DISSERTATION

MAXIMAL CURVES, ZETA FUNCTIONS, AND DIGITAL SIGNATURES

Submitted by

Beth Malmskog

Department of Mathematics

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Spring 2011

Doctoral Committe:

Advisor: Rachel Pries

Jeffrey Achter
Tim Penttila
Jacob Roberts

ABSTRACT

MAXIMAL CURVES, ZETA FUNCTIONS, AND DIGITAL SIGNATURES

Curves with as many points as possible over a finite field $\mathbb{F}_q$ under the Hasse-Weil bound are called maximal curves. Besides being interesting as extremal objects, maximal curves have applications in coding theory. A maximal curves may also have a great deal of symmetry, i.e. have an automorphism group which is large compared to the curve's genus. In Part 1, we study certain families of maximal curves and find a large subgroup of each curve's automorphism group. We also give an upper bound for the size of the automorphism group.

In Part 2, we study the zeta functions of graphs. The Ihara zeta function of a graph was defined by Ihara in the 1960s. It was modeled on other zeta functions in its form, an infinite product over primes, and has some analogous properties, for example convergence to a rational function. The knowledge of the zeta function of a regular graph is equivalent to knowledge of the eigenvalues of its adjacency matrix. We calculate the Ihara zeta function for an infinite family of irregular graphs and consider how the same technique could be applied to other irregular families. We also discuss ramified coverings of graphs and a joint result with Michelle Manes on the divisibility properties of zeta functions for graphs in ramified covers.

Part 3 is joint work with Jeremy Muskat. Gauss's curve, with equation $x^2t^2 + y^2t^2 + x^2y^2 - t^4 = 0$ defined over $\mathbb{F}_p$ was the subject of the last entry in Gauss' mathematical diary. For $p \equiv 3 \bmod 4$, we give a proof that the zeta function of C is

$$Z_C(u) = \frac{(1+pu^2)(1+u)^2}{(1-pu)(1-u)}.$$

Using this, we find the global zeta function for C.

The best algorithms for solving some lattice problems, like finding the shortest vector in an arbitrary lattice, are exponential in run-time. This makes lattice problems a potentially good basis for cryptographic protocols. Right now, lattices are especially important in information security because there are no known quantum computer algorithms that solve lattice problems any faster than traditional computing. The learning with errors problem (LWE) is provably as hard as certain lattice problems. Part 4 of the dissertation is a description of a digital signature scheme based on the learning with errors problem over polynomial rings. The search version of LWE is to find a hidden vector s, given access to many pairs of noisy inner products with random vectors $(a_i, b_i = a_i \cdot s + e_i)$. The context can be shifted to a polynomial ring over $\mathbb{Z}/q$, giving rise to the problem of learning with errors over a ring (R-LWE). In this joint work with Kristin Lauter, Michael Naehrig, and Vinod Vaikuntanathan, we devise a digital signature scheme based on R-LWE and outline a proof of security for certain parameter choices.

Contents

Part I

Automorphisms of a family of maximal curves

I began this work with my advisor Rachel Pries in 2007 when Guilietti and Korchmaros posted a paper on the Arxiv about a new family of maximal curves (GK curves). Besides proving maximality, Guilietti and Korchmaros determined the automorphism groups for their curves. I was reading their paper in preparation for my qualifying exam when Garcia, Guneri, and Stichtenoth posted a paper on a family of maximal curves that generalized the GK curves. We decided to work on finding the automorphism groups of the curves in the general family.

Chapter 1

Introduction

Let $n \geq 3$ be odd and let $q = p^h$ be a power of a prime. Let $m = (q^n + 1)/(q + 1)$. Define $\mathcal{C}_n$ to be the normalization of a fiber product over $\mathbb{P}^1$ of the covers of curves $\mathcal{H}_q \to \mathbb{P}^1_y$ and $\mathcal{X}_n \to \mathbb{P}^1_y$, where $\mathcal{H}_q$ and $\mathcal{X}_n$ have affine equations

$$\mathcal{H}_q \quad : \quad x^q + x - y^{q+1} = 0 \tag{1.1}$$

$$\mathcal{X}_n \quad : \quad y^{q^2} - y - z^m = 0. \tag{1.2}$$

The Hasse-Weil bound states that for a smooth connected projective curve $\mathcal{X}$ with genus g, defined over $\mathbb{F}_{q^2}$, the number of points on $\mathcal{X}$ defined over $\mathbb{F}_{q^2}$ is bounded above by $q^2 + 1 + 2gq$. A curve which attains this bound is called an $\mathbb{F}_{q^2}$-maximal curve. The curve $\mathcal{H}_q$ is known as the Hermitian curve and has been well studied [16]. It is a maximal over $\mathbb{F}_{q^2}$, and thus maximal over $\mathbb{F}_{q^{2n}}$ for $n \geq 3$ odd. It has genus $q(q-1)/2$, the highest genus which is attainable for an $\mathbb{F}_{q^2}$-maximal curve. The curve $\mathcal{X}_n$ is a member of a class that has been studied by Stichtenoth [45]. Abdon, Bezerra, and Quoos proved that the genus of $\mathcal{X}_n$ is $(q-1)(q^n - q)/2$, and that $\mathcal{X}_n$ is $\mathbb{F}_{q^{2n}}$-maximal [1].

For a given q, the curve $\mathcal{C}_3$ coincides with Giulietti and Korchmaros' maximal curve [12]. Giulietti and Korchmaros proved that $\mathcal{C}_3$ was maximal using the natural embedding theorem, a result from Korchmaros and Torres [21] that states that every

$\mathbb{F}_{q^2}$-maximal curve is isomorphic to a curve of degree $q+1$ embedded in a Hermitian variety of bounded dimension. Giulietti and Korchmaros also proved that the curve $\mathcal{C}_3$ is not covered by any Hermitian curve and determined the $\mathbb{F}_{q^2}$-automorphism group $\mathrm{Aut}_{\mathbb{F}_{q^2}}(\mathcal{C}_3)$ if $q \equiv 1 (\mathrm{mod}\ 3)$ and a normal subgroup of index 3 in $\mathrm{Aut}_{\mathbb{F}_{q^2}}(\mathcal{C}_3)$ if $q \equiv 2 (\mathrm{mod}\ 3)$. Giulietti and Korchmaros prove these facts by showing that elements of these groups give rise to automorphisms of the curve $\mathcal{C}_3$, then by bounding the size of $\mathrm{Aut}_{\mathbb{F}_{q^2}}(\mathcal{C}_3)$. When $q \equiv 1 (\mathrm{mod}\ 3)$, this completely determines the $\mathbb{F}_{q^2}$-automorphism group. This group is very large compared to the genus $g_{\mathcal{C}_3}$ of $\mathcal{C}_3$, i.e. $\mathrm{Aut}(\mathcal{C}_3) \geq 24 g_{\mathcal{C}_3}(g_{\mathcal{C}_3} - 1)$.

Garcia, Guneri, and Stichtenoth prove that $\mathcal{C}_n$ is $\mathbb{F}_{q^{2n}}$-maximal for $n \geq 3$ [11]. Recently, Duursma and Mak proved that $\mathcal{C}_n$ is not Galois covered by the Hermitian curve $\mathcal{H}_{q^{2n}}$ for q odd, and exhibited a Galois covering for q even [9]. For $n > 3$, the automorphism groups of the curves $\mathcal{X}_n$ and $\mathcal{C}_n$ do not appear in the literature. We prove the following:

Theorem 1. *The automorphism group $Aut(\mathcal{X}_n)$ fixes the point at infinity on $\mathcal{X}_n$ and is a semi-direct product of the form $(\mathbb{Z}/p)^{2h} \rtimes \mathbb{Z}/(q^n+1)(q-1)$.*

The curve $\mathcal{C}_n$ has a single point P_∞ at infinity. Let $I_{\mathcal{C}_n} \subseteq \mathrm{Aut}(\mathcal{C}_n)$ be the inertia group of $\mathcal{C}_n$ at P_∞.

Theorem 2. *The group $I_{\mathcal{C}_n}$ is a semi-direct product of the form $Q \rtimes \mathbb{Z}/(q^n+1)(q-1)$, where Q is a non-abelian group of order q^3 and exponent p.*

We describe the structure of $\Gamma = Q \rtimes \mathbb{Z}/(q^n+1)(q-1)$ more precisely in Chapter 3. These results are the focus of the first chapters of this part of the thesis. See the appendix, Chapter 6, for a background on algebraic curves, maximal curves, and the particular families that are the focus of this paper.

Chapter 2

Geometry of $\mathcal{C}_n$ and $\mathcal{X}_n$

Though the curve $\mathcal{C}_3$ was initially presented [12] as the intersection in $\mathbb{P}^3$ of two hypersurfaces, it is useful to view the generalized curve $\mathcal{C}_n$ as a fiber product, as illustrated in the following diagram:

$$\begin{array}{ccccc} \mathcal{C}_n = & \mathcal{X}_n \tilde{\times}_{\mathbb{P}^1} \mathcal{H}_q & \rightarrow & \mathcal{X}_n \\ & \downarrow & & \downarrow \\ & \mathcal{H}_q & \longrightarrow & \mathbb{P}^1_y. \end{array}$$

This construction lets us see that $\mathcal{C}_n$ has exactly one point at infinity.

Proposition 1. *The curve $\mathcal{C}_n$ has a single point at infinity.*

Proof. The fibers of $\mathcal{H}_q \rightarrow \mathbb{P}^1_y$ and $\mathcal{X}_n \rightarrow \mathbb{P}^1_y$ over infinity each consist of a unique point, meaning that the point at infinity is fully ramified in each of these covers. Since the degrees of the covers are relatively prime, we see that $q(q^n+1)/(q+1)$ divides the ramification index of any point on $\mathcal{C}_n$ over infinity in the cover $\mathcal{C}_n \rightarrow \mathbb{P}^1_y$. Since the degeree of $\mathcal{C}_n \rightarrow \mathbb{P}^1_y$ is $q(q^n+1)/(q+1)$, this implies that there can only be a single point at ∞ on $\mathcal{C}_n$. □

Viewing the curve $\mathcal{C}_n$ as a fiber product allows us to use the Riemann-Hurwitz formula to show that the genus of $\mathcal{C}_n$ is $(q-1)(q^{n+1}+q^n-q^2)/2$ [11]. Garcia, Guneri, and Stichtenoth also use the fiber product construction to study the sizes of the fibers above $\mathbb{F}_{q^{2n}}$-points in $\mathcal{H}_q$ and $\mathcal{X}_n$, and thereby prove that $\mathcal{C}_n$ is $\mathbb{F}_{q^{2n}}$ maximal.

Remark 1. *The projective curve in* $\mathbb{P}^3$ *given by the homogenization of the equations* $x^q + x - y^{q+1} = 0$ *and* $y^{q^2} - y - z^{(q^n+1)/(q+1)} = 0$ *is smooth only when* $n = 3$. *For* $n \geq 5$, *the curve has a cusp type singularity at* ∞.

The following summarizes the genera and numbers of $\mathbb{F}_{q^{2n}}$-points for the curves $\mathcal{X}_n$ and $\mathcal{C}_n$:

$$\begin{aligned} g_{\mathcal{X}_n} &= (q-1)(q^n-q)/2, & \#\mathcal{X}_n(\mathbb{F}_{q^{2n}}) &= q^{2n+1} - q^{n+2} + q^{n+1} + 1, \\ g_{\mathcal{C}_n} &= (q-1)(q^{n+1}+q^n-q^2)/2, & \#\mathcal{C}_n(\mathbb{F}_{q^{2n}}) &= q^{2n+2} - q^{n+3} + q^{n+2} + 1. \end{aligned}$$

Chapter 3

The subgroups Q and G of $\mathbf{Aut}(\mathcal{C}_n)$

Let $a, b \in \mathbb{F}_{q^2}$ be such that $a^q + a = b^{q+1}$. Define

$$Q_{a,b} := \begin{pmatrix} 1 & b^q & a \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix}.$$

Let $Q = \{Q_{a,b} : a, b \in \mathbb{F}_{q^2}, a^q + a = b^{q+1}\}$. Note that with the operation of matrix multiplication, Q is a subgroup of the special unitary group $\mathrm{SU}(3, q^2)$. Notice that we have

$$Q_{a,b}Q_{c,d} = \begin{pmatrix} 1 & (b+d)^q & a+c+b^q d \\ 0 & 1 & b+d \\ 0 & 0 & 1 \end{pmatrix} = Q_{a+c+b^q d, b+d}.$$

This implies that Q is not abelian, since $Q_{a,b}Q_{c,d} = Q_{c,d}Q_{a,b}$ means that $b^q d = d^q b$, which is not true for arbitrary $b, d \in \mathbb{F}_{q^2}$.

Since there is a bijection between Q and the $\mathbb{F}_{q^2}$-rational affine points of $\mathcal{H}_q$, we see that $|Q| = q^3$. It is known that Q has exponent p if $p \neq 2$, exponent 4 if $p = 2$. The center of Q is $Q_0 := \{Q_{a,0}\} \subset Q$.

Lemma 1. *The subgroup Q_0 is isomorphic to $(\mathbb{Z}/p)^h$.*

Proof. Since Q has exponent p, so does Q_0. Also, $|Q_0| = q = p^h$, since $a^q + a = \mathrm{Tr}(a) = 0$ has q solutions $a \in \mathbb{F}_{q^2}$. Then we check that Q_0 is abelian:

$$Q_{\alpha,0}Q_{\beta,0} = Q_{\alpha+\beta,0} = Q_{\beta,0}Q_{\alpha,0}.$$

□

Since $Q_0 \lhd Q$, the subgroup Q/Q_0 is well defined and has order q^2.

