function satisfies

$$\mathcal{N}(N, X) \asymp \#\{x \in \mathcal{X}_0(N)(\mathbb{Q}) \mid \mathrm{Ht}_{\lambda}^{12}(x) < X\}. \tag{3.6}$$

3.5.1 Computing heights on stacks

Throughout this subsection, $\mathcal{X}$ will be a proper Artin stack over $\operatorname{Spec} \mathbb{Z}$ with finite diagonal. A $\mathbb{Q}$-rational point x of $\mathcal{X}$ is a map $x : \operatorname{Spec} \mathbb{Q} \to \mathcal{X}$. Let $\mathcal{V}$ be a vector bundle on $\mathcal{X}$. Consider for a moment the case where $\mathcal{X} = X$, a proper scheme, and $\mathcal{V}$ is an ample line bundle on it. When computing the height of a point on X, we use a power of $\mathcal{V}$ to embed $X \hookrightarrow \mathbb{P}^n$ for some n, and then use the naive height of the image of the point on $\mathbb{P}^n$. This makes computations easier. For a stack, the analogue would be mapping it into weighted projective space. In [13], the authors show that this works. We recall the specific result below.

Consider the special case where $\mathcal{V}$ is a metrized line bundle $\mathcal{L}$ (see [13] for precise definition). Suppose $s_1, s_2, \ldots, s_k$ are sections of $\mathcal{L}$. Then, $\mathcal{L}$ is said to be *generically globally generated by* $s_1, \ldots, s_k$ if the cokernel of the corresponding morphism

$$\mathcal{O}_{\mathcal{X}}^{\oplus k} \to \mathcal{L}$$

vanishes over the generic point of $\operatorname{Spec}\mathbb{Z}$. In particular, this implies that the cokernel is supported at finitely many places.

Proposition 3.5.2 ([13], Proposition 2.27)**.** *Let $\mathcal{X}$ be a stack over* $\operatorname{Spec}\mathbb{Z}$*, let $\mathcal{L}$ be a line bundle on $\mathcal{X}$ such that $\mathcal{L}^{\otimes n}$ is generically globally generated by sections $s_1, s_2 \cdots s_k$. Let $x : \operatorname{Spec}\mathbb{Q} \to \mathcal{X}$ and for each i, let $x_i = x^*(s_i)$ (after picking an identification of $x^*\mathcal{L}$ with $\mathbb{Q}$). Scale $x_1, \ldots, x_k$ so that each $x_i \in \mathbb{Z}$ and for every prime p, there is some x_i such that $v_p(x_i) < n$. Then*

$$\mathrm{ht}_{\mathcal{L}}(x) = \frac{1}{n} \log \max_i \{|x_1|, |x_2| \ldots |x_k|\} + O_{\mathcal{X}(\mathbb{Q})}(1)$$

where $|\cdot|$ is the usual archimedean absolute value.

Note here that we have only stated the version of the proposition that we require, i.e. for $\operatorname{Spec}\mathbb{Q}$ and $\operatorname{Spec}\mathbb{Z}$. A more general version of this proposition holds for other global fields. We will say that the tuple $(x_1, \ldots x_k) \in \mathbb{Z}^k$ is *minimal* if it satisfies the last condition in the theorem: for each prime p, there is some $i \in \{1 \ldots k\}$ such that $p^n \nmid x_i$.

Example 3.5.3. Let $\mathcal{L} = \lambda$, the Hodge bundle on $\mathcal{X}(1)$. Then the global sections of $\lambda^{\otimes 12}$ are weight 12 modular forms, and it is a classical fact that the Eisenstein series E_4^3, E_6^2 generically globally generate $\lambda^{\otimes 12}$. An elliptic curve $E : y^2 = x^3 + Ax + B$ gives a $\mathbb{Q}$-point $x : \operatorname{Spec}\mathbb{Q} \to \mathcal{X}(1)$. The assumption about scaling the sections corresponds to choosing a minimal Weierstrass equation for E. Proposition 3.5.2 then says that

$$\mathrm{ht}_{\lambda}(x) = \frac{1}{12} \log \max\{|A|^3, |B|^2\} + O_{\mathcal{X}(1)}(1),$$

which is, up to the constant $O_{\mathcal{X}(1)}(1)$, a twelfth of the logarithmic naive height of E. Thus,

$\mathrm{Ht}_\lambda^{12}(x)$ is a constant multiple of the naive height $\mathrm{ht}(E)$.

3.5.2 The ring of modular forms of low level

Since modular forms are sections of powers of the Hodge bundle, we will rely on the structure of the rings of modular forms of $\mathcal{X}_0(N)$ quite heavily. This subsection summarizes part of the work of Hayato and Tomohiko in [18].

Notation 3.5.4. Let $M_k(N)$ denote the space of modular forms for $\Gamma_0(N)$ of weight k. We let $M(N) = \bigoplus_k M_k(N)$ be the entire ring of modular forms for $\Gamma_0(N)$.

- E_k: classical Eisenstein series of weight k. Note that $E_k \in M_k(1)$. For k an even integer, E_k is given by
$$1 - \frac{2k}{B_k} \sum_{n=1}^{\infty} \sigma_{k-1}(n) q^n$$
where B_k is the k-th Bernoulli number and σ_k is sum of k-th powers of divisors function.
- For a modular form f and an integer h, let $f^{(h)}(q) = f(qh)$.
- For any $N \geq 1$, let $C_N = \frac{1}{\gcd(N-1,24)}(NE_2^{(N)} - E_2) \in M_2(N)$.
- For a prime p, let $\alpha_p = \frac{1}{240}(E_4 - E_4^{(p)}) \in M_4(p)$.
- For the definitions of β_N for general N and for α_N for N composite, we refer the reader to [18].

Proposition 3.5.5 ([18], Theorems 1,2)**.** *Maintaining the above notation, the rings of modular forms for* $\Gamma_0(N)$ *for* $N \in \{2, 3, 4, 5, 6, 8, 9\}$ *are as follows.*

N	*Degrees of generators*	$M(N)$
2	$(2,4)$	$\mathbb{C}[C_2,\alpha_2]$
3	$(2,4,6)$	$\mathbb{C}[C_3,\alpha_3,\beta_3]/(O_3)$
4	$(2,2)$	$\mathbb{C}[C_2,C_4]$
5	$(2,4,4)$	$\mathbb{C}[C_5,\alpha_5,\beta_5]/(O_5)$
6	$(2,2,2)$	$\mathbb{C}[C_3^{(2)},\alpha_6,\beta_6]/(O_6)$
8	$(2,2,2)$	$\mathbb{C}[C_4^{(2)},\alpha_4,\alpha_4^{(2)}]/(O_8)$
9	$(2,2,2)$	$\mathbb{C}[C_3,\alpha_9,\beta_9]/(O_9)$

Table 4: Rings of modular forms of low level

Here the O_n's are explicit polynomials whose form we will mention later.