Lemma 2. *The quotient group Q/Q_0 is isomorphic to $(\mathbb{Z}/p)^{2h}$.*

Proof. Since Q has exponent p, the factor group Q/Q_0 also has exponent p. Then Q/Q_0 is abelian since for $Q_{a,b}, Q_{c,d} \in Q$, the commutator $Q_{a,b}^{-1}Q_{b,c}^{-1}Q_{a,b}Q_{c,d}$ is in Q_0:

$$\begin{aligned} Q_{a,b}^{-1}Q_{b,c}^{-1}Q_{a,b}Q_{c,d} &= Q_{a^q+c^q+b^qd,-b-d}Q_{a+c+b^qd,b+d} \\ &= Q_{a^q+a+c^q+c+2b^qd-(b+d)^{q+1},0} \in Q_0. \end{aligned}$$

□

We now consider how $Q_{a,b}$ acts on the curve $\mathcal{C}_n$. Let (x, y, z) denote an affine point of $\mathcal{C}_n$. Define

$$Q_{a,b}: \qquad x \mapsto x + b^q y + a, \qquad y \mapsto y + b, \qquad z \mapsto z.$$

Proposition 2. *The group Q is contained in $Aut(\mathcal{C}_n)$ and the quotient curve $\mathcal{C}_n/Q$ is a projective line. Further, $\mathcal{C}_n/Q_0 = \mathcal{X}_n$.*

Proof. First, note that Q preserves $\mathcal{H}_q$. Let (x, y) be an affine point of $\mathcal{H}_q$. Then

$$\begin{aligned} Q_{a,b}(x)^q + Q_{a,b}(x) - Q_{a,b}(y)^{q+1} &= x^q + x + b^{q^2}y^q + b^qy + a^q + a - (y+b)^{q+1} \\ &= y^{q+1} + by^q + b^qy + b^{q+1} - (y+b)^{q+1} \\ &= 0. \end{aligned}$$

Next, we check that Q preserves $\mathcal{X}_n$. Let (y, z) be an affine point of $\mathcal{X}_n$. Then

$$\begin{aligned} Q_{a,b}(y)^{q^2} - Q_{a,b}(y) - Q_{a,b}(z)^m &= y^{q^2} - y + b^{q^2} - b - z^m \\ &= y^{q^2} - y - z^m \\ &= 0. \end{aligned}$$

So Q preserves $\mathcal{X}_n$, and so preserves the fiber product $\mathcal{C}_n$.

The quotient curve of $\mathcal{C}_n$ by Q_0 is $\mathcal{X}_n$ because $\mathbb{K}(\mathcal{X}_n)$ is fixed by Q_0 and $|\mathbb{K}(\mathcal{C}_n) : \mathbb{K}(\mathcal{X}_n)| = q = |Q_0|$. A similar argument shows that $\mathbb{K}(\mathcal{C}_n/Q) \cong \mathbb{K}(z)$, and so $\mathcal{C}_n/Q$ is a projective line, denoted $\mathbb{P}^1_z$.

□

Let $\zeta \in \mu_{(q^n+1)(q-1)}$ be a $(q^n+1)(q-1)$-st root of unity. Define g_ζ by

$$g_\zeta : \quad x \mapsto \zeta^{q^n+1}x \qquad y \mapsto \zeta^m y, \qquad z \mapsto \zeta z.$$

Note that ζ^{q^n+1} is a $(q-1)$-st root of unity, so an element of $\mathbb{F}_q$. Therefore $(\zeta^{q^n+1})^q = \zeta^{q^n+1}$. Let $G = \{g_\zeta : \zeta \in \mu_{(q^n+1)(q-1)}\}$, a group of order $(q^n+1)(q-1)$. Define M to be the subgroup of G of order $m = (q^n+1)/(q+1)$, i.e. $M = \{g_\zeta : \zeta^m = 1\}$. Let N be the subgroup of G of order q^n+1.

Proposition 3. *The group G is contained in Aut($\mathcal{C}_n$). The quotient curves $\mathcal{C}_n/N$ and $\mathcal{C}_n/G$ are projective lines and $\mathcal{C}_n/M = \mathcal{H}_q$.*

Proof. First, we check that G preserves $\mathcal{H}_q$. Let (x, y) be an affine point of $\mathcal{H}_q$. Then

$$\begin{aligned} g_\zeta(x)^q + g_\zeta(x) - g_\zeta(y)^{q+1} &= \zeta^{q(q^n+1)}x^q + \zeta^{q^n+1}x - \zeta^{q^n+1}y^{q+1} \\ &= \zeta^{q^n+1}(x^q + x - y^{q+1}) \\ &= 0 \end{aligned}$$

Now we check that G preserves $\mathcal{X}_n$. Let (y, z) be an affine point of $\mathcal{X}_n$. Then

$$\begin{aligned} g_\zeta(y)^{q^2} - g_\zeta(y) - g_\zeta(z)^m &= \zeta^{mq^2}y^{q^2} - \zeta^m y - \zeta^m z^m \\ &= \zeta^m y(\zeta^{m(q^2-1)}y^{q^2-1} - 1) - \zeta^m z^m \\ &= \zeta^m y(y^{q^2-1} - 1) - \zeta^m z^m \\ &= \zeta^m(y^{q^2} - y - z^m) \\ &= 0 \end{aligned}$$

So G preserves the fiber product $\mathcal{C}_n$. Therefore $G \subset \mathrm{Aut}(\mathcal{C}_n)$.

As before we see that $\mathbb{K}(\mathcal{C}_n/M) \cong \mathbb{K}(\mathcal{H}_q)$ and that $\mathbb{K}(\mathcal{C}_n/N) \cong \mathbb{K}(x)$. So $\mathcal{C}_n/N$ is a projective line denoted $\mathbb{P}^1_x$. We have that $\mathbb{K}(\mathcal{C}_n/G) \cong \mathbb{K}(u)$, where $u = x^{q-1}$, and so $\mathcal{C}_n/G$ is a projective line denoted $\mathbb{P}^1_u$. □

Proposition 4. *The group generated by G and Q in Aut($\mathcal{C}_n$) is a semi-direct product of the form $Q \rtimes_\phi G$. Further, the homomorphism $\phi : G \to Aut(Q)$ is given by $g_\zeta \mapsto \psi_\zeta$, where $\psi_\zeta : Q_{a,b} \mapsto g_\zeta Q_{a,b} g_\zeta^{-1}$.*

Proof. Note first that $|Q|$ and $|G|$ are relatively prime, so $Q \cap G = \{\mathrm{id}\}$. To have a semidirect product, the only other requirement is that G normalizes Q [8, Section 5.5].

Let $g_\zeta \in G$ and $Q_{a,b} \in Q$ as above. Writing ψ_ζ more explicitly, we have

$$\begin{aligned} \psi_\zeta(Q_{a,b}) : \quad x &\mapsto x + b^q \zeta^{q^n+1-m} y + a\zeta^{q^n+1} \\ y &\mapsto y + \zeta^m b \\ z &\mapsto z \end{aligned}$$

Since $\zeta^{q^n+1-m} = (\zeta^m)^q$, that means that $\psi_\zeta(Q_{a,b}) = Q_{\zeta^{q^n+1}a, \zeta^m b}$. This is an element of Q because

$$(\zeta^m b)^{q+1} = \zeta^{q^n+1} b^{q+1} = \zeta^{q^n+1}(a + a^q) = \zeta^{q^n+1} a + (\zeta^{q^n+1} a)^q.$$

Therefore G normalizes Q, and the group $\langle G, Q \rangle$ is a semidirect product.

□

Let $\Gamma = Q \rtimes_\phi G$. This semidirect product is not a direct product, as the following lemma makes specifically clear.

Lemma 3. *Subgroups isomorphic to $Q \times M$ and $Q_0 \times N$ are contained in Γ.*

Proof. Notice that

$$g_\zeta Q_{a,b} g_\zeta^{-1} = Q_{\zeta^{q^n+1}a, \zeta^m b}.$$

For $b \neq 0$, we have $g_\zeta Q_{a,b} g_\zeta^{-1} = Q_{a,b}$ if and only if $g_\zeta \in M$. Therefore M commutes with every element of Q. If $b = 0$, then $g_\zeta Q_{a,0} g_\zeta^{-1} = Q_{a,0}$ if and only if $g_\zeta \in N$. Therefore N commutes with every element of Q_0. □

The above propositions prove the following:

Proposition 5. *The group Aut($\mathcal{C}_n$) contains a subgroup isomorphic to $Q \rtimes_\phi G$.*

This gives a lower bound on the size of $\mathrm{Aut}(\mathcal{C}_n)$:

$$|\mathrm{Aut}(\mathcal{C}_n)| \geq |Q \rtimes_\phi G| = q^3(q^n+1)(q-1).$$

The genus of $\mathcal{C}_n$ is $(q-1)(q^{n+1}+q^n-q^2)/2$, meaning that $|\mathrm{Aut}(\mathcal{C}_n)|$ grows more than linearly in $g(\mathcal{C}_n)$, surpassing the Hurwitz bound for large q.

We can draw the following diagram to summarize the coverings we have described:

$$\begin{array}{cccccc}
 & & q & & q^2 & \\
 & \mathcal{C}_n & \to & \mathcal{X}_n & \to & \mathbb{P}^1_z \\
\frac{q^n+1}{q+1} = m & \downarrow & & \downarrow & & \downarrow \\
 & \mathcal{H}_q & \to & \mathbb{P}^1_y & \to & \mathbb{P}^1_t \\
q+1 & \downarrow & & \downarrow & & \\
 & \mathbb{P}^1_x & \to & \mathbb{P}^1_w & & \\
q-1 & \downarrow & & & & \\
 & \mathbb{P}^1_u & & & &
\end{array}$$

The numbers next to the arrows are the degrees of the coverings. The projective line $\mathbb{P}^1_w$ denotes the curve $\mathcal{C}_n/(Q_0 \times N)$, where $\mathbb{K}(\mathcal{C}_n/(Q_0 \times N)) \cong \mathbb{K}(w)$ with $w = y^{q+1}$. The projective line $\mathbb{P}^1_t$ denotes the curve $\mathcal{C}_n/(Q \times M)$, where $\mathbb{K}(\mathcal{C}_n/(Q \times M)) \cong \mathbb{K}(t)$ with $t = z^m$.

Chapter 4

Ramification

Ramification theory provides a major tool in understanding the possibilities for the automorphism group of $\mathcal{C}_n$.

4.1 Ramification groups

Definition 1. *Let $\mathcal{Y}' \to \mathcal{Y}$ be a Galois covering of curves with Galois group G. Let P be a point on $\mathcal{Y}$ and P' be a point on $\mathcal{Y}'$ in the fiber over P. For $i \geq -1$ define the i-th ramification group of P' by*

$$G_i(P'|P) := \{\sigma \in G : v_{P'}(\sigma(z) - z) \geq i + 1 \ \ \forall \ \ z \in \mathcal{O}_{P'}\}.$$

For a reference about ramification groups, see Sections 2 and 3 of Chapter 4 of Serre [38]. Some facts about ramification groups in positive characteristic are:

- $G \geq G_{-1} \trianglerighteq G_0 \trianglerighteq ... \trianglerighteq G_i \trianglerighteq G_{i+1} \trianglerighteq ...$ and $G_m = \{\mathrm{id}\}$ for m sufficiently large.
- $|G_0| = e(P'|P)$ is the ramification index of P' over P.
- The order of G_1 is a power of p, G_0/G_1 is cyclic of order relatively prime to p, and G_i/G_{i+1} is elementary abelian of exponent p for $i \geq 1$.

The sequence of indices i which are above assigned to the subgroups is called the lower numbering of the ramification groups. An integer i for which $G_i \neq G_{i+1}$ is called a lower jump.

Lower numbering can be said to behave well with respect to subgroups of G (see [38, Chapter 4]).

Proposition 6. *Let $H \leq G$ with K the fixed field of H in $\mathbb{K}'$. Let H_i denote the ith ramification group for the extension K'/K. Then $H_i = G_i \cap H$.*

The lower jumps of H are a subsequence of the lower jumps of G.

The ramification groups can also be numbered by another system, known as the upper numbering. To find the upper numbering, we first need to extend the domain of our index set for the ramification groups. For $u \geq -1$ a real number, let G_u be the ramification group G_i, where $i = \lceil u \rceil$, i.e. i is the smallest integer $\geq u$. We then define the functions ϕ and ψ by

$$\phi(u) := \int_0^u \frac{dt}{(G_0 : G_t)},$$

$$\psi(u) := \phi^{-1}(u).$$

We now define the upper numbering of the ramification groups:

$$G^v := G_{\psi(v)},$$

or, equivalently

$$G^{\phi(u)} := G_u.$$

From this definition, it can be derived that

$$\psi(v) = \int_0^v (G^0 : G^w) dw.$$

An index j for which $G^j \neq G^k$ for all $k > j$ is called an upper jump. Upper numbering behaves well with respect to quotient groups of G, in the following sense.

Proposition 7. *H is a normal subgroup of G, then*

$$(G/H)^v = (G^v H)/H$$

for all v.

Serre proves this proposition [38, Chapter 4, Section 3] using Herbrand's Theorem.

4.2 Filtrations at infinity

In determining the ramification groups, it is enough to consider the action of $\sigma \in G$ on a uniformizer of the curve at the given point [45].

Proposition 8. *There is one break in the ramification filtration of $\mathcal{H}_q \to \mathbb{P}^1_y$ and it occurs at index $q+1$ in the lower numbering.*

Proof. Recall that this is a degree q Artin-Schreier cover. The associated Galois group is $Q_0 \cong (\mathbb{Z}/p)^h$ where $p^h = q$. The elements of Q_0 are $Q_{a,0}$, where $Q_{a,0}(x) = x + a$ for $a \in \mathbb{F}_{q^2}$ with $a^q + a = \mathrm{Tr}_{\mathbb{F}_{q^2}/\mathbb{F}_q}(a) = 0$.

Let ∞_y be the point at infinity of $\mathbb{P}^1_y$. Let ∞' be the point at infinity of $\mathcal{H}_q$. Stichtenoth analyzes generalized Artin-Schreier covers corresponding to function field extensions $K(x)/K$ of the form $A(x) = u$ for some $u \in K$ with certain properties and $A(x)$ a separable, additive polynomial with all its roots in the base field. He uses the valuation at infinity in K of $u - A(r)$, where r is any element of K, to determine the ramification groups at infinity [45, Proposition 3.7.10]. Here, we have $K = \mathbb{K}(\mathbb{P}^1_y)$, $K = \mathbb{K}(\mathcal{H}_q)$, with $A(x) = x^q + x$ and $u = y^{q+1}$.

For every each P on $\mathbb{P}^1_y$, define an integer m_P by

$$m_P := \begin{cases} m & \text{if there is an element } r \in \mathbb{F}_{q^{2n}} \text{ such that} \\ & v_P(y^{q+1} - (r^q + r)) = -m < 0 \text{ and } m \not\equiv 0 \bmod p \\ \\ -1 & \text{if } v_P(y^{q+1} - (r^q + r)) \geq 0 \text{ for some } r \in \mathbb{F}_{q^{2n}} \end{cases}$$

We can choose r to be an element of trace 0 in $\mathbb{F}_{q^2}$. We then have

$$m_{\infty_y} = -v_{\infty_y}(y^{q+1} - r^q - r) = -v_{\infty_y}(y^{q+1}) = -(q+1)(v_{\infty_y}(y)) = q+1.$$

Let τ be a uniformizer at ∞'. Following [45, Proposition 3.7.8] we see that since m_P is prime-to-p, we have $v_{\infty'}(\tau - Q_{a,0}(\tau)) = m_\infty + 1 = q+2$ for all $a \neq 0$ as above. Also, $v_{\infty'}(\tau - Q_{0,0}(\tau)) = v_{\infty'}(0) = \infty$ by definition. That means that

$$G_i := \begin{cases} Q_0 & \text{for } -1 \leq i \leq q+1 \\ \{id\} & \text{for } i \geq q+2 \end{cases},$$

so the lower jump for the extension $\mathbb{K}(\mathcal{H}_q)/\mathbb{K}(y)$ is $q+1$.

□

Proposition 9. *There is one break in the ramification filtration of $\mathcal{C}_n \to \mathcal{X}_n$ and it occurs at index q^n+1 in the lower numbering.*

Proof. The same ideas apply to the covering $\mathcal{C}_n \to \mathcal{X}_n$, with the modification that ∞' now represents the place at infinity in $\mathbb{K}(\mathcal{C}_n)$, and ∞ the place at infinity in $\mathbb{K}(\mathcal{X}_n)$. As before, let ∞_y represent the point at infinity on $\mathbb{P}^1_y$. Then we have

$$v_\infty(y) = e(\infty|\infty_y)(v_{\infty_y}(y)) = -\frac{q^n+1}{q+1}.$$

So, again choosing r of trace 0 in $\mathbb{F}_{q^2}$, we have

$$m_\infty = -v_\infty(y^{q+1} - r^q - r) = -v_\infty(y^{q+1}) = q^n + 1.$$

Therefore $|G_i| = q$ for $-1 \leq i \leq q^n+1$ and $|G_i| = 1$ for $q^n+2 \leq i$, giving a lower jump of q^n+1. □

Next we find the ramification filtration for the extension $\mathbb{K}(\mathcal{X}_n)/\mathbb{K}(\mathbb{P}^1_z)$, a generalized Artin-Schreier extension of order q^2. The Galois group Q/Q_0 is isomorphic to $(\mathbb{Z}/p)^{2h} \cong \mathbb{F}_{q^2}$ (as an additive group). Let $\sigma_b \in Q/Q_0$ be defined by $\sigma_b(y,z) = (y+b, z)$ for (y,z) a point of $\mathcal{X}_n$ and $b \in \mathbb{F}_{q^2}$.

Proposition 10. *There is one break in the ramification filtration of $\mathcal{X}_n \to \mathbb{P}^1_z$ and it occurs at index m in the lower numbering.*

Proof. Let ∞' be the place at infinity in $\mathbb{K}(\mathcal{X}_n)$ and let ∞ be the point at infinity on $\mathbb{P}^1_z$. Here, we have

$$m_\infty = -v_\infty(z^m) = -m(v_\infty(z)) = m.$$

Since this is prime-to-p, that means if τ is a uniformizer at ∞, we have

$$v_{\infty'}(\tau - \sigma_b(\tau)) = \frac{q^n+1}{q+1} + 1$$

for all $b \neq 0$. Therefore $|G_i| = q^2$ for $-1 \leq i \leq m = (q^n+1)/(q+1)$ and $|G_i| = 1$ for all larger i, giving a lower jump of $m = (q^n+1)/(q+1)$. □

Now that we've determined the ramification filtrations and lower jumps for each of the extensions above, we determine the upper jumps for the degree q^2 extension from Proposition 10. This is so that we can determine the jumps for the larger extension $\mathbb{K}(\mathcal{C}_n)/\mathbb{K}(\mathbb{P}^1_z)$.

Proposition 11. *There are two jumps in the ramification filtration of $\mathcal{C}_n \to \mathbb{P}^1_z$, and they occur at indices $m = (q^n+1)/(q+1)$ and q^n+1 in the lower numbering.*

Proof. Begin with $\mathbb{K}(\mathcal{X}_n)/\mathbb{K}(\mathbb{P}^1_z)$. Since the lower jump in extension is m, the upper jump is given by

$$\begin{aligned} \phi(m) &= \int_0^m \frac{dt}{G_0 : G_t} \\ &= \sum_{i=1}^{m} 1 \\ &= m. \end{aligned}$$

This illustrates the fact that, if there is only one jump in a p-group extension, its upper jump is the same as its lower jump.

Now, we use proposition 6 to say that the q^n+1, the lower jump for $\mathbb{K}(\mathcal{C}_n)/\mathbb{K}(\mathcal{X}_n)$, is also a lower jump for $\mathbb{K}(\mathcal{C}_n)/\mathbb{K}(\mathbb{P}^1_z)$. Proposition 7 requires that m, the upper jump

for $\mathbb{K}(\mathcal{X}_n)/\mathbb{K}(\mathbb{P}^1_z)$, is also an upper jump for $\mathbb{K}(\mathcal{C}_n)/\mathbb{K}(\mathbb{P}^1_z)$. The question now becomes whether or not these represent the same jump in the filtration, and if not, which one is larger.

We know that the filtration has at most two jumps, because more jumps in the filtration they would require additional jumps in the filtrations for $\mathbb{K}(\mathcal{C}_n)/\mathbb{K}(\mathcal{X}_n)$ or $\mathbb{K}(\mathcal{X}_n)/\mathbb{K}(\mathbb{P}^1_z)$–if $G_i \neq G_{i+1}$, then either $G_i \cap H \neq G_{i+1} \cap H$, which creates a jump in the filtration for $\mathbb{K}(\mathcal{C}_n)/\mathbb{K}(\mathcal{X}_n)$, or if not, then $HG_i/H \neq HG_{i+1}/H$, which implies a jump i in the filtration for $\mathbb{K}(\mathcal{X}_n)/\mathbb{K}(\mathbb{P}^1_z)$. There must be in fact two jumps, because if there is only one jump in the extension, it will have the same upper and lower numbering.

Say that the jump with lower numbering of $q^n + 1$ represents the smaller jump. Then, its upper numbering would be found by

$$\begin{aligned}\phi(q^n+1) &= \int_0^{q^n+1} \frac{dt}{G_0 : G_t} \\ &= \Sigma_{i=1}^{q^n+1} \frac{1}{G_0 : G - t} = q^n + 1\end{aligned}$$

This contradicts the assumption that it is the smaller jump, since $q^n + 1 > m$. Therefore the jump with upper numbering $(q^n + 1)(q + 1)$ is the smaller jump, and therefore its lower numbering is also m.

□

Chapter 5

Automorphisms of $\mathcal{X}_n$ and $\mathcal{C}_n$

5.1 Automorphisms of $\mathcal{X}_n$

Let $I_{\mathcal{X}_n}$ be the inertia group at the point at infinity of $\mathcal{X}_n$.

Theorem 3. *The automorphism group of $\mathcal{X}_n$ is Aut($\mathcal{X}_n$) = $Q/Q_0 \rtimes_\phi G$.*

Proof. The curve $\mathcal{X}_n$ is defined by the equation $A(y) = B(z)$, where

$$A(y) = y^{q^2} - y$$

and

$$B(z) = z^m.$$

Notice that $A(y)$ has the property that $A(y + a) = A(y) + A(a)$. Theorem 12.11 from [16] states that all automorphisms of this type of curve fix the point at infinity, so in fact the inertia group $I_{\mathcal{X}_n}$ is the entire automorphism group of $\mathcal{X}_n$.

Let G_i for $i \geq 0$ be the i-th ramification group of $I_{\mathcal{X}_n}$. The tame (prime-to-p) automorphisms in the inertia group are described $H < G_0$, where $H \cong G_0/G_1$. Notice that since the inertia group admits such a ramification filtration, there must be a unique Sylow-p subgroup of the inertia group. In fact, an inertia group G_0 must be the semi-direct product of a cyclic subgroup of order prime-to-p with a normal

(and hence unique) Sylow-p subgroup G_1. Applying [16] Theorem 12.7 (i) and (iii) to $\mathcal{X}_n$, we know that $|H|$ divides $(q-1)(q^n+1)$ and that $|G_1| = q^2$. So $|I_{\mathcal{X}_n}|$ divides $q^2(q-1)(q^n+1)$.

We know that $(Q \rtimes_\phi G)/Q_0 \cong Q/Q_0 \rtimes_{\tilde{\phi}} G \leq I_{\mathcal{X}_n}$ since $\mathcal{X}_n = \mathcal{C}_n/Q_0$. Since $|Q/Q_0 \rtimes_\phi G| = q^2(q-1)(q^n+1)$, we have that $I_{\mathcal{X}_n} = \mathrm{Aut}(\mathcal{X}_n) = Q/Q_0 \rtimes_\phi G$.

□

5.2 Automorphisms of $\mathcal{C}_n$

Let $I_{\mathcal{C}_n}$ be the inertia group at P_∞ for $\mathcal{C}_n$ and let S be the Sylow-p subgroup of $I_{\mathcal{C}_n}$. Let $\mathcal{W} = \mathcal{C}_n/S$.

We refer to the results of Section 4 and make use of the following theorem of Serre [38]:

Theorem 4 (Serre)**.** *If $s \in G_i, t \in G_j$, and $i, j \geq 1$, then $sts^{-1}t^{-1} \in G_{i+j+1}$.*

Proposition 12. *The group Q_0 is in the center of S.*

Proof. Consider the extension $\mathbb{K}(\mathcal{C}_n)/\mathbb{K}(\mathcal{W})$. Let s be an element of $S - Q$ with maximal lower jump J in this extension. Since G_i/G_{i+1} is elementary abelian it must be that s and s^p have different lower jumps in this extension. The jump of s^p must be greater than that of s, so either $s^p = \mathrm{Id}$ or $s^p \in Q$. That means that $|s| = p$ or $|s| = p^2$.

Case 1: $|s| = p$. Assume that $J > q^n + 1$. Then s commutes with Q_0, since for $q \in Q_0$ Theorem 4 implies that $sqs^{-1}q^{-1} = \mathrm{Id}$. Therefore s descends to $\overline{s} \in I_{\mathcal{X}_n} = Q/Q_0$, a contradiction since $s \notin Q$. That means that $J \leq q^n + 1$. Then Q_0 is the last non-trivial ramification group and therefore in the center of S.

Case 2: $|s| = p^2$. Then $\langle s \rangle \cap Q = \langle s^p \rangle$, where $\langle s^p \rangle \cong \mathbb{Z}/p$. That means that in the ramification filtration of $\mathbb{K}(\mathcal{C}_n)/\mathbb{K}(\mathbb{P}^1_z)$, the lower jump of s^p must be $q^n + 1$ or

m. Since $Q \leq S$, Proposition 6 implies that the lower jumps of elements of Q will the same in the larger extension $\mathbb{K}(\mathcal{C}_n)/\mathbb{K}(\mathcal{W})$. Since the lower jump of s is less than or equal to that of s^{p^a}, we must have that $J \leq q^n + 1$. Therefore Theorem 4 implies that Q_0 is in the center of S.

□

Theorem 5. *The inertia group at infinity of $\mathcal{C}_n$ is $Q \rtimes_\phi G$.*

Proof. Assume that there exists $s \in S$ but $s \notin Q$. Then, by Proposition 12, we have that s commutes with Q_0 and s descends to an automorphism in $I_{\mathcal{X}_n}$, and $I_{\mathcal{X}_n} = S/Q_0$. Since $I_{\mathcal{X}_n} = S/Q_0 = Q/Q_0$, the third Isomorphism theorem implies that $S = Q$.