Since the ring of modular forms is graded by weight, the degrees in Table 4 refer to the weights in which the corresponding rings are generated. In what follows, we will use the structure of the ring of modular forms of the levels in Table 4 to count points of bounded height. The reason we restrict to these cases is that for such N, the rings of modular forms are easier to handle. For some other N (e.g. see Chapter 4), this method reduces to a problem of counting integral points on more complicated varieties.

3.5.3 Counting results

Our counting results will be split into three parts: the first, for $N = 2, 4$ corresponds to the N for which $M(N)$ is freely generated. The second part, is for $N = 3, 6, 8, 9$. These are the levels N for which the corresponding O_N's in Table 4 have a similar form. The last part is for $N = 5$, which has to be dealt with separately because O_5 has a starkly different form, and thus requires different counting techniques.

Notation 3.5.6. Let $(a_1, \dots a_k) \in \mathbb{Z}^k$, $\mathbf{p} = (p_1, \dots p_k) \in \mathbb{Z}^k_{>0}$ be two tuples of integers. Let $n \in \mathbb{Z}_{>0}$ such that $\operatorname{lcm}(p_1, \dots p_k) | n$. We will say that the pair, $((a_1, \dots a_k), \mathbf{p})$ satisfies condition (†) if for any prime p,

$$p^n \nmid \gcd_i(|a_i|^{p_i}).$$

Condition (†) reflects the minimality condition in Proposition 3.5.2. For the rest of this section, we will only consider (†) for $n = 12$, unless mentioned otherwise.

The cases $N = 2, 4$: From Table 4, we see that $M(2) \cong \mathbb{C}[x, y]_{(2,4)}$ and $M(4) \cong \mathbb{C}[x, y]_{(2,2)}$. For $N = 2$, $\lambda^{\otimes 12}$ is globally generated by C_2^6 and and α_2^3. Let $x : \operatorname{Spec} \mathbb{Q} \to \mathcal{X}_0(2)$. Let $a = x^*(C_2)$ and $b = x^*(\alpha_2)$. Taking these to be in minimal form implies that $a, b \in \mathbb{Z}$ with $p^{12} \nmid \gcd(a^6, b^3)$. Then, by Proposition 3.5.2, we see that

$$\mathrm{ht}_\lambda(x) = \frac{1}{12} \log \max\{|a|^6, |b|^3\} + O_{\mathcal{X}_0(2)(\mathbb{Q})}(1).$$

Thus we have that,

$$\begin{aligned} \mathcal{N}(2, X) &\asymp \#\{x \in \mathcal{X}_0(2)(\mathbb{Q}) \mid \mathrm{Ht}_\lambda^{12}(x) < X\} \\ &\asymp \#\{(a, b) \in \mathbb{Z}^2 \mid ((a, b), (6, 3)) \text{ satisfies } (\dagger), \max\{|a|^6, |b|^3\} < X\}. \end{aligned}$$

By a similar argument, observing that C_2^6 and C_4^6 globally generate λ^{12} on $\mathcal{X}_0(4)$, we set $a = x^*(C_2)$ and $b = x^*(C_4)$. Thus,

$$\mathcal{N}(4, X) \asymp \#\{(a, b) \in \mathbb{Z}^2 \mid ((a, b), (6, 6)) \text{ satisfies } (\dagger), \max\{|a|^6, |b|^6\} < X\}.$$

In each of these cases, our counting problem reduces to counting integers in a box with certain divisibility conditions. The set we need to count has the form $\{(a, b) \in \mathbb{Z}^2 \mid |a| < M, |b| < N, p^{12} \nmid \gcd(a^{p_1}, b^{p_2})\}$ for some constants p_1, p_2, M and N. The set $\{(a, b) \in \mathbb{Z}^2 \mid |a| < M, |b| < N, \gcd(a, b) = 1\}$ is always a subset of this set, and in particular, has size a constant multiple of MN. Thus the condition (†) does not affect the asymptotic growth rate.

Proposition 3.5.7. *Maintaining the above notation, we have:*

$$\mathcal{N}(2, X) \asymp X^{1/2}$$
$$\mathcal{N}(4, X) \asymp X^{1/3}.$$

Note that this agrees with the asymptotics in [17], [29] as well as the conclusions of Table 3 in Section 3.4.

The cases $N = 3, 6, 8, 9$**:** These cases are similar because of the similarity in form on the O_N's in Table 4. More precisely, we have from [18]:

- $O_3 = \alpha_3^2 - C_3\beta_3$
- $O_6 = \alpha_6^2 - C_3^{(2)}\beta_6$
- $O_8 = \alpha_4^2 - C_4^{(2)}\alpha_4^{(2)}$
- $O_9 = \alpha_9^2 - C_3\beta_9$.

In order to deal with these cases uniformly, we must introduce some notation. For $(a, b, c) \in \mathbb{Z}^3$ and $\mathbf{p} = (p_a, p_b, p_c) \in \mathbb{Z}^3_{>0}$, define

$$\mathrm{Ht}^{\mathbf{p}}(a, b, c) = \max\{|a|^{p_a}, |b|^{p_b}, |c|^{p_c}\}.$$

Later, for each N, we will fix a choice of $\mathbf{p}$ that makes this height compatible with $\mathrm{Ht}^{12}_{\lambda}$ on $\mathcal{X}_0(N)$. We will be interested in the following counting functions:

$$\mathcal{N}(\mathbf{p}, X) := \#\{(a, b, c) \in \mathbb{Z}^3 \mid \mathrm{Ht}^{\mathbf{p}}(a, b, c) < X, b^2 = ac\},$$
$$\mathcal{N}(\mathbf{p}, X, \dagger) := \#\{(a, b, c) \in \mathbb{Z}^3 \mid \mathrm{Ht}^{\mathbf{p}}(a, b, c) < X, b^2 = ac, \text{ and } ((a, b, c), \mathbf{p}) \text{ satisfies } (\dagger)\}.$$

Lemma 3.5.8. *Let n be any integer such that $\mathrm{lcm}(p_a, p_b, p_c)|n$. There is a positive constant C that depends only on $\mathbf{p}$ and n such that:*

$$\mathcal{N}(\mathbf{p}, X) = C\, X^{1/p_b} \log(X) + X^{1/p_c} + O(X^{1/p_a}).$$

Proof. We start by noting that we must have $|a| < X^{\frac{1}{p_a}}$, $|b| < X^{\frac{1}{p_b}}$ and $|c| < X^{\frac{1}{p_c}}$. Now

suppose $a \neq 0$. Then

$$\begin{aligned}\sum_{\substack{|a|<X^{\frac{1}{p_a}} \\ a\neq 0}} \sum_{|c|<X^{\frac{1}{p_c}}} \sum_{\substack{|b|<X^{\frac{1}{p_b}} \\ b^2=ac}} 1 &= \sum_{\substack{|a|<X^{\frac{1}{p_a}} \\ a\neq 0}} \sum_{|b|<X^{\frac{1}{p_b}}, a|b^2} 1 \\ &= \sum_{\substack{|a|<X^{\frac{1}{p_a}} \\ a\neq 0}} \left(\frac{X^{1/p_b}}{a} + O(1)\right) \\ &= CX^{1/p_b}\log(X) + O(X^{1/p_a}).\end{aligned}$$

One might worry here that the 'error' term, X^{1/p_a}, is actually bigger than the main terms. However, for all of our cases $p_a \geq p_b, p_c$, so X^{1/p_a} will indeed be an error term. If $a = 0$, then b is necessarily 0 too. Thus we are reduced to counting the set $\{c \in \mathbb{Z} \mid c < X^{1/p_c}\}$, which has size $X^{1/p_c} + O(1)$. □

Claim 3.5.9. *If p_a, p_b and p_c are not all equal to n, then $\mathcal{N}(\mathbf{p}, X, \dagger)$ is a positive proportion of $\mathcal{N}(\mathbf{p}, X)$.*

Proof. Without loss of generality, suppose $p_a \neq n$. If $a = 0$, then

$$\#\{c \in \mathbb{Z} \mid c \text{ is } p_c\text{th power free}, |c| < X^{1/p_c}\}$$

is a positive proportion of $\#\{c \in \mathbb{Z} \mid |c| < X^{1/p_c}\}$. In particular these sizes differ by a factor of $\zeta(p_c)$. Now suppose $a \neq 0$. Then the set of triples satisfying (†) contains those for which a is squarefree. The proof of Lemma 3.5.8 shows that the set of such triples has size a constant times $X^{1/p_b}\log(X)$ as well. This proves the claim. □

We have therefore proved the following proposition.

Proposition 3.5.10. *Maintaining the above notation:*

$$\begin{aligned}\mathcal{N}(3, X) &\asymp X^{1/2}, \\ \mathcal{N}(6, X) &\asymp X^{1/6}\log(X),\end{aligned}$$

$$\mathcal{N}(8, X) \asymp X^{1/6} \log(X),$$
$$\mathcal{N}(9, X) \asymp X^{1/6} \log(X).$$

Proof. Since we only care about the 12th power of the Hodge bundle, we will take $n = 12$. From Table 4, we observe that for the following choices of $\mathbf{p}$, $\mathcal{N}(N, X) \asymp \mathcal{N}(\mathbf{p}, X, \dagger)$:

- $N = 3$: $\mathbf{p} = (6, 3, 2)$,
- $N = 6, 8, 9$: $\mathbf{p} = (6, 6, 6)$.

The proposition now follows from Claim 3.5.9 and Lemma 3.5.8. □

Note that for $N = 3$, the elliptic point on $\mathcal{X}_0(3)$ corresponds to the point where a is 0. Indeed, one may think of $\mathcal{X}_0(3)$ as being cut out by the octic $b^2 - ac$ inside $\mathbb{P}(2, 4, 6)$. The point $[0 : 0 : 1]$ has an automorphism group of size 6. If $a \neq 0$, then from Lemma 3.5.8, we get an asymptotic growth rate of $X^{1/3} \log(X)$ which agrees with that obtained from Theorem 3.3.3.

The case $N = 5$: Note that this is one of the cases that cannot be tackled by the methods in Sections 3.2 and 3.3. We first give an upper bound for $\mathcal{N}(5, X)$, and then use a simple sieving argument to refine it into an asymptotic.

The ring of modular forms, $M(5)$ is generated by three modular forms, C_5, α_5 and β_5 of weights 2,4 and 4 respectively. The relation between these forms is

$$O_5 = \alpha_5^2 - \beta_5(C_5^2 + 4\alpha_5 - 8\beta_5). \tag{3.7}$$

Set $n = 12$ and $\mathbf{p} = (6, 3, 3)$. Proceeding analogously as before, we must count integers (a, b, c) with $\mathrm{Ht}^{\mathbf{p}}(a, b, c) < X$ such that

$$b^2 - a^2c - 4bc + 8c^2 = 0, \tag{3.8}$$

and the pair $((a, b, c), \mathbf{p})$ satisfies the minimality condition (†). If $\alpha_5 = 0$, then $\beta_5 = 0$, and we

get $\asymp X^{1/6}$ elliptic curves, which is the trivial lower bound. If $C_5 = 0$, we get the two points of $\mathcal{X}_0(5)$ that have automorphism group μ_4. Each of these is defined over $\mathbb{Q}(i)$ and doesn't contribute to the rational points on $\mathcal{X}_0(5)$.

We obtain the upper bound by counting integer triples (a, b, c) without the minimality condition (†). Equation 3.8 can be rearranged to one of the form:

$$(4b - 8c)^2 + (8c - a^2)^2 = a^4.$$

For any integer n, let $r_2(n)$ denote the number of ways of writing an integer as a sum of two squares. An upper bound can be proved by summing $r_2(a^4)$ over all $a < X^{1/6}$.

Lemma 3.5.11 ([4], Chapter XV)**.** *Let $n \in \mathbb{Z}_{>0}$ have factorization:*

$$n = 2^{a_0} p_1^{e_1} \dots p_r^{e_r} q_1^{2f_1} q_2^{2f_2} \dots q_s^{2f_s}$$

where the p_i's are $\equiv 1 \mod 4$ and the q_i's are $\equiv 3 \mod 4$. Define $B(n) = \prod_{i=1}^{r}(e_i + 1)$. Then:

$$r_2(n) = 4B(n)$$

Remark 3.5.12. This is a well known result. Note that the constant in front of B is different depending on whether one takes into account signs and order. But this will not make a difference to our result, since we are only interested in the asymptotic growth rate.