Now we consider the tame automorphisms in $I_{\mathcal{C}_n}$. Let T be the tame part of $I_{\mathcal{C}_n}$. Since T is cyclic group, all tame automorphisms commute with M and so descend to $\mathcal{H}_q$. Since the tame part of $I_{\mathcal{H}_q} = G/M$ [2], we have $T/M = G/M$ and the third isomorphism theorem implies that $T = G$. Since we already know that Q is the Sylow-p subgroup of $I_{\mathcal{C}_n}$, we have that $Q \rtimes_\phi G = I_{\mathcal{C}_n}$. □

5.3 Further restrictions on $\mathrm{Aut}(\mathcal{C}_n)$

The inertia group $I_{\mathcal{C}_n}$ gives a lower bound on the size of $\mathrm{Aut}(\mathcal{C}_n)$, namely

$$|\mathrm{Aut}(\mathcal{C}_n)| \geq |I_{\mathcal{C}_n}| = q^3(q^n + 1)(q - 1).$$

What else can be said to restrict the possibilities for $\mathrm{Aut}(\mathcal{C}_n)$? Theorem 11.127 from [16] gives one upper bound on the size. Let $\mathcal{Y}$ be a curve of genus $g_\mathcal{Y} \geq 2$ with automorphism group H. If $|H| \geq 8g_\mathcal{Y}^3$, then $\mathcal{Y}$ must be birationally equivalent to one of four specific curves:

1. the hyperelliptic curve with equation $y^2 + y + x^{2^k+1} = 0$ for $p = 2$ and $g = 2^{k-1}$;

2. the hyperelliptic curve with equation $y^2 = x^q + x$ for $p > 2$ and $g = (q - 1)/2$;

3. the Hermitian curve $\mathcal{H}_q$ with genus $g = (q^2 - q)/2$;

4. the DLS curve with equation $x^{q_0}(x^q + x) = y^q - y$ with $p = 2$, $q_0 = 2^r$, $q = 2^{2r+1}$, and $g = q_0(q - 1)$.

In our case, these can all be eliminated since $g_{\mathcal{C}_n} = (q-1)(q^{n+1} + q^n - q^2)/2$. This gives an upper bound on the size of the automorphism group:

$$q^3(q^n + 1)(q - 1) \leq |\mathrm{Aut}(\mathcal{C}_n)| < (q - 1)^3(q^{n+1} + q^n - q^2)^3.$$

Giulietti and Korchmaros determined the $\mathbb{F}_{q^2}$-automorphism group for the curve $\mathcal{C}_3$ if $q \equiv 1 (\mathrm{mod}\ 3)$ and found a normal subgroup of index 3 if $q \equiv 2 (\mathrm{mod}\ 3)$ [12].

Theorem 6 (Giulietti and Korchmaros)**.** *If $q \equiv 1$ (mod 3), then*

$$Aut_{\mathbb{F}_{q^6}}(\mathcal{C}_n) \cong SU(3, q^2) \times \mathbb{Z}/(q^3 + 1)/(q + 1).$$

If $q \equiv 2$ (mod 3), then there exists $G \lhd Aut_{\mathbb{F}_{q^2}}(\mathcal{C}_3)$ such that $|Aut_{\mathbb{F}_{q^6}}(\mathcal{C}_3) : G| = 3$ and

$$G \cong SU(3, q^2) \times \mathbb{Z}/(q^3 + 1)/(3(q + 1)).$$

Giulietti and Korchmaros prove this by showing that the given groups preserve $\mathcal{C}_3$ then using geometry to bound the size of $\mathrm{Aut}_{\mathbb{F}_{q^6}}(\mathcal{C}_3)$, which in the first case completely determines the automorphism group. We make similar progress towards the $\mathbb{F}_{q^{2n}}$-automorphism group of $\mathcal{C}_n$. Notice that the automorphisms in $\Gamma = Q \rtimes_\phi G$ are defined over $\mathbb{F}_{q^{2n}}$. Therefore Γ acts on the $\mathbb{F}_{q^{2n}}$-points of $\mathcal{C}_n$. Let H be a group acting on a set A. For any $a \in A$, the orbit-stabilizer theorem states that $|H| = |\mathrm{Stab}_H(a)||a^H|$. Since $|P_\infty^\Gamma| \leq \#\mathcal{C}_n(\mathbb{F}_{q^{2n}}) = q^{2n+2} - q^{n+3} + q^{n+2} + 1$, we get the upper bound

$$|\mathrm{Aut}_{\mathbb{F}_{q^{2n}}}(\mathcal{C}_n)| \leq |I_{\mathcal{C}_n}| \cdot \#\mathcal{C}_n(\mathbb{F}_{q^{2n}}) = q^3(q^n + 1)(q - 1)(q^{2n+2} - q^{n+3} + q^{n+2} + 1).$$

Though the leading terms of the two upper bounds are the same, comparing the next term of the two bounds shows that the second is a slight improvement over the first. We can more information by considering the possible orbits of P_∞, the point at infinity on $\mathcal{C}_n$.

Lemma 4. *Two affine $\mathbb{F}_{q^{2n}}$-points (x_1, y_1, z_1) and (x_2, y_2, z_2) are in the same orbit under $\Gamma \leq Aut_{\mathbb{F}_{q^{2n}}}(\mathcal{C}_n)$ if and only if $z_2 = \zeta z_2$ for ζ a $q^n + 1)(q-1)$st root of unity. The number of orbits of the $\mathbb{F}_{q^{2n}}$-points of $\mathcal{C}_n$ is $(q^{n-1} - 1)/(q-1) + 2$:*

$$O_\infty = \{P_\infty\}, \qquad O_{\mathbb{F}_{q^2}} = \{(x, y, 0) \text{ affine on } \mathcal{C}_n\}$$

$$O_i = \{(x, y, z) \text{ affine on } \mathcal{C}_n \text{ with } z \neq 0\}, 1 \leq i \leq (q^{n-1} - 1)/(q-1)$$

Proof. We have seen that Γ fixes P_∞. Next, consider an affine point $(x, y, 0)$ on $\mathcal{C}_n$. Notice that $y^{q^2} - y = 0$ if and only if $y \in \mathbb{F}_{q^2}$, and since $x^q + x = y^{q+1}$, it must be that $y \in \mathbb{F}_{q^2}$ if and only if $x \in \mathbb{F}_{q^2}$. If ζ is a $(q^n+1)(q-1)$-st root of unity then $\zeta^m \mathbb{F}_{q^2}$ and $\zeta^{q^n+1} \in \mathbb{F}_q$. Therefore $g_\zeta \in G$ fixes $O_{\mathbb{F}_{q^2}}$ set-wise. The group Q also fixes $O_{\mathbb{F}_{q^2}}$ set-wise because Q is defined over $\mathbb{F}_{q^2}$. We see that Q acts transitively on O_2 as follows. Let $P_1 = (x_1, y_1, 0)$ and $P_2 = (x_2, y_2, 0)$ be any two points in $O_{\mathbb{F}_{q^2}}$. Let $y_2 - y_1 = b \in \mathbb{F}_{q^2}$. Choose any a such that $a^q + a = b^{q+1}$. Then $Q_{a,b}(P_1) = (x_1 + b^q y_1 + a, y_2)$. Then let $x_1 + b^q y_1 + a = x'$. There are q solutions to the equation $x^q + x = y_2^{q+1}$, and both x' and x_2 are solutions. The group Q_0 fixes y and z. Since there are q elements of Q_0 there must be some element $Q_{\alpha,0} \in Q_0$ which sends x' to x_2. Therefore $Q_{\alpha,0} Q_{a,b}(P_1) = P_2$.

Consider the affine $\mathbb{F}_{q^{2n}}$-points of $\mathcal{C}_n$ with $z \neq 0$. Let $P_1 = (x_1, y_1, z_1)$ and $P_2 = (z_1, y_1, z_1)$ be two such points. Then P_1 and P_2 are in the same orbit if and only if $z_1 = \zeta z_2$ for ζ a $(q^n + 1)(q-1)$-st root of 1. This is true since Q fixes z and there exists $g_\zeta \in G$ which multiplies z by ζ. Two points $S_1 = (x_1, y_1, z)$ and $S_2 = (x_2, y_2, z)$ with the same z value are in the same orbit if and only if there is an automorphism $Q_{a,b}$ such that $Q_{a,b}(S_1) = S_2$. This gives a total of $q^3(q^n + 1)(q-1)$ points in each orbit. Since there are $q^{2n+2} - q^{n+3} + q^{n+2} - q^3$ such points, there must be $(q^{2n+2} - q^{n+3} + q^{n+2} - q^3)/(q^3(q^n+1)(q-1)) = (q^{n-1} - 1)/(q-1)$ different such orbits. From the description of the orbits it is clear that Γ acts transitively on these points.

□

Proposition 13. *The order of the $\mathbb{F}_{q^{2n}}$-automorphism group of $\mathcal{C}_n$ is given by*

$$|Aut_{\mathbb{F}_{q^{2n}}}(\mathcal{C}_n)| = |\Gamma|(1 + \alpha q^3 + \beta q^3(q^n + 1)(q-1)),$$

where $\alpha \in \{0,1\}$ *and* $\beta \in \{0,1,...,(q^{n-1}-1)/(q-1)\}$.

Proof. The orbit of P_∞ under $\mathrm{Aut}_{\mathbb{F}_{q^2}}(\mathcal{C}_n)$ must be some combination of the orbits of the $\mathbb{F}_{q^{2n}}$-points of $\mathcal{C}_n$ under Γ. Lemma 4 and the orbit-stabilizer theorem then imply the result. □

Since $1 + \alpha q^3 + \beta q^3(q^n+1)(q-1)$ is prime-to-p, we get the following corollary:

Corollary 1. *The group Q is a Sylow-p subgroup of $Aut_{\mathbb{F}_{q^{2n}}}(\mathcal{C}_n)$.*

Chapter 6

Appendix: Background on Algebraic Curves and Maximal Curves

6.1 Summary of notation

Here, let $q = p^h$ where p is a prime and $h \in \mathbb{N}$. Let $\mathbb{F}_q$ be the field with q elements. Let $\overline{\mathbb{F}}$ be an algebraic closure of $\mathbb{F}$. Let $\mathbb{P}^n$ denote n-dimensional projective space. If $\mathcal{X}$ is a curve in $\mathbb{P}^n$ defined over $\mathbb{F}_q$, for any field $\mathbb{F}_{q^i} \subset \overline{\mathbb{F}_q}$, let $\mathcal{X}(\mathbb{F}_{q^i})$ denote the set of points of $\mathcal{X}$ which are defined over $\mathbb{F}_{q^i}$. Then let $N_i = \#\mathcal{X}(\mathbb{F}_{q^i})$ be the cardinality of that set.

6.2 Algebraic curves

This brief background survey can be supplemented by Dummit and Foote [8], Stichtenoth [45], and Hirschfeld, Korchmaros, and Torres [16] [8].

An algebraic variety in projective space is the zero set of some set of homogeneous polynomials. All functions in the ideal generated by these polynomials will vanish at

these same values, so we can associate this variety with the entire homogeneous ideal generated by the defining polynomials. A point on a variety is a point of projective space which satisfies the defining polynomials.

An algebraic curve is a one-dimensional algebraic variety. In the simplest case, a curve is cut out of $\mathbb{P}^2$ (two dimensional projective space) as the zero set of a single homogeneous polynomial in 3 variables, i.e. it is the vanishing of the principal ideal generated by this polynomial. Curves of this sort are called plane curves. Curves can also be embedded in higher dimensional projective spaces but then require more than one defining polynomial (corresponding to non-principal ideals). It takes at least $n-1$ polynomials to carve a curve out of $\mathbb{P}^n$. Though curves are always considered to exist in projective space, it will sometimes be useful to work with the non-homogenous, affine versions of the defining equations.

The coefficients of the defining polynomials will lie in some field $\mathbb{L}$, and we could consider points whose coordinate values lie in any extension of $\mathbb{L}$. We will think of the curve $\mathcal{X}$ as existing over $\overline{\mathbb{L}}$, with all satisfactory tuples in $\overline{\mathbb{L}}$ being points of the curve. It will also be interesting to restrict our view to smaller fields. For a field $\mathbb{M}$ with $\mathbb{L} \subseteq \mathbb{M} \subseteq \overline{\mathbb{L}}$, we call the points of $\mathcal{X}$ with coordinates in $\mathbb{M}$ the $\mathbb{M}$-points of $\mathcal{X}$. If $\mathcal{X}$ is defined over a finite field $\mathbb{F}_q$, $\mathcal{X}$ will have an infinite number of points considered over $\overline{\mathbb{F}_q}$, but only a finite number of $\mathbb{F}_{q^n}$-points for any value of n. Counting these points is an important part of this paper.

6.3 Irreducibility and Smoothness

For the purposes of this paper, it is desirable that a curve be both irreducible and smooth. Vaguely, irreducible means that the curve is made up of a single component, and smooth means that the curve has a well defined unique tangent line at every point.

Let $\mathcal{X} = V(I)$, where I is a homogeneous ideal of $\mathbb{L}[x_0, x_1, ..., x_n]$. The variety $\mathcal{X}$ is said to be irreducible if and only if it can not be written as the union of two proper subvarieties.

$$\mathcal{X} \text{ is irreducible } \Leftrightarrow \mathcal{X} \neq V_1 \cup V_2 \text{ for all varieties } V_1, V_2 \neq \mathcal{X}.$$

If $\mathcal{X}$ is a plane curve, so $I = (F)$ for some single polynomial F, then $\mathcal{X}$ is irreducible if and only if F doesn't factor over $\mathbb{L}[X_0, X_1, X_2]$. The curve is called absolutely irreducible if it is irreducible when considered over $\overline{\mathbb{L}}$.

From here on, *curve* will be assumed to mean *absolutely irreducible curve* unless stated otherwise. This does not give away much ground because curves which are not irreducible are made up of a finite number of irreducible components.

A singularity of a curve is a point on the curve without a well defined tangent line. On a the graph of a real, affine plane curve, a singularity might look like a point where the line of the graph crosses itself or is pinched to a point. This picture doesn't translate directly to curves over finite fields, but the algebraic formulation of smoothness does.