We will now focus on the sum:

$$\sum_{|a|<X^{1/6}} B^{(4)}(n)$$

where we define $B^{(4)}(n)$ to be $B(n^4)$, for notational convenience. Note that if $p \equiv 1 \mod 4$, $B^{(4)}(p^k) = 4k + 1$. If $p = 2$ or $3 \mod 4$, then $B^{(4)}(p^k) = 1$ for any k. Thus, $B^{(4)}(n)$ is a multiplicative (although not completely multiplicative) function.

Proposition 3.5.13. *Maintaining the above notation, there is a constant $c > 0$ such that for any $0 < \delta < 1/6$,*

$$\sum_{|n|<X^{1/6}} B^{(4)}(n) = cX^{1/6}(\log(X))^2 + O(X^{1/6-\delta}).$$

Proof. Consider the Dirichlet series: $\displaystyle\sum_{n\geq 1} \frac{B^{(4)}(n)}{n^s}$. By multiplicativity, this can be written as the Euler product:

$$\prod_{p\equiv 1 \mod 4} \left(\sum_{k\geq 0}(4k+1)p^{-ks}\right) \prod_{p\equiv 3 \mod 4} \left(\sum_{k\geq 0} p^{-ks}\right)\left(\sum_{k\geq 0} 2^{-ks}\right).$$

We now simplify this expression.

$$\prod_{p\equiv 3 \mod 4} \left(\sum_{k\geq 0} p^{-ks}\right) = \prod_{p\equiv 3 \mod 4} \frac{1}{1-p^{-s}}.$$

$$\begin{aligned}
\prod_{p\equiv 1 \mod 4} \left(\sum_{k\geq 0}(4k+1)p^{-ks}\right) &= \prod_{p\equiv 1 \mod 4} \left(4\sum_{k\geq 0} kp^{-ks} + \sum_{k\geq 0} p^{-ks}\right) \\
&= \prod_{p\equiv 1 \mod 4} \left(\frac{4p^{-s}}{(1-p^{-s})^2} + \frac{1}{1-p^{-s}}\right) \\
&= \prod_{p\equiv 1 \mod 4} \left(\frac{1+3p^{-s}}{(1-p^{-s})^2}\right).
\end{aligned}$$

Thus:

$$\begin{aligned}
\sum_{n\geq 1} \frac{B^{(4)}(n)}{n^s} &= \prod_{p} \left(\frac{1}{1-p^{-s}}\right) \prod_{p\equiv 1 \mod 4} \left(\frac{1+3p^{-s}}{1-p^{-s}}\right) \\
&= \zeta(s) \prod_{p\equiv 1 \mod 4} \left(\frac{1+3p^{-s}}{1-p^{-s}}\right).
\end{aligned}$$

Now, let $\chi(p)$ denote the usual Legendre Symbol $\left(\frac{-1}{p}\right)$. Let $K(s) = \frac{1-2^{-s}}{1+3.2^{-s}}$. Then,

$$\begin{aligned}\Psi(s) := \prod_{p\equiv 1 \mod 4}\left(\frac{1+3p^{-s}}{1-p^{-s}}\right) &= K(s)\prod_p\left(\frac{1+3p^{-s}}{1-p^{-s}}\right)^{\frac{1+\chi(p)}{2}} \\ &= K(s)\prod_p\left(1+\frac{4p^{-s}}{1-p^{-s}}\right)^{\frac{1+\chi(p)}{2}} \\ &= K(s)\prod_p\left(1+\frac{1}{2}(1+\chi(p))\frac{4p^{-s}}{1-p^{-s}}+\dots\right) \\ &= K(s)\prod_p\left(1+2(1+\chi(p))p^{-s}+\text{higher powers of } p^{-s}\right).\end{aligned}$$

Consider the Dirichlet L-function $L(s,\chi) = \prod_p(1-\chi(p)p^{-s})^{-1}$. Since

$$\left(1+2(1+\chi(p))p^{-s}+\dots\right)\left(1-\chi(p)p^{-s}\right)^2 = 1+2p^{-s}\dots,$$

we see that $\Psi(s)L(s,\chi)^{-2}$ has a pole of order 2 at $s=1$ and converges for $Re(s)>1$. We know that $L(s,\chi)$ is holomorphic at $s=1$. Thus $\sum_{n\geq 1}\frac{B^{(4)}(n)}{n^s}$ has a pole of order 3 at $s=1$. The proposition now follows from the standard Tauberian theorem (2.5.1). □

Proposition 3.5.14. *There is an absolute constant $K>0$ such that for $X>K$, $\mathcal{N}(5,X) \asymp X^{1/6}\log(X)^2$.*

Proof. The main ingredient here is the upper bound proved in Proposition 3.5.13. To refine this to give an asymptotic growth rate, we must count only the minimal (a,b,c). If a triple is *non-minimal*, then there exists a prime p such that $p^2|a$, $p^4|b$ and $p^4|c$. Let p be such a prime. Then the number of such triples is in bijection with the number of ways of writing a^4 as a sum of two squares, say $a^4 = A^2+B^2$, such that $p^4|A$ and $p^4|B$. This is the same as the number of ways of writing $(a/p^2)^4$ as a sum of two squares. Therefore the number of triples that are

non-minimal at p can be calculated by:

$$\sum_{|n|<X^{1/6}/p^2} B^{(4)}(n).$$

By Proposition 3.5.13, this has the same asymptotic growth rate as

$$c.\frac{X^{1/6}}{p^2}\log\left(\frac{X}{p^{12}}\right)^2 = (c/p^2)X^{1/6}\left(\log(X)^2 - 2\log(X)\log(p^{12}) + \log(p^{12})^2\right).$$

where c is independent of p. Thus,

$$cX^{1/6}\log(X)^2 - c\frac{X^{1/6}}{p^2}\log\left(\frac{X}{p^{12}}\right)^2 = cX^{1/6}\log(X)^2\left(1 - \frac{1}{p^2} - 24\log(X)^{-1}\frac{\log(p)}{p^2} + 144\log(X)^{-2}\frac{\log(p)^2}{p^2}\right).$$

We now examine the product

$$\prod_{p^2<X^{1/6}}\left(1 - \frac{1}{p^2} - 24\log(X)^{-1}\frac{\log(p)}{p^2} + 144\log(X)^{-2}\frac{\log(p)^2}{p^2}\right).$$

For large enough X, this product is bounded both above and below by positive constants. This proves the proposition.

□

Chapter 4

Future work

4.1 On counting curves with given p-torsion

For $p = 2$ the Artin-Schreier locus $\mathscr{AS}_g$ coincides with the hyperelliptic locus $\mathscr{H}_g$. However, in general, $\mathscr{AS}_g$ is not irreducible. In [30], the authors give the following characterization for the irreducibility of $\mathscr{AS}_g$:

Proposition 4.1.1 ([30] Corollary 1.2)**.** *The moduli space $\mathscr{AS}_g$ is irreducible in exactly the following cases: (a) $p = 2$, or (b) $g = 0$ or $g = \frac{p-1}{2}$, or (c) $p = 3$ and $g = 2, 3, 5$.*

It is interesting to ask whether the reducibility of $\mathscr{AS}_g$ completely explains the probability obtained in theorem 2.5.6. That is, for each g, let $\overline{\mathscr{AS}}_{g,g}$ denote the closure of the ordinary locus inside $\mathscr{AS}_g$. Then, is:

$$\lim_{X\to\infty} \frac{\#\{\mathcal{C} \in \mathscr{AS}_{g,g} \mid q^g < X\}}{\#\{\mathcal{C} \in \overline{\mathscr{AS}}_{g,g}(\mathbb{F}_q) \mid q^g < X\}} \tag{4.1}$$

positive?

On a different note, for $p \geq 3$

$$\tilde{\psi}_{p,Q}(s) = \left(1 + \sum_{i=0}^{p-2} |Q|^{(i+1)-(i+2)s} - \sum_{i=0}^{p-2} |Q|^{i-(i+2)s}\right) \prod_{i=0}^{p-2} (1 - |Q|^{(i+1)-(i+2)s})$$

and let $\tilde{\psi}_p(s) = \prod_Q \tilde{\psi}_{p,Q}(s)$. Let $s_0 = \lceil \frac{2}{p}(g + p - 1) \rceil$. Then one can show, by a similar calculating as in Chapter 2, that the probabilty that an Artin-Schreier curve has p-rank $\geq s_0$ is

bounded below by the quantity

$$\frac{\tilde{\psi}_p(1)}{\psi_p(1)}.$$

This quantity comes from counting the rational points on the $d-1$-dimensional components of each p-rank stratum $\mathscr{AS}_{g,s}$ with $s \geq s_0$. As expected, for any fixed p, this number gets closer to 1 as q increases. Further, if $\tilde{s} > s_0$, then the probability that an Artin-Schreier curve has p-rank $\geq \tilde{s}$ is 0. For $\tilde{s} < s_0$, it would be interesting to calculate the probability that the p-rank is $\geq \tilde{s}$. This would give us a better understanding of the distribution of p-ranks in the Artin-Schreier locus.

4.2 On counting points on $\mathcal{X}_0(N)$

This report raises multiple questions, some that we believe can be answered by pushing further the methods used here, and some that require different approaches. The first question is about $\mathcal{X}_0(7)$. We believe that the ideas of Section 3.3 can be generalized to count points on $\mathcal{X}_0(7)$, since $\mathcal{X}_{1/2}(7)$ is a stacky curve with two stacky points. In this case, one must generalize Proposition 3.3.1 to the case where f and g are not necessarily coprime. The tricky bit here turns out to be the analogue of Lemma 3.3.7.

One might wonder whether one can count rational points on $\mathcal{X}_0(7)$ via the framework in [13], as we did for some values of N in Section 3.5. The issue with this is that for each level not listed in Table 4, the ring of modular forms is quite complicated. Using relations between the generators of these rings to count points on $\mathcal{X}_0(N)$ can lead to very hard counting problems. For instance, the problem of counting rational points on $\mathcal{X}_0(7)$ can be rephrased in terms of counting *integral* points on the intersection of one cubic and two quadric hypersurfaces in $\mathbb{A}^5$. More precisely, one must count tuples (a, b, c, d, e) of integers satisfying

- $c^2 - ae = 0$,
- $ce - bd = 0$,
- $b^2 - c(a^2 + 7b - 19c) = 0$,

- No 12th power divides $\gcd(|a|^6, |b|^3, |c|^3, |d|^2, |e|^2)$.

This gets more complicated with higher N, at least as far using the description in [18] goes. For these higher N, if one were to find a smaller set of modular forms that could *both* globally generate $\lambda^{\otimes 12}$ and had simpler relations among them, then one could perhaps count points on the corresponding $\mathcal{X}_0(N)$ more easily. We do not know at this time if that is indeed possible.

There is of course the question of an exact asymptotic as opposed to an asymptotic growth rate. More precisely, one can ask if the limit:

$$c_N := \lim_{X \to \infty} \frac{\mathcal{N}(N, X)}{h_N(X)}$$

exists and what its value is. The case $N = 2$ is known due to [17], $N = 3$ due to [28] and $N = 4$ due to [29]. It would be interesting to calculate the values for other N, perhaps using the precise definition of logarithmic height from [13].

The stacky Batyrev-Manin-Malle conjecture. As mentioned in Chapter 1, for a scheme V and an ample line bundle L on it, the Batyrev-Manin conjecture predicts that there are constants $a(L)$ and $b(L)$ such that the number of rational points on V of height bounded by a number B grows like

$$B^{a(L)} \log(B)^{b(L)}.$$

Here the height refers to the height with respect to the line bundle L. The weaker analogue states that the number of rational points should grow like $B^{a(L)+\epsilon}$. In [13], the authors make a similar conjecture for stacks, which they call the 'Weak stacky Batyrev-Manin-Malle conjecture'. For each of the modular curves considered in this paper, as well as those in [17], the asymptotic growth rate seems to be of the same form as predicted, but it would be interesting to verify if the constants match the constants in [13]. This is work in progress.

Appendix A

Explicit description of $\mathcal{X}_{1/2}(N)$

A.1 Modular descriptions of cusps

The main reference for this section is [40]. Let C_n be a Néron n-gon. Each irreducible component of C_n is isomorphic to $\mathbb{P}^1$. For each $i < n$, the i-th component is glued to the $i+1$-th component by gluing $\infty \in \mathbb{P}^1{}_{(i)}$ to $0 \in \mathbb{P}^1{}_{(i+1)}$, and analogously for the n-th component. The smooth part of C_n, denoted C_n^{sm}, is isomorphic to $\mathbb{G}_m \times \mathbb{Z}/n\mathbb{Z}$. The group structure on C_n^{sm} is given by the usual group structures on each component. The automorphism group of C_n is given by $\mu_n \times \langle \text{inv} \rangle$, where $\zeta \cdot (x, i) = (\zeta^i x, i)$ for ζ a primitive n-th root of unity and $\text{inv} : (x, i) \mapsto (x^{-1}, -i)$. A *generalized elliptic curve* E over S is a flat, finitely presented map $E \to S$ whose geometric fibers are either

- elliptic curves (hence smooth and equipped with a group structure), or
- a Néron n-gon for some $n \geq 1$ equipped with a group structure on the smooth part.