To determine if a plane curve has a single, well defined tangent line at a point, we can calculate the formal partial derivatives of the defining homogeneous equation. The tangent line to the curve defined by $F(X, Y, Z) = 0$ at a point (A, B, C) has the homogeneous linear equation

$$\frac{\partial F}{\partial X}(A, B, C)(X - A) + \frac{\partial F}{\partial Y}(A, B, C)(Y - B) + \frac{\partial F}{\partial Z}(A, B, C)(Z - C) = 0.$$

It is not hard to show that the constant terms will cancel, resulting in the homogeneous linear equation

$$\frac{\partial F}{\partial X}(A, B, C)(X) + \frac{\partial F}{\partial Y}(A, B, C)(Y) + \frac{\partial F}{\partial Z}(A, B, C)(Z) = 0.$$

This will be a well defined line as long some partial derivative is non-zero when evaluated at (A, B, C).

For non-plane curves a similar but slightly more complicated procedure is required. For a curve in $\mathbb{P}^n$ defined by the $n-1$ equations $F_i(X_0, X_1, ..., X_n)$ $1 \leq i \leq n-1$, define the Jacobian matrix J of the curve to be the matrix of formal partial derivatives

$$J := \begin{pmatrix} \frac{\partial F_1}{\partial X_0} & \frac{\partial F_1}{\partial X_1} & \cdots & \frac{\partial F_1}{\partial X_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial F_{n-1}}{\partial X_0} & \frac{\partial F_{n-1}}{\partial X_1} & \cdots & \frac{\partial F_{n-1}}{\partial X_n} \end{pmatrix}.$$

The tangent space to this projective curve at the point $(a_0, ..., a_n)$ is given by the linear system

$$J(a_0, ..., a_n) \begin{pmatrix} X_0 \\ X_1 \\ \vdots \\ X_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

This defines a line as long as $J(a_0, ..., a_n)$ has rank $n-1$. So a space curve is non-singular if J has rank $n-1$ when evaluated at all points on the curve.

We can also check for singularities by checking the affine equations for the curve, then checking for singularities at infinity.

6.4 The Function Field and Automorphism Group of a Curve

Given a curve $\mathcal{X}$ defined over a field $\mathbb{K}$, there is an associated function field $\mathbb{K}(\mathcal{X})$. For a plane curve given by $F(X, Y, Z) = 0$ where $F \in \mathbb{K}[X, Y, Z]$, the function field is given by

$$\mathbb{K}(\mathcal{X}) \cong \text{Frac}(\mathbb{K}[X, Y, Z]/(F)).$$

An automorphism of $\mathcal{X}$ is a morphism from $\mathcal{X}$ to itself which induces an automorphism of $\mathbb{K}(\mathcal{X})$. The set of automorphisms of $\mathbb{K}(\mathcal{X})$ which fix $\mathbb{K}$ and are defined

over some field $\mathbb{L}$ form a group under composition, called the $\mathbb{L}$-automorphism group of $\mathcal{X}$ and denoted by $\mathrm{Aut}_{\mathbb{L}}(\mathcal{X})$. Unless it is otherwise noted, we will assume that $\mathbb{L} = \mathbb{K} = \overline{\mathbb{K}}$ and simply write $\mathrm{Aut}(\mathcal{X})$. A basic fact about automorphism groups is that if the genus of $\mathcal{X}$ is greater than 1, then $\mathrm{Aut}(\mathcal{X})$ is finite. In fact this can be greatly improved. If $\mathcal{X}$ is an irreducible curve of genus $g \geq 2$ defined over a field of characteristic 0, then $|\mathrm{Aut}(\mathcal{X})| \leq 84(g-1)$. This is known as the Hurwitz bound. There are many exceptions in positive characteristic, but automorphism groups larger than this can be considered fairly large. It has been proven that in most cases $|\mathrm{Aut}(\mathcal{X})| \leq 24g^2$. Automorphism groups surpassing this bound are considered to be very large.

For curves $\mathcal{X}$ and $\mathcal{Y}$, a covering morphism $\phi : \mathcal{X} \to \mathcal{Y}$ defined over $\mathbb{K}$ corresponds to an injective homomorphism $\psi : \mathbb{K}(\mathcal{Y}) \to \mathbb{K}(\mathcal{X})$. So $\mathbb{K}(\mathcal{X})$ is an extension of $\mathbb{K}(\mathcal{Y})$.

$$\begin{array}{cc} \mathcal{X} & \mathbb{K}(\mathcal{X}) \\ \downarrow & \uparrow \\ \mathcal{Y} & \mathbb{K}(\mathcal{Y}) \end{array}$$

Two curves $\mathcal{X}$ and $\mathcal{Y}$ are said to be birationally equivalent if there are dominant rational maps $\phi : \mathcal{X} \to \mathcal{Y}$ and $\psi : \mathcal{Y} \to \mathcal{X}$ such that $\phi \circ \psi$ is the identity on an open subset of $\mathcal{Y}$ and $\psi \circ \phi$ is the identity on an open subset of $\mathcal{X}$. Birationally equivalent curves have isomorphic function fields. Two curves are said to be isomorphic if the maps ϕ and ψ are morphisms, i.e. maps that are defined at every point of the curves.

6.5 Galois Extensions and Quotient Curves

If the extension $\mathbb{K}(\mathcal{X})/\mathbb{K}(\mathcal{Y})$ is Galois, $\mathcal{X}$ is said to be a Galois cover of $\mathcal{Y}$. In the case of a Galois cover $\mathrm{Aut}(\mathbb{K}(\mathcal{X})/\mathbb{K}(\mathcal{Y}))$ is called the Galois group of $\mathbb{K}(\mathcal{X})/\mathbb{K}(\mathcal{Y})$ and denoted $\mathrm{Gal}(\mathbb{K}(\mathcal{X})/\mathbb{K}(\mathcal{Y}))$.

In the Galois case we can use Galois theory to see further structure. Let $G = \mathrm{Gal}(\mathbb{K}(\mathcal{X})/\mathbb{K}(\mathcal{Y}))$. Any subgroup H of G corresponds to a subfield $\mathbb{F}$ with $\mathbb{K}(\mathcal{X}) \subset \mathbb{F} \subset \mathbb{K}(\mathcal{X})$, where $\mathbb{F}$ is the fixed field of the group H. If H is a normal subgroup of G then $\mathbb{F}/\mathbb{K}$ is a Galois extension and

$$\mathrm{Gal}(\mathbb{F}/\mathbb{K}) \cong G/H.$$

Any intermediate subfield $\mathbb{K}(\mathcal{Y}) \subset \mathbb{F} \subset \mathbb{K}(\mathcal{X})$ is a finite algebraic extension of $\mathbb{K}(\mathcal{Y})$, so it is the function field of some curve $\mathcal{Z}$. If $\mathbb{F} = \mathbb{K}(\mathcal{Z})$ is the fixed field of H for some $H \leqslant G$, then $\mathcal{Z}$ is called the quotient curve of $\mathcal{X}$ by H and we write $\mathcal{Z} = \mathcal{X}/H$.

6.6 Maximal curves and Zeta functions

For a smooth projective curve $\mathcal{X}$ defined over $\mathbb{F}_q$, the zeta function $Z(\mathcal{X}, t)$ of $\mathcal{X}$ is given by

$$Z(\mathcal{X}, t) := \exp(\Sigma_{i=1}^{\infty} \frac{N_i}{i} t^i) \tag{6.1}$$

By the Weil conjectures, proven by Andre Weil himself in the case of curves, it is known that the zeta function of a nonsingular projective curve converges to a rational function

$$Z(\mathcal{X}, t) = \frac{L(t)}{(1-t)(1-qt),} \tag{6.2}$$

where

$$L(t) = \Pi_{i=1}^{2g}(1 - \omega_i t) \in \mathbb{Z}[t].$$

Here, g is the genus of the curve. By the Riemann hypothesis for curves, again a portion of the Weil conjectures, we have $|\omega_i| = \sqrt{q}$ for all i.

By equating the forms of $Z(\mathcal{X}, t)$ given in (1) and (2), then taking the logarithm of both and equating coefficients, we obtain the relation

$$N_n = q^n + 1 - \Sigma_{i=1}^{2g} \omega_i^n \tag{6.3}$$

Consider a curve $\mathcal{X}$ defined over $\mathbb{F}_q$. The fact that $|\omega_i| = \sqrt{q}$ imposes that

$$-2g\sqrt{q} \leq \Sigma_{i=1}^{2g}\omega_i \leq 2g\sqrt{q}.,$$

giving the bound

$$q + 1 - 2g\sqrt{q} \leq N_1 \leq q + 1 + 2g\sqrt{q}.$$

If $\mathcal{X}$ is maximal over $\mathbb{F}_q$ according to this bound, then

$$N_1 = q + 1 + 2g\sqrt{q}$$

which requires that $\omega_i = -\sqrt{q}$ for all i.

An $\mathbb{F}_q$-maximal curve has $q + 1 + 2g\sqrt{q}$ points over $\mathbb{F}_q$. If this is to be an integer, q must be a perfect square. For this reason, we will generally consider curves defined over $\mathbb{F}_{q^2}$.

We then have the following formula for the zeta function of an $\mathbb{F}_{q^2}$-maximal curve $\mathcal{X}$

$$Z(\mathcal{X}, t) = \frac{(1 + qt)^{2g}}{(1 - t)(1 - q^2t)}.$$

Equation (3) gives the number of points on $\mathcal{X}$ over all extension fields of $\mathbb{F}_{q^2}$. Since $\omega_i = -q$ for all i, we have

$$N_n = q^{2n} + 1 - 2g(-q)^n.$$

Therefore a curve which is maximal over a given finite field will be maximal over odd degree extensions of that field and be minimal (attain the lower Hasse-Weil bound) over extensions of even degree.

6.7 Upper Bound on Genera of Maximal Curves

The zeta function of a maximal curve gives an upper bound on possible genera for maximal curves over a given finite field.

Theorem 7. *Let $\mathcal{X}$ be a $\mathbb{F}_{q^2}$-maximal curve. Then $g(\mathcal{X}) \leq \frac{q(q-1)}{2}$.*

Proof. Since $\mathbb{F}_{q^2} \subset \mathbb{F}_{q^4}$, it must be that

$$\#\mathcal{X}(\mathbb{F}_{q^2}) \leq \#\mathcal{X}(\mathbb{F}_{q^4}).$$

That means

$$\begin{aligned} q^2 + 1 + 2qg(\mathcal{X}) &\leq q^4 + 1 - 2q^2 g(\mathcal{X}) \\ 2g(\mathcal{X})(q + q^2) &\leq q^4 - q^2 = q^2(q+1)(q-1) \\ g(\mathcal{X}) &\leq \frac{q(q-1)}{2}. \end{aligned}$$

□

6.8 Newton Polygon of a Maximal Curve

Newton polygons encode information about polynomials or power series over local fields. The Newton polygon of an algebraic curve is defined based on its zeta function, more specifically the numerator of the zeta function, known as its L-polynomial. Though we use the fact that $L(t) \in \mathbb{Z}[t]$ in the definition of the Newton polygon this is not necessary. The method of Newton polygons can be applied to any polynomial with coefficients in a complete local field, for example the completion of the p-adics. We will determine the Newton polygon of a maximal curve, then state a result that shows maximal curves are supersingular. Background on Newton polygons and elliptic curves can be found in Koblitz [20] and Silverman [40], respectively.

Definition 2. *Let $L(t) = a_0 + a_1 t + a_2 t^2 + ... + a_n t^n$ be the L-polynomial for an smooth projective curve $\mathcal{X}$ defined over $\mathbb{F}_q$, where $q = p^h$ for some prime p. By the Weil conjectures, we know that $L(t) \in \mathbb{Z}[t]$. Let $\nu_p(x)$ denote the p-adic valuation of $x \in \mathbb{Q}$. Let $\nu_p'(x) = \frac{\nu_p(x)}{h}$ be the normalized valuation. Define the Newton polygon of*

$\mathcal{X}$ with respect to q be the lower convex hull of the points $P_i = (i, \nu'(a_i))$ for $0 \leq i \leq n$, plotted on a standard xy-plane.

To illustrate the lower convex hull, one can imagine placing pins in the points described on a graph, then stretching a rubber band from below around the pins, coming from the negative y direction. The resulting shape of the rubber band is the lower convex hull. So a Newton polygon is made up of several straight line segments between points.

We will introduce a few ideas to make an interesting statement about maximal curves based on their Newton polygons. For more on abelian varieties and Jacobians, see Serre [37].

An abelian variety is an algebraic variety which is also an abelian group. Given a smooth, irreducible projective curve $\mathcal{X}$, one can define an abelian variety called the Jacobian of $\mathcal{X}$ which is denoted $\mathrm{Jac}(\mathcal{X})$.

If A and B are abelian varieties, an isogeny $A \to B$ is a surjective morphism with a finite kernel. We say A is isogenous to B if such an isogeny exists. Isogeny is an equivalence relation, coarser than isomorphism.

An elliptic curve is a smooth irreducible curve of genus 1 with a rational point. An elliptic curve is isomorphic to its own Jacobian, i.e. it comes equipped with a group structure of its own. For n a natural number, the n-torsion points of an elliptic curve over the algebraic closure of the base field form a subgroup. An elliptic curve defined over $\mathbb{F}_q$, where $q = p^h$ with p prime, has n-torsion over $\overline{\mathbb{F}_q}$ isomorphic to $C_n \times C_n$ for all n relatively prime to p. For n such that $p|n$, the n-torsion can vary. In particular, the p-torsion of an elliptic curve can either be trivial or isomorphic to C_p. Elliptic curves of the latter type are called ordinary, while elliptic curves with trivial p-torsion are called supersingular.