Let $\overline{\text{Ell}}_n$ denote the moduli space of generalized elliptic curves whose degenerate fibers are n-gons. In general, for a moduli stack $\mathcal{X}$ of generalized elliptic curves, and a positive integer n, let $\mathcal{X}_{(n)}$ denote the substack of $\mathcal{X}$ that parametrizes generalized elliptic curves whose degenerate fibers are n-gons.

A.1.1 $\Gamma_1(N)$ and $\Gamma_0(N)$ structures

Let N be a positive integer, $n|N$ and E an elliptic curve over S. A $\Gamma_1(N)$ structure on E is the following data:

- A homomorphism $\alpha : \mathbb{Z}/N\mathbb{Z} \to E^{sm}(S)$ such that $D = \sum_{a \in \mathbb{Z}/N\mathbb{Z}} [\alpha(a)]$ is an effective

Cartier divisor on E forming an S-subgroup scheme of E.

- If the fiber over some point in S is an n-gon, then the divisor D intersects every irreducible component of the n-gon. This criterion is equivalent to the ampleness of D.

The stack $\mathcal{X}_1(N)$ parametrizes generalized elliptic curves with a $\Gamma_1(N)$ structure. Further, we have that $\mathcal{X}_1(N) = \bigcup_{n|N} \mathcal{X}_1(N)_{(n)}$.

Unlike the definition of a $\Gamma_1(N)$ structure, which is fairly intuitive given our understanding of $\mathcal{Y}_1(N)$, the definition of a $\Gamma_0(N)$ structure takes more work. To explain this, let us first define a naive $\Gamma_0(N)$ structure on a generalized elliptic curve E/S. This consists of the following data:

- A homomorphism $\alpha : \mathbb{Z}/N\mathbb{Z} \to E^{sm}$ such that $D = \sum_{a \in \mathbb{Z}/n\mathbb{Z}} [\alpha(a)]$ is an ample, effective divisor on E.
- The image of α is an S subgroup scheme of E^{sm}.

One defines $\mathcal{X}_0(N)^{naive}$ as the moduli space parametrizing generalized elliptic curves with an naive $\Gamma_0(N)$ structure. As before, $\mathcal{X}_0(N)^{naive} = \bigcup_{n|N} \mathcal{X}_0(N)^{naive}$.

Consider the modular curve $\mathcal{X}_0(p^2)$ for some prime p. Let E/S be a generalized elliptic curve whose degenerate fiber is a p-gon, equipped with a naive $\Gamma_0(p^2)$ structure G_E. On the degenerate fiber, the group G generated by $(\zeta_{p^2}, 1)$ gives a naive $\Gamma_0(p^2)$ structure, and the pair (C_p, G) has automorphism group $\mu_p \times \langle \text{inv} \rangle$. On the other hand, the image of (E, G_E) in $\mathcal{X}_0(1)$ is a generalized elliptic curve whose degenerate fiber has automorphism group $\langle \text{inv} \rangle$. In particular, the map $\mathcal{X}_0(N)^{naive} \to \mathcal{X}_0(1)$ is not representable (see criterion for representability in Lemma 3.2.2 in [40]). This does not agree with the construction of $\mathcal{X}_0(N)$ in [10], which is what we are using.

The correct definition of a $\Gamma_0(N)$ structure is a little bit long-winded, so we do not define it here. Instead, we explain how to construct one from a naive $\Gamma_0(N)$ structure, which is sufficient for our purposes. Let $n|N$ and let $d(n) = \frac{n}{\gcd(n, N/n)}$, and let E/S be a generalized elliptic curve.

Let G be a naive $\Gamma_0(N)$ structure on E, let E^∞ denote a degenerate fiber of E that is an n-gon. Let G^∞ denote the fiber of G on E^∞. Consider the torsion subgroup $E^{\infty,sm}[d(n)] \in E^{\infty,sm}$. Define the contraction of E along $E^{sm}[d(n)]$ by leaving the non-degenerate fibers intact, and on each E^∞ in the degenerate n-gon locus, by contracting any components that don't intersect $E^{\infty,sm}[d(n)]$ to a point. Thus, the image of E^∞ is a $d(n)$-gon. A new elliptic curve E'/S may now be constructed by gluing together the contractions of E/S for each $n|N$ along the non-degenerate locus. The image of G under these contractions, gives a $\Gamma_0(N)$ structure. Note that a $\Gamma_0(N)$ structure remembers G as well as the images of all the degenerate fibers of G under the contractions.

Let $\mathcal{X}_0(N)$ be the modular curve paramterizing generalized elliptic curves with a $\Gamma_0(N)$ structure. The following lemma from [40] is probably the best way to understand the relation between the n-part of $\mathcal{X}_0(N)^{naive}$ and the $d(n)$ part of $\mathcal{X}_0(N)$.

Lemma A.1.1 ([40], Lemma 5.1.2)**.** *There is a commutative diagram:*

$$\begin{array}{ccc} \mathcal{X}_0(N)^{naive}_{(n)} & \longrightarrow & \overline{\mathrm{Ell}}_n \\ \downarrow & & \downarrow \\ \mathcal{X}_0(N)_{(d(n))} & \longrightarrow & \overline{\mathrm{Ell}}_{d(n)} \end{array}$$

where the vertical maps are contractions.