A proof of the following can be found in Yuri Manin's 1963 thesis [28].

Fact 1. *If the Newton polygon of a curve $\mathcal{X}$ has all slopes equal to $\frac{1}{2}$, then $Jac(\mathcal{X})$ is isogenous to a product of supersingular elliptic curves.*

A curve of genus ≥ 1 which is isogenous to the product of supersingular elliptic curves is also called supersingular. The Jacobian of a supersingular curve has trivial p-torsion, however supersingularity is a stronger condition than trivial p-torsion.

This leads to a statement about the structure of the Jacobian of a maximal curve:

Theorem 8. *Maximal curves are supersingular. If $\mathcal{X}$ is an $\mathbb{F}_{p^{2h}}$-maximal curve for some p prime, then the p-torsion of $Jac(\mathcal{X})(\overline{\mathbb{F}_p})$ is trivial.*

Proof. The Newton polygon of an $\mathbb{F}_{q^2}$-maximal curve $\mathcal{X}$ with respect to $\mathbb{F}_{q^2}$ can be fairly easily determined by expanding the L-polynomial determined above. Since

$$L(t) = (1+qt)^{2g} = \Sigma_{i=0}^{2g} \binom{2g}{i} q^i t^i,$$

we can see that $a_i = \binom{2g}{i} q^i$. We can check that $\nu'(\binom{2g}{i}) = \nu'(1) = 0$ for $i = 0$ and $i = 2g$, and $\nu'(\binom{2g}{i}) \geq 0$ for all $1 \leq i \leq 2g-1$. Then, considering that $\nu'(q^i) = \frac{i}{2}$, we can see that $P_1 = (0,0)$, $P_{2g} = (2g, g)$, and all P_i for $1 \leq i \leq 2g-1$ lie above the straight line between these two points. Thus the lower convex hull is a straight line of slope $\frac{1}{2}$ from $(0,0)$ to $(2g,g)$. This is the Newton polygon of $\mathcal{X}$. Thus by fact 1, $\mathcal{X}$ is supersingular. □

6.9 Example: The Hermitian Curve

6.9.1 The Hermitian curve $\mathcal{H}_q$

Perhaps the best known example of a maximal curve is the Hermitian curve. For $q = p^n$, define the Hermitian curve $\mathcal{H}_q$ to be the projective curve with the following affine equation

$$\mathcal{H}_q : h_q(x,y) = x^q + x - y^{q+1} = 0.$$

Theorem 9. *The Hermitian curve $\mathcal{H}_q$ is maximal over $\mathbb{F}_{q^2}$.*

Proof. The number of points possible for a curve depends on its genus, so the first step is calculating the genus of $\mathcal{H}_q$. The Plucker formula gives the genus g of a smooth projective plane curve $\mathcal{X}$ of degree d as

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

To see that $\mathcal{H}_q$ is smooth, it is sufficient to show that there is no point at which all partial derivatives of the homogenized form of the equation defining $\mathcal{H}_q$ vanish. Consider the homogenized equation for $\mathcal{H}_q$

$$H_q(X, Y, Z) := X^q Z + X Z^q - Y^{q+1} = 0.$$

The conditions imposed by the partial derivatives vanishing are as follows:

$$\begin{aligned} 0 = \frac{\partial H_q}{\partial X} &= Z^q \Rightarrow Z = 0 \\ 0 = \frac{\partial H_q}{\partial Y} &= -Y^q \Rightarrow Y = 0 \\ 0 = \frac{\partial H_q}{\partial Z} &= X^q \Rightarrow X = 0. \end{aligned}$$

Since $X = Y = Z = 0$ is not a point of $\mathbb{P}^2$, $\mathcal{H}_q$ is smooth. The Plucker formula gives that $g(\mathcal{H}_q) = \frac{q(q-1)}{2}$. Therefore the Hasse-Weil bound requires that

$$\begin{aligned} \#\mathcal{H}_q(\mathbb{F}_{q^2}) &\leq q^2 + 1 + 2gq \\ &= q^3 + 1. \end{aligned}$$

Now we can count the points on the curve using the field trace and norm maps.

First note that $\mathcal{H}_q$ has a single point at infinity since $Z = 0$ implies that $Y^{q+1} = 0$, so $Y = 0$, and $[1 : 0 : 0]$ is the only non-affine point on the curve. The affine points over $\mathbb{F}_{q^2}$ are solutions to

$$x^q + x = y^{q+1}$$

with $x, y \in \mathbb{F}_{q^2}$. Note that $x^q + x = \mathrm{Tr}_{\mathbb{F}_{q^2}/\mathbb{F}_q}(x)$, where the trace map $\mathrm{Tr}_{\mathbb{F}_{q^2}/\mathbb{F}_q}$ is a degree q homomorphism of additive groups mapping $\mathbb{F}_{q^2} \to \mathbb{F}_q$. That means that for every $\alpha \in \mathbb{F}_q$, there are q values of $x \in \mathbb{F}_{q^2}$ such that $x^q + x = \alpha$.

Next, notice that $y^{q+1} = \mathrm{N}_{\mathbb{F}_{q^2}/\mathbb{F}_q}(y)$, where the norm map $\mathrm{N}_{\mathbb{F}_{q^2}/\mathbb{F}_q}$ is a degree $q+1$ homomorphism of multiplicative groups mapping $\mathbb{F}_{q^2}^* \to \mathbb{F}_q^*$. So for each nonzero $\alpha \in \mathbb{F}_q$, there are $q+1$ values of $y \in \mathbb{F}_{q^2}$ such that $y^{q+1} = \alpha$. For $\alpha = 0$, the only solution is $y = 0$.

Counting these possibilities, we have $q-1$ non-zero α in $\mathbb{F}_q$, each of which are mapped to by q values of x and $q+1$ values of y in $\mathbb{F}_{q^2}$. For $\alpha = 0$ there are still q values of x but only 1 value of y. Adding the point at infinity gives a total of

$$(q-1)(q+1)(q) + q + 1 = q^3 + 1$$

points over $\mathbb{F}_{q^2}$.

□

Notice that the Hermitian curve $\mathcal{H}_q$ attains the upper bound for the genus of an $\mathbb{F}_{q^2}$ maximal curve, proving that the bound in Theorem 1 is tight.

The following fact was proven in 1994 by Rück and Stichtenoth [35], and by Fuhrmann and Torres in 1996 using different methods [10].

Fact 2. *Any $\mathbb{F}_{q^2}$-maximal curve with genus $\frac{q(q-1)}{2}$ is isomorphic over $\mathbb{F}_{q^2}$ to $\mathcal{H}_q$.*

Not every genus below the bound of $\frac{q(q-1)}{2}$ is obtainable for an $\mathbb{F}_{q^2}$-maximal curve. The Hermitian curve takes on the upper bound, but Fuhrmann and Torres proved that the next largest obtainable genus is $\frac{(q-1)^2}{4}$ [10], the genus of the $\mathbb{F}_{q^2}$-maximal curve $\mathcal{F}$ given by the affine equation

$$f(x, y) := x^q + x - y^{\frac{(q+1)}{2}} = 0,$$

which is covered by $\mathcal{H}_q$ under the map

$$\begin{aligned} x &\mapsto x \\ y &\mapsto y^2. \end{aligned}$$

By the following result of Serre, this map forces $\mathcal{F}$ to be maximal as well.

Fact 3. *Any curve which is $\mathbb{F}_{q^2}$-covered by an $\mathbb{F}_{q^2}$ maximal curve is also $\mathbb{F}_{q^2}$-maximal.*

In particular, the quotient curves of a maximal curve $\mathcal{X}$ by subgroups of $\mathrm{Aut}_{\mathbb{F}_{q^2}}(\mathcal{X})$ are also maximal. Curves with large automorphism groups have the potential to generate many quotient curves, so are a good source of maximal curves.

Many known examples of maximal curves are covered by the Hermitian curve, and until recently no maximal curve had been proven not to be covered by the Hermitian curve or one of three other curves, namely the Deligne-Lustzig curve associated to the Suzuki group (DLS), the Deligne-Lustzig curve associated to the Ree group (DLR), and the Garcia-Stichtenoth curve. In 2007, Giuletti and Korchmaros introduced a family of curves maximal over $\mathbb{F}_{q^6}$ which they proved were not covered by the any of these curves [12]. A 2008 article by Garcia, Guneri, and Stichtenoth generalizes Giulietti and Korchmaros' family, introducing curves maximal over $\mathbb{F}_{q^{2n}}$ for all odd $n \geq 3$ [11]. It is not known whether these curves are covered by a known maximal curve for $n \geq 5$. These families of curves will be the main focus of this portion of the paper.

6.9.2 Other equations for the Hermitian curve

The equation considered above is not the only one valid for the Hermitian curve. A Hermitian curve can be understood by its relationship to a Hermitian form, defined to be a symmetric sesquilinear form $h : V \times V \mapsto \mathbb{K}$ such that $h(v, w) = \overline{h(w, v)}$, where V is a $\mathbb{K}$-vector space and $\overline{x}$ is the conjugate of $x \in \mathbb{K}$. If $\mathbb{K} = \mathbb{F}_{q^2}$, the conjugate considered is the element's image under the q-th power map, i.e. $\overline{x} = x^q$. For more

on Hermitian forms, see Grove [14]. Given a basis for V, a Hermitian form can be expressed as a Hermitian matrix, by definition an $n \times n$ matrix M such that

$$M = \overline{M}^T,$$

that is, if $m_{i,j}$ is the i,j-th entry of M, then $m_{i,j} = \overline{m_{j,i}}$.

A Hermitian matrix gives rise to a homogeneous polynomial and projective curve as follows. Let $(X, Y, Z) = v$ and M be a non-singular 3×3 Hermitian matrix defined over $\mathbb{F}_{q^2}$. Then $v^T M \overline{v}$ is a homogenous polynomial, the projective vanishing of which is a Hermitian curve. To see $\mathcal{H}_q$ in this light, let

$$M = \begin{pmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

Then

$$\begin{aligned} \begin{pmatrix} X & Y & Z \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} \overline{X} \\ \overline{Y} \\ \overline{Z} \end{pmatrix} &= \begin{pmatrix} X & Y & Z \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} X^q \\ Y^q \\ Z^q \end{pmatrix} \\ &= \begin{pmatrix} Z & -Y & Z \end{pmatrix} \begin{pmatrix} X^q \\ Y^q \\ Z^q \end{pmatrix} \\ &= X^q Z + X Z^q - Y^{q+1}. \end{aligned}$$

The identity matrix is also a Hermitian matrix, corresponding to the curve $\mathcal{H}'_q$ with equation $X^{q+1} + Y^{q+1} + Z^{q+1} = 0$. This curve has the same genus as $\mathcal{H}_q$ and is also $\mathbb{F}_{q^2}$-maximal, so we know that the curves are isomorphic over $\mathbb{F}_{q^2}$. To see that the curves are isomorphic it would also serve to construct an invertible morphism from one curve to the other, as below.

Let the points of $\mathcal{H}'_q$ be given by $R^{q+1} + S^{q+1} + T^{q+1} = 0$. Then choose $\delta, \gamma \in \mathbb{F}_{q^2}$ so that $\delta^{q+1} = \gamma^q + \gamma = -1$ and define

$$X := \delta(1+\gamma)R - \gamma S \qquad Y := \delta T \qquad Z := S - \delta R.$$

We can then calculate directly that

$$X^q Z + X Z^q - Y^{q+1} = R^{q+1} + S^{q+1} + T^{q+1} = 0.$$

We have just seen that $\mathcal{H}_q$ is isomorphic to $\mathcal{H}'_q$. Now we will see that all curves arising from Hermitian forms are isomorphic to $\mathcal{H}'_q$, and so isomorphic to $\mathcal{H}_q$.

Proposition 14. *All non-singular Hermitian matrices with entries in $\mathbb{F}_{q^2}$ give rise to $\mathbb{F}_{q^2}$ isomorphic curves.*

Proof. To demonstrate that the curves are $\mathbb{F}_{q^2}$-isomorphic, it is sufficient to give an invertible homogeneous linear transformation defined over $\mathbb{F}_{q^2}$ between them. We use the fact that any Hermitian matrix with entries in $\mathbb{F}_{q^2}$ can be diagonalized by a unitary matrix with entries in $\mathbb{F}_{q^2}$, and the resulting diagonal matrix will take on values in $\mathbb{F}_q$ (thus is also a Hermitian matrix).

Let M be a non-singular Hermitian matrix, U be a unitary matrix which diagonalizes M. Let v be a vector in $\mathbb{P}^2$ such that $v^T M \overline{v} = 0$. Then

$$0 = v^T U U^{-1} M U U^{-1} \overline{v}.$$

Since U is unitary, $U^{-1} = \overline{U}^T$. If $w = U^T v$, then $v^T U = w^T$ and $U^{-1}\overline{v} = \overline{U}^T \overline{v} = \overline{w}^T$. That means

$$0 = w^T (U^{-1} M U) \overline{w}.$$

So the matrix U defines an invertible linear transformation from points on the curve $\mathcal{M}$, associated to M, to points on the curve $\mathcal{D}_q$ associated to the diagonal Hermitian matrix $D = U^{-1} M U$, meaning that $\mathcal{M}$ and $\mathcal{D}_q$ are isomorphic over $\mathbb{F}_{q^2}$. Now we prove that $\mathcal{D}_q$ is isomorphic to $\mathcal{H}'_q$.

The defining equation for the curve $\mathcal{D}_q$ associated to D is as follows:

$$D = \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix} \qquad \leftrightarrow \qquad \mathcal{D}_q : aX^{q+1} + bY^{q+1} + cZ^{q+1} = 0,$$

where $a, b, c \in \mathbb{F}_q^*$.