A.1.2 Construction of $\mathcal{X}_{1/2}(N)$ at the cusps

Recall the definition of $\mathcal{Y}_{1/2}(N)$ from Chapter 3. We claimed, in Section 3.2.1, that the construction makes sense for $\mathcal{X}_{1/2}(N)$, i.e. for generalized elliptic curves, via a similar process. We now outline a proof of the claim. The process is the same as obtaining a $\Gamma_0(N)$ structure via a naive $\Gamma_0(N)$ structure. In order to make our description less wordy, we will let C_n denote the cusp parametrizing generalized elliptic curves whose degenerate fibers are n-gons, and describe the construction on C_n directly. The fiber of $\mathcal{X}_1(N) \to \mathcal{X}_0(N)^{naive}$ over C_n consists of generators of the $\Gamma_0(N)$ structure. Thus, it makes sense to define $\mathcal{X}_{1/2}(N)^{naive}$ as the fiberwise quotient of $\mathcal{X}_1(N) \to \mathcal{X}_0(N)^{naive}$ by an index two subgroup of $(\mathbb{Z}/N\mathbb{Z})^\times$. We now declare that

the fiber of $\mathcal{X}_{1/2}(N) \to \mathcal{X}_0(N)$ over a cusp consists of the data:

- The fiber above the corresponding point $\mathcal{X}_{1/2}(N)^{naive} \to \mathcal{X}_0(N)^{naive}$.
- The $\Gamma_0(N)$ structure on the cusp.

The second condition helps rigidify our structure.

As an example, consider the case $N = 9$. In this case, we are dealing with 3 possible n-gons in the degenerate fiber: C_1, C_3 and C_9.

- There is exactly one (naive as well as not) $\Gamma_0(9)$ structure on C_1, namely the subgroup generated by ζ_9, a primitive 9th root of unity. The fiber of the map $\Phi_9 : \mathcal{X}_1(N) \to \mathcal{X}_0(N)$ corresponds to the generators of this subgroup, namely $\{\zeta_9^i \mid i = 1, 2, 4, 5, 7, 8\}$. The two points in the fiber of $\mathcal{X}_{1/2}(9) \to \mathcal{X}_0(9)$, therefore, correspond to the cosets $\{\zeta_9, \zeta_9^4, \zeta_9^7\}$ and $\{\zeta_9^2, \zeta_9^5, \zeta_9^8\}$.
- Consider the cusp C_3. One naive $\Gamma_0(9)$ structure on C_3 is generated by the pair $(\zeta_9, 1)$. Further, $d(3) = 1$ and so $E^{sm}[d(3)] = 0$. To obtain the contraction corresponding $\Gamma_0(9)$ structure therefore, one contracts the degenerate fiber to a C_1, and the image of $\langle(\zeta_9, 1)\rangle$ under this contraction, is the subgroup generated by $(\zeta_9^3, 0)$. The data of the $\Gamma_0(9)$ structure consists of both the data of the original naive structure, and its contraction.

  To obtain the fibers of $\mathcal{X}_{1/2}(9) \to \mathcal{X}_0(9)$, consider the points over $\mathcal{X}_{1/2}(9)^{naive} \to \mathcal{X}_0(9)^{naive}$. The fibers above $\langle(\zeta_9, 1)\rangle$ correspond to the cosets $\{(\zeta_9, 1), (\zeta_9^4, 1), (\zeta_9^7, 1)\}$ and $\{(\zeta_9^2, 2), (\zeta_9^5, 2), (\zeta_9^8, 2)\}$ respectively. Note as an aside, that each of these cosets has an automorphism group of size 3. We rigidify these points by adding in the data of the $\Gamma_0(9)$ structure above.
- Consider the naive $\Gamma_0(9)$ structure on C_9 generated by the element $(\zeta_9, 1)$. This is also a $\Gamma_0(9)$ structure, since $d(9) = 9$. The fiber of $\mathcal{X}_{1/2}(9)^{naive} \to \mathcal{X}_0(9)^{naive}$ thus corresponds to the two cosets $\{(\zeta_9, 1), (\zeta_9^4, 4), (\zeta_9^7, 7)\}$ and $\{(\zeta_9^2, 2), (\zeta_9^5, 5), (\zeta_9^8, 8)\}$.

A.2 Equations for $\mathcal{X}_{1/2}(N)$ for some N

For $N \in \{8, 9, 12, 16, 18\}$, we give the equations for $\mathcal{X}_{1/2}(N)$ using the hauptmoduln from [25]. We use the notation of Chapter 3 where in we give polynomials $f_N(t), g_N(t) \in \mathbb{Q}[t]$ such that every elliptic curve arising as a rational point on $\mathcal{X}_{1/2}(N)$ is isomorphic to one of the form,

$$y^2 = x^3 + f_N(t)x + g_N(t).$$

N	$f_N(t)$	$g_N(t)$
8	$\frac{1}{3}(t^4 + 256t^3 + 5120t^2 + 32768t + 65536)$	$\frac{1}{2}(t^2 + 32t + 128)(t^4 - 512t^3 - 10240t^2 - 65536t - 131072)$
9	$\frac{1}{3}(t + 9)(t^3 + 243t^2 + 2187t + 6561)$	$\frac{1}{2}(t^6 - 486t^5 - 24057t^4 - 367416t^3 - 2657205t^2 - 9565938t - 14348907)$
12	$\frac{1}{3}(t^2 + 12t + 24)(t^6 + 252t^5 + 4392t^4 + 31104t^3 + 108864t^2 + 186624t + 124416)$	$\frac{1}{2}(t^4 + 36t^3 + 288t^2 + 864t + 864)^2(t^8 - 504t^7 - 14832t^6 - 179712t^5 - 1175040t^4 - 4478976t^3 - 9953280t^2 - 11943936t - 5971968)^2$
16	$\frac{1}{3}(t^8+256t^7+5632t^6+53248t^5+282624t^4+917504t^3 + 1835008t^2 + 2097152t + 1048576)$	$\frac{1}{2}(t^4 + 32t^3 + 192t^2 + 512t + 512)(t^8 - 512t^7 - 11264t^6 - 106496t^5 - 565248t^4 - 1835008t^3 - 3670016t^2 - 4194304t - 2097152)$
18	$\frac{1}{3}(t^3+12t^2+36t+36)(t^9+252t^8+4644t^7+39636t^6+198288t^5+629856t^4+1294704t^3+1679616t^2 + 1259712t + 419904)$	$\frac{1}{2}(t^6 + 36t^5 + 324t^4 + 1404t^3 + 3240t^2 + 3888t + 1944)(t^12 - 504t^11 - 15336t^10 - 208872t^9 - 1700352t^8 - 9206784t^7 - 34836480t^6 - 94058496t^5 - 181398528t^4 - 245223936t^3 - 221709312t^2 - 120932352t - 30233088)$

Table 5: Equations for $\mathcal{X}_{1/2}(N)$

Bibliography

[1] Jeffrey Achter, *The distribution of class groups of function fields*, Journal of Pure and Applied Algebra **204** (2006), 316–333.

[2] ———, *The distribution of p-torsion subschemes of abelian varieties*, Journal of the Institute of Mathematics of Jussieu **15** (2016), 693–710.

[3] Jeffrey Achter and Rachel Pries, *Superspecial rank of supersingular abelian varieties and jacobians*, Journal de théorie des nombres de Bordeaux **27** (2015), no. 3, 605–624.

[4] Albert H. Beiler, *Recreations in the theory of numbers*, 2nd ed., Dover Publications Inc., New York, 1966.

[5] Alina Bucur, Chantal David, Brooke Feigon, and Matlide Lálin, *Statistics for traces of cyclic trigonal curves over finite fields*, International Mathematics Research Notices **2010** (2010), no. 5, 932–967.

[6] Bryden Cais, Jordan S. Ellenberg, and David Zureick-Brown, *Random dieudonné modules, random p-divisible groups, and random curves over finite fields*, Journal of the Institute of Mathematics of Jussieu **12** (2013), no. 3, 651–676.

[7] Antoine Chambert-Loir and Yuri Tschinkel, *Fonctions zêta des hauteurs des espaces fibrés*, Rational points on algebraic varieties, pp. 71–115.

[8] Richard Crew, *Etale p-covers in characteristic p*, Compositio Mathematica **52** (1984), no. 1, 31–45.

[9] Harold Davenport, *On a principle of lipschitz*, Journal of the London Mathematical Society **s1-26** (1951), 179–183.

[10] Pierre Deligne and Michael Rapoport, *Les schémas de modules de courbes elliptiques*, Modular functions of one variable II. Lecture notes in mathematics, 1973.

[11] Fred Diamond and Jerry Shurman, *A first course in modular forms*, Vol. 228, Springer-Verlag New York, New York, 2005.

[12] Arsen Elkin, *The rank of the Cartier operator on cyclic covers of the projective line*, Journal of Algebra **237** (2011), no. 1, 1–12.

[13] Jordan S. Ellenberg, Matthew Satriano, and David Zureick-Brown, *Heights on stacks and a generalized Batyrev–Manin–Malle conjecture* (2020).

[14] Carel Faber and Gerard van der Geer, *Complete subvarieties of moduli spaces and the Prym map*, JJournal für die reine und angewandte Mathematik (Crelles Journal) **2004** (2004), no. 573, 117–137.

[15] Darren Glass and Rachel Pries, *Hyperelliptic curves with prescribed p-torsion*, Manuscripta Mathematica **117** (2005), no. 3, 299–317.

[16] Ralph Greenberg, Karl Rubin, Alice Silverberg, and Michael Stoll, *On elliptic curves with an isogeny of degree 7*, American Journal of Mathematics **136** (2014), 77–109.

[17] Robert Harron and Andrew Snowden, *Counting elliptic curves with prescribed torsion*, Journal für die reine und angewandte Mathematik (Crelles Journal) **729** (2017), 151–170.

[18] Saito Hayato and Suda Tomohiko, *An explicit structure of the graded ring of modular forms of small level* (2011). Arxiv: 11108.3933.

[19] Nicholas M. Katz and Barry Mazur, *Arithmetic moduli of elliptic curves*, Princeton University Press, 1985.

[20] M. A Kenku, *On the number of $\mathbb{Q}$-isomorphism classes of elliptic curves in each $\mathbb{Q}$-isogeny class*, Journal of Number Theory **15** (1982), 199–202.

[21] Ja Kyung Koo, *On holomorphic differentials of some algebraic function fields of one variable over $\mathbb{C}$*, Bulletin of the Australian Mathematical Society **43** (1991), no. 3, 399–405.

[22] Daniel Sion Kubert, *Universal bounds on the torsion of elliptic curves*, Proceedings of the London Mathematical Society **3** (1976), no. 2, 193–237.

[23] Sylvain Maugeais, *Quelques résultats sur les déformations équivariantes des courbes stables*, Manuscripta Mathematica **120** (2006), no. 53.

[24] Barry Mazur and Dorian Goldfeld, *Rational isogenies of prime degree*, Inventiones mathematicae **44** (1978), 129–162.

[25] Ken McMurdy, *Equations for $\mathcal{X}_0(N)$*. Available at `https://phobos.ramapo.edu/~kmcmurdy/research/Models/index.html`.

[26] Shoichi Nakajima, *On generalized Hasse-Witt invariants of an algebraic curve*, 1983, pp. 69–88.

[27] Tadao Oda, *The first de Rham cohomology group and Dieudonné modules*, Annales scientifiques de l'École Normale Supérieure **2** (1969), 63–135.

[28] Maggie Pizzo, Carl Pomerance, and John Voight, Proceedings of the American Mathematical Society, Series B **7** (2020), 28–42.

[29] Carl Pomerance and Edward F. Schaefer, *Elliptic curves with Galois-stable cyclic subgroups of order 4*. Arxiv: 2004.14947.

[30] Rachel Pries and Hui June Zhu, *The p-rank stratification of Artin-Schreier curves*, Annales de l'Institut Fourier **62** (2012), no. 2, 707–726.

[31] Jeremy Rouse and David Zureick-Brown, *Elliptic curves over $\mathbb{Q}$ and 2-adic images of Galois*, Research in Number Theory **12** (2015).

[32] Joseph H. Silverman, *The Arithmetic of Elliptic Curves*, 2nd ed., Graduate Texts in Mathematics, vol. 106, Springer-Verlag New York, 2009.

[33] The Stacks Project Authors, *Stacks Project*, 2018.

[34] Henning Stichtenoth, *Über die Automorphismengruppe eines algebraischen Funktionenkörpers von Primzahlcharakteristik. Teil 1: Eine Abschätzung der Ordnung der Automorphismengruppe*, Archiv der Mathematik **24** (1973), 524–544.

[35] ———, *Über die Automorphismengruppe eines algebraischen Funktionenkörpers von Primzahlcharakteristik. Teil 2: Ein spezieller Typ von Funktionenkörpern*, Archiv der Mathematik **24** (1973), 615–631.

[36] ———, *Algebraic function fields and codes*, Springer-Verlag, 2009.

[37] Doré Subrao, *The p-rank of Artin-Schreier curves*, Manuscripta Mathematica **16** (1975), 169–194.

[38] John Tate, *p-divisible groups*, Proceedings of a conference on local fields, 1967.

[39] John Voight and David Zureick-Brown, *The canonical ring of a stacky curve* (2019), available at `Arxiv: 1501.04657v3`.

[40] Kęstutis Česnavičius, *A modular description of* $\mathscr{X}_0(N)$, Algebra Number Theory **11** (2017), 2001–2089.