The curve $\mathcal{D}_q$ then maps to the curve $\mathcal{H}'_q$ defined as follows. Choose $\alpha, \beta, \sigma \in \mathbb{F}_{q^2}^*$ so that

$$\alpha^{q+1} = a$$
$$\beta^{q+1} = b$$
$$\sigma^{q+1} = c.$$

The values α, β, and γ exist in $\mathbb{F}_{q^2}^*$ because $x^{q+1} = \mathrm{N}_{\mathbb{F}_{q^2}/\mathbb{F}_q}(x)$ maps $\mathbb{F}_{q^2}^*$ onto $\mathbb{F}_q^*$. Define the matrix N:

$$N = \begin{bmatrix} \alpha & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & \sigma \end{bmatrix}.$$

Given $v = (X, Y, Z)$ a point on $\mathcal{D}_q$, let $Nv = w$. Then $w = (\alpha X, \beta Y, \sigma Z)$, and

$$(\alpha X)^{q+1} + (\beta Y)^{q+1} + (\sigma Z)^{q+1} = aX^{q+1} + bY^{q+1} + cZ^{q+1} = 0,$$

So Nw is a point of $\mathcal{H}'_q$.

Therefore NU is a map from the points of $\mathcal{M}$, the curve arising from a Hermitian matrix M, to $\mathcal{H}'_q$, a standard version of the Hermitian curve. This is an isomorphism since U and N are invertible.

□

In the special case of $M = \begin{bmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$ discussed earlier, we could use $U =$

$$\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \text{ and } N = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \gamma & 0 \\ 0 & 0 & 1 \end{bmatrix}, \text{ where } \gamma \text{ is an element of } \mathbb{F}_{q^2} \text{ with } \gamma^{q+1} = -1.$$

We now know that any Hermitian form generates a curve isomorphic to $\mathcal{H}'_q$. Now, given a curve $\mathcal{F}$ in $\mathbb{P}^2$ and an isomorphism $\mathtt{T} : \mathcal{F} \to \mathcal{H}'_q$, can we find a Hermitian matrix H so that the points of $\mathcal{F}$ are the vanishing of the form H?

Let $w = (X, Y, Z)$ be a point of $\mathcal{F}$. The isomorphism $\mathtt{T}$ is an invertible linear transformation mapping points of $\mathcal{F}$ to points of $\mathcal{H}'_q$, so may be given as an invertible matrix. Then let $\mathtt{T}w = v$, where $v = (X', Y', Z')$ is a point of $\mathcal{H}_q$. Since the defining property of v is that

$$v^T \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \overline{v} = 0,$$

we have that

$$(\mathtt{T}w)^T \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \overline{\mathtt{T}w} = 0.$$

Simplifying, this gives

$$w^T \mathtt{T}^T \overline{\mathtt{T}} \overline{w} = 0.$$

Let $H = \mathtt{T}^T \overline{\mathtt{T}}$. Then

$$\overline{H}^T = \overline{\mathtt{T}^T \overline{\mathtt{T}}}^T = \mathtt{T}^T \overline{\mathtt{T}} = H$$

so H is Hermitian. Therefore given an isomorphism from a curve in $\mathbb{P}^2$ to the Hermitian curve $\mathcal{H}'_q$, we can write the curve as the vanishing of a corresponding Hermitian form.

Since all curves arising from this type of Hermitian form are isomorphic it makes sense to use the most convenient form for a given problem. The curve $\mathcal{H}_q$ is convenient because it is a very commonly used form and has only a single point at infinity. In

contrast, the curve $\mathcal{H}'_q$, with corresponding affine equation $x^{q+1} + y^{q+1} + 1 = 0$, has $q + 1$ points at infinity.

6.10 Proving Maximality in Two Families of Maximal Curves

We have already seen two methods of proving that a curve is maximal. The first, as demonstrated for the Hermitian curve, is simply to count points. A second method, used for $\mathcal{F}$ above, is to show that the curve in question is $\mathbb{F}_{q^2}$-covered by another maximal curve. We will consider two more methods, which were employed by Giulietti and Korchmaros and Garcia, Guneri, and Stichtenoth in the papers mentioned above.

The natural embedding theorem, due to Korchmaros and Torres in 2001, completely characterizes maximal curves, though it gives no explicit equations [16].

Fact 4. *Let $\mathcal{X}$ be a smooth projective curve defined over $\mathbb{F}_{q^2}$. Then $\mathcal{X}$ is maximal over $\mathbb{F}_{q^2}$ if and only if $\mathcal{X}$ is isomorphic over $\mathbb{F}_{q^2}$ to a smooth, absolutely irreducible curve of degree $q + 1$ lying on a non-degenerate Hermitian variety $\mathcal{H}_{m,q}$.*

Here $\mathcal{H}_{m,q}$ is the m-dimensional analogue of the Hermitian curve. For example, consider the 4×4 identity matrix, which gives rise to the 2-dimensional projective Hermitian variety associated with the equation $W^{q+1} + X^{q+1} + Y^{q+1} + Z^{q+1} = 0$. A curve can be proven to be maximal by showing that it is isomorphic to a curve on this surface which is smooth, absolutely irreducible, and of degree $q+1$. Giulietti and Korchmaros employ this method to prove that the family of curves we now define are maximal.

6.10.1 Giulietti and Korchmaros' family of maximal curves

Define the curve $\mathcal{Z}$ to be the intersection of the surfaces Σ and $\mathcal{H}_q$ with affine equations given by

$$\Sigma \quad : \quad \sigma(x,y,z) := z^{\frac{q^3+1}{q+1}} - y\frac{x^{q^2}-x}{x^q+x} = 0, \tag{6.4}$$

$$\mathcal{H}_q \quad : \quad h_q(x,y,z) := x^q + x - y^{q+1} = 0. \tag{6.5}$$

Notice that $\frac{x^{q^2}-x}{x^q+x}$ is a polynomial with degree q^2-q which has zeros at exactly the points of $\mathbb{F}_{q^2}$ whose trace in $\mathbb{F}_q$ is non-zero. Projective equations can be obtained by homogenizing both equations with a single variable. These equations are

$$\begin{aligned} \Sigma \quad &: \quad S(X,Y,Z,W) := Z^{\frac{q^3+1}{q+1}} - Y\frac{X^{q^2}W - XW^{q^2}}{X^qW + XW^q} = 0, \\ \mathcal{H}_q \quad &: \quad H_q(X,Y,Z,W) := X^qW + XW^q - Y^{q+1} = 0. \end{aligned}$$

The intersection has the single point $Y = Z = W = 0$, $X = 1$ at infinity. Giulietti and Korchmaros demonstrate that $\mathcal{Z}$ is $\mathbb{F}_{q^6}$-maximal. This will be outlined in several claims.

Claim 1. *The curve $\mathcal{Z}$ has degree q^3+1.*

For a non-plane curve such as this one we need a new definition of degree. The degree of a curve $\mathcal{X} \subset \mathbb{P}^n$ is denoted $\deg(\mathcal{X})$ and is defined to be the maximum number of points (counted with proper multiplicity) of intersection between $\mathcal{X}$ and a hyperplane not containing any component of $\mathcal{X}$. We will make use of the notion of the notion if the dimension of a variety $\mathcal{V}$, denoted $\dim(\mathcal{V})$ also. Without launching into a discussion of dimension, we can get by with the key fact that the dimension of a projective variety defined by a single homogeneous equation is 1 less than that of the ambient space. For a deeper discussion of dimension and proofs of the facts used below, see Shafarevich [39].

Proof. Consider a hyperplane $\mathcal{P}$ in $\mathbb{P}^3$, defined to be the vanishing of a single linear homogeneous polynomial in X, Y, Z, and W. Choose $\mathcal{P}$ so that it has a non-empty intersection with each of Σ and $\mathcal{H}_q$, is not tangent with either, and contains no components of either. We know that $\mathcal{P}$ exists because $\dim(\mathcal{P}) = \dim(\Sigma) = \dim(\mathcal{H}_q) = 2$, and two projective varieties must have a non-empty intersection if their dimensions sum to at least the dimension of the ambient space. We can avoid both of their tangent spaces because, as smooth surfaces, each has a tangent space of dimension 2. We can avoid containing components of the surfaces because each is irreducible and does not itself lie in any hyperplane.

Having chosen $\mathcal{P}$ so carefully, we can see that the intersection in $\mathbb{P}^3$ of Σ and $\mathcal{P}$ will be a curve of degree $\frac{q^3+1}{q+1}$, perhaps reducible. Similarly, the intersection in $\mathbb{P}^3$ of $\mathcal{H}_q$ and $\mathcal{P}$ will be a curve of degree $q+1$. The intersection of $\Sigma \cap \mathcal{H}_q$ with $\mathcal{P}$ will be the intersection of these curves. Bezout's theorem implies that the number of points of intersection (counted with multiplicity) of a curve of degree d with a curve of degree e is de. So $\deg(\mathcal{Z}) = \frac{q^3+1}{q+1}(q+1) = q^3 + 1$. □

Claim 2. *The curve $\mathcal{Z}$ is smooth.*

Proof. First we'll look at the affine patch with equations given in (4) and (5). Consider the matrix of partial derivatives, calculated here:

$$J := \begin{pmatrix} \frac{\partial\sigma}{\partial x} & \frac{\partial\sigma}{\partial y} & \frac{\partial\sigma}{\partial z} \\ \frac{\partial h_q}{\partial x} & \frac{\partial h_q}{\partial y} & \frac{\partial h_q}{\partial z} \end{pmatrix} = \begin{pmatrix} -y\frac{d}{dx}\left(\frac{x^{q^2}-x}{x^q+x}\right) & -\frac{x^{q^2}-x}{x^q+x} & \frac{q^3+1}{q+1}z^{\frac{q^3+1}{q+1}-1} \\ 1 & -y^q & 0 \end{pmatrix}$$

The only possibility for the rank of J to be less than 2 is if $z = 0$. If $z = 0$, then from (4) we know that either $y = 0$ or $\frac{x^{q^2}-x}{x^q+x} = 0$. If $y = 0$, then $x^q + x = 0$ from (5). But that means that the formal polynomial $\frac{x^{q^2}-x}{x^q+x}$ evaluated at x is non-zero, so the matrix again has rank 2. If $\frac{x^{q^2}-x}{x^q+x} = 0$, then $x^q + x \neq 0$, so $y^{q+1} \neq 0$ from (5), and $-y^q \neq 0$ either, and J has rank 2 as long as $\frac{\partial\sigma}{\partial x} \neq 0$. The roots of $\frac{x^{q^2}-x}{x^q+x}$ are the elements of $\mathbb{F}_{q^2}$ with trace in $\mathbb{F}_q$ not equal to zero, so we know they are distinct, i.e.

$\frac{x^{q^2}-x}{x^q+x}$ is a separable polynomial. Therefore it shares no roots with its derivative, and since $y \neq 0$, we have $y\frac{d}{dx}(\frac{x^{q^2}-x}{x^q+x}) \neq 0$. Therefore J has rank 2 in any case and $\mathcal{Z}$ has no singular affine points.

Now consider the point at infinity. We can cover this with the affine patch where $X \neq 0$, in which case we can dehomogenize as follows:

$$\sigma'(w, y, z) := z^{\frac{q^3+1}{q+1}} + y\frac{w^{q^2} - w}{w^q + w} = 0, \tag{6.6}$$

$$h_q(w, y, z) := w^q + w - y^{q+1} = 0. \tag{6.7}$$

These equations are extremely similar to those in (4) and (5), and a similar reasoning process leads us to conclude that the curve is smooth on this affine patch as well. So $\mathcal{Z}$ is smooth everywhere. □

To prove that $\mathcal{Z}$ lies on a Hermitian surface, we will need a polynomial identity.

Claim 3. *In* $\mathbb{F}_{q^2}[x]$,

$$(\frac{x^{q^2} - x}{x^q + x})^{q+1}(x^q + x) = x^{q^3} + x - (x^q + x)^{\frac{q^3+1}{q+1}}.$$

Proof. Calculation shows that

$$(x^q - x)^q(x^{q^3} - x + (x^q - x)^{\frac{q^3+1}{q+1}}) = (x^{q^2} - x)^{q+1}. \tag{6.8}$$

Now choose $\rho \in \mathbb{F}_{q^2}$ so that $\rho \neq 0$ but $\mathrm{Tr}(\rho) = \rho^q + \rho = 0$. We know that such a ρ exists because the trace map surjects onto $\mathbb{F}_q$. Since (8) holds for all $x \in \overline{\mathbb{F}_{q^2}}$, replace x by ρx to obtain

$$\begin{aligned}
(\rho^q x^q - \rho x)^q(\rho^{q^3} x^{q^3} - \rho x + (\rho^q x^q - \rho x)^{\frac{q^3+1}{q+1}}) &= (\rho^{q^2} x^{q^2} - \rho x)^{q+1} \\
-\rho^2(x^q + x)^q((x^{q^3} + x) - (x^q + x)^{\frac{q^3+1}{q+1}}) &= -\rho^2(x^{q^2} - x)^{q+1} \\
x^{q^3} + x - (x^q + x)^{\frac{q^3+1}{q+1}} &= (\frac{x^{q^2} - x}{x^q + x})^{q+1}(x^q + x).
\end{aligned}$$

□

Claim 4. *The curve $\mathcal{Z}$ lies on the Hermitian surface $\mathcal{H}$ with affine equation*

$$x^{q^3} + x = y^{q^3+1} + z^{q^3+1}.$$

Note that this is an affine form of the Hermitian surface corresponding to the matrix $\begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$, which could be denoted by $\mathcal{H}_{2,q^3}$.

Proof. Let $P := (x, y, z)$ be an affine point of $\mathcal{Z}$. Since P lies on Σ, we know that

$$z^{\frac{q^3+1}{q+1}} = y\frac{x^{q^2} - x}{x^q + x}.$$

Raising both sides of this equality to the $(q+1)$-st power, we get

$$z^{q^3+1} = y^{q+1}(\frac{x^{q^2} - x}{x^q + x})^{q+1}.$$

Using the fact that $y^{q+1} = x^q + x$ and the identity from claim 3, we have

$$z^{q^3+1} = (x^q + x)(\frac{x^{q^2} - x}{x^q + x})^{q+1} = x^{q^3} + x - (x^q + x)^{\frac{q^3+1}{q+1}}.$$

Using $y^{q+1} = x^q + x$ again, we obtain

$$z^{q^3+1} = x^{q^3} + x - y^{q^3+1}.$$

The point at infinity on $\mathcal{Z}$ is $[1 : 0 : 0 : 0]$, which is also the point at infinity on $\mathcal{H}$. □

Claim 5. *$\mathcal{Z}$ is absolutely irreducible.*

Proof. This proof uses results from intersection theory. More on this can be found in Hirschfeld, Korchmaros and Torres [16]. This claim is proven by considering the function field of an absolutely irreducible component of $\mathcal{Z}$. Let $\mathcal{Y}$ be the absolutely irreducible component of $\mathcal{Z}$ containing the affine point $(x, y, z) = (0, 0, 0)$. It is straightforward to check that $\mathcal{Z}$ is nonsingular at $(0, 0, 0)$. Let $\mathbb{K} = \overline{\mathbb{F}_{q^2}}$, so $\mathbb{K}(\mathcal{Y})$ is

the function field of $\mathcal{Y}$. Since $\mathcal{Z}$ is embedded in $\mathbb{P}^3$ we also have an embedding of $\mathcal{Y}$ in $\mathbb{P}^3$. Let x, y, z, t $\in \mathbb{K}(\mathcal{Y})$ be the coordinate functions of this embedding and consider the affine model of the curve where $t = 1$, so $\mathbb{K}(\mathcal{Y}) \cong \text{Frac}(\mathbb{K}[\text{x, y, z}]/\sim)$ where $\sim$ indicates algebraic relations between x, y, and z.

For a non-zero function f(x, y, z), let $v_{(0,0,0)}(f(\text{x, y, z}))$ denote the valuation of f(x, y, z) $\in \mathbb{K}(\mathcal{Y})$ at $(0, 0, 0)$. The valuation can be understood as the order of vanishing of a function at the given point. Since $\mathcal{Y}$ lies on $\mathcal{H}$, we know that

$$\begin{aligned} \text{x}^{q^3} + \text{x} &= \text{y}^{q^3+1} + \text{z}^{q^3+1} \\ \text{x}(\text{x}^{q^3-1} + 1) &= \text{y}^{q^3+1} + \text{z}^{q^3+1} \end{aligned}$$

Since $(0, 0, 0)$ is a zero, we have

$$v_{(0,0,0)}(\text{x}) = v_{(0,0,0)}(\text{x}^{q^3} + \text{x}) = v_{(0,0,0)}(\text{y}^{q^3+1} + \text{z}^{q^3+1}) \geq q^3 + 1$$

Consider the plane π defined by $x = 0$. Then the degree of the intersection of π and $\mathcal{Y}$ at $(0, 0, 0)$, called the intersection number and denoted $I((0, 0, 0), \pi \cap \mathcal{Y})$, is at least q^3+1. However, $I((0, 0, 0), \pi \cap \mathcal{Z}) \leq \deg(\pi)\deg(\mathcal{Z}) = q^3+1$. So either $\mathcal{Y} = \mathcal{Z}$ or $\mathcal{Y} \subset \pi$. But by examining the equations for $\mathcal{Z}$, we can see that $\mathcal{Z} \cap \pi = \{(0, 0, 0) \cup \infty\}$. So $\mathcal{Y} = \mathcal{Z}$, and $\mathcal{Z}$ is absolutely irreducible.

□

By the natural embedding theorem, these claims imply that $\mathcal{Z}$ is maximal. Notice that we are in the strange position of knowing that the curve is maximal without knowing how many points it has, as by this reasoning we do not yet know the genus of $\mathcal{Z}$. Giulietti and Korchmaros determine the genus by finding that $\mathcal{Z}$ has the Hermitian curve as a quotient, then using the Riemann-Hurwitz formula.

6.10.2 Garcia, Guneri, and Stichtenoth's family $\mathcal{C}_n$

Recall for $n \geq 3$ odd $\mathcal{C}_n$ is the curve defined in Section 1. Giulietti and Korchmaros' curve $\mathcal{Z}$ discussed in Section 6.10.1 is also the curve $\mathcal{C}_3$. In proving the maximality of $\mathcal{C}_n$ for arbitrary n, the following cover is considered:

$$\begin{array}{cc} \mathcal{C}_n & (x,y,z) \\ \pi' \downarrow & \downarrow \\ \mathcal{X}_n & (y,z) \\ \psi \downarrow & \downarrow \\ \mathbb{P}_1 & y \end{array}$$

Abdon, Bezerra, and Quoos earlier determined the genus of $\mathcal{X}_n$ is $\frac{(q-1)(q^n-q)}{2}$, and proved that $\mathcal{X}_n$ is $\mathbb{F}_{q^{2n}}$-maximal [1]. By proving that every $\mathbb{F}_{q^{2n}}$-point of $\mathcal{X}_n$ except the point at infinity splits completely in $\mathcal{C}_n$, Garcia, Guneri, and Stichtenoth show that

$$\begin{aligned} \#\mathcal{C}_n(\mathbb{F}_{q^{2n}}) &= 1 + \deg(\pi')(\#\mathcal{X}_n(\mathbb{F}_{q^{2n}}) - 1) && (6.9) \\ &= 1 + q(\#\mathcal{X}_n(\mathbb{F}_{q^{2n}}) - 1). && (6.10) \end{aligned}$$

They then determine the genus of the curve $\mathcal{C}_n$. Note that since we do not have one explicit equation for $\mathcal{C}_n$, we can not use the Plucker formula as we did for $\mathcal{H}_q$. Then, considering the genus, the equality in (1.12) forces $\mathcal{C}_n$ to be $\mathbb{F}_{q^{2n}}$-maximal. We will lead up to this proof in three claims, after stating some polynomial identities in characteristic p that are important to the proof. In the following facts, let $\mathbb{F}$ be a field of characteristic p, with q a power of p.

Fact 5. *Let $n = 2k+1 \geq 1$ for $0 \leq k \in \mathbb{Z}$. Let $S := y^{q^2} - y \in \mathbb{F}[y]$. Then*

$$\begin{aligned} \sum_{i=1}^{k} S^{q^n+q^{2i}} &= y^{q^{n+2}+q^{n+1}} - y^{q^{n+2}+q^2} - y^{q^{n+1}+q^n} + y^{q^n+q^2} \\ \sum_{i=0}^{k} S^{1+q^{2i+1}} &= y^{q^{n+2}+q^2} - y^{q^{n+2}+1} - y^{q^2+q} + y^{q+1}. \end{aligned}$$

Garcia, Guneri, and Stichtenoth prove this fact by simple induction on n.

Fact 6. *If $i, j \in \mathbb{Z}$ are not congruent modulo 2, then $q+1$ divides $q^i + q^j$.*

Fact 7. *Let $n > 1$ be an odd integer of the form $n = 3m + r$ for some $r \in \{0, 1, -1\}$, $m \geq 1$. Define*

$$T := y^{q+1} \qquad S := y^{q^2} - y \qquad T_n := T^{\frac{q^n+q^2}{q+1}} - T^q + T$$
$$B_n := T^{q^{n-1}} - T^{n-2} + ... - T^q + T \qquad Q_n := \sum_{j=0}^{m-1} (-1)^{r+j} (S^{q+1})^{q^{r+3j}}.$$

Then

$$B_n - Q_n - T_n = P_n,$$

where P_n is a polynomial in $\mathbb{F}[S^{q+1}]$ with coefficients in $\{0, 1, -1\}$.

This fact is proven by induction on n, with separate base cases for each value of n modulo 3. Facts 6 and 7 are also employed.

Now for the claims that make up the proof of maximality.

Claim 6. *The degree of the map π' is q.*

Proof. To see this, consider a point $(y_0, z_0) \in \mathcal{X}_n$, with $y_0 \neq 0$. The degree of the map will be the number of points $(x_0, y_0) \in \mathcal{H}_q$ so that $\psi((x_0, y_0)) = \phi((y_0, z_0)) = y_0$. This is the number of values for x so that $x + x^q = y_0$. Since $\mathrm{Tr}_{\mathbb{F}_{q^2}/\mathbb{F}_q}(x) = x + x^q$ is a separable, degree q morphism, we have q such x values. Thus $\deg(\pi') = q$. □

Claim 7. *Every $\mathbb{F}_{q^{2n}}$-rational affine point of $\mathcal{X}_n$ splits completely in $\mathcal{C}_n$ in the covering π'.*

Proof. Since we are first concerned with the affine points of $\mathcal{C}_n$, let (x, y, z) be a $\mathbb{F}_{q^{2n}}$-rational affine point of $\mathcal{C}_n$, meaning that $x, y, z \in \mathbb{F}_{q^{2n}}$ satisfy the affine equations given in (9) and (10). Then let $\epsilon \in \mathbb{F}_{q^2}$ be an element of trace zero, i.e. $\epsilon^q + \epsilon = 0$. Notice that $(x + \epsilon, y, z)$ is also an $\mathbb{F}_{q^{2n}}$-rational point of $\mathcal{C}_n$, simply because

$$(x + \epsilon)^q + x + \epsilon = x^q + x + \epsilon^q + \epsilon = x^q + x.$$

Also notice that for affine points, the projection π' maps (x, y, z) to (y, z), so if (α, β, γ) is a point of $\mathcal{C}_n$, then (β, γ) is a point of $\mathcal{X}_n$. So let $(\alpha, \beta, \gamma) \in \mathcal{C}_n(\overline{\mathbb{F}_q})$ with $\beta, \gamma \in \mathbb{F}_{q^{2n}}$. If we can show that this implies $\alpha \in \mathbb{F}_{q^{2n}}$ also, that means $(\beta, \gamma) \in \mathcal{X}_n(\mathbb{F}_{q^{2n}})$ splits completely in $\mathcal{C}_n(\mathbb{F}_{q^{2n}})$.

This is where we will use facts 7 and 8. Consider $(x, y, z) = (\alpha, \beta, \gamma)$. Then in terms of α, β, and γ, we have

$$T = \beta^{q+1}, S = \beta^{q^2} - \beta,$$

and T_n, B_n, and Q_n can be rewritten as functions of β. Then

$$\begin{aligned} B_n^{q^n} - B_n &= (T^{q^{n-1}} - T^{q^{n-2}} + ... - T^q + T)^{q^n} - (T^{q^{n-1}} - T^{q^{n-2}} + ... - T^q + T) \\ &= T^{q^{2n-1}} - T^{q^{2n-2}} + ... - T^{q^{n+1}} + T^{q^n} - T^{q^{n-1}} + T^{q^{n-2}} - ... + T^q - T \\ &= \Sigma_{i=0}^{2n-1}(-1)^{i+1}T^{q^i}. \end{aligned}$$

Since $T = \beta^{q+1} = \alpha^q + \alpha$ by (9), this becomes

$$\begin{aligned} B_n^{q^n} - B_n &= \Sigma_{i=0}^{2n-1}(-1)^{i+1}(\alpha^q + \alpha)^{q^i} \\ &= \Sigma_{i=0}^{2n-1}(-1)^{i+1}(\alpha^{q^{i+1}} + \alpha^{q^i}) \\ &= -\alpha - \alpha^q + \alpha^q + \alpha^{q+1} - \alpha^{q+1} - \alpha^{q+2} + ... + \alpha^{q^{2n-1}} + \alpha^{q^{2n}} \\ &= \alpha^{q^{2n}} - \alpha. \end{aligned}$$

Notice that $\alpha \in \mathbb{F}_{q^{2n}}$ if and only if $\alpha^{q^{2n}} - \alpha = 0$, meaning $B_n^{q^n} - B_n = 0$, which happens if and only if $B_n \in \mathbb{F}_{q^n}$. So we wish to prove that $B_n \in \mathbb{F}_{q^n}$.

If $S = \beta^{q^2} - \beta = 0$, then $\beta \in \mathbb{F}_{q^2}$, and $\beta^{q+1} = \alpha^q + \alpha$ means $\alpha \in \mathbb{F}_{q^2} \subset \mathbb{F}_{q^{2n}}$. So assume $S \neq 0$.

Adapted to this situation, Hilbert's Theorem 90 states that $a \in \mathbb{F}_{q^{2n}}$ has $\mathrm{Tr}_{\mathbb{F}_{q^{2n}}/\mathbb{F}_{q^2}}(a) = a + a^{q^2} + a^{q^4} + ... + a^{q^{2(n-1)}} = 0$ if and only if $a = b^{q^2} - b$ for some $b \in \mathbb{F}_{q^{2n}}$. So since $S = \beta^{q^2} - \beta$,

$$S + S^{q^2} + S^{q^4} + ... + S^{q^{2(n-1)}} = 0.$$

Multiplying by the non-zero value S^{q^n}, we have

$$S^{q^n+1} + S^{q^n+q^2} + S^{q^n+q^4} + ... + S^{q^n+q^{2(n-1)}} = 0.$$

Since $\gamma \in \mathbb{F}_{q^{2n}}$ by assumption, $\gamma^{q^n+1} = \mathrm{N}_{\mathbb{F}_{q^{2n}}/\mathbb{F}_{q^n}}(\gamma) \in \mathbb{F}_{q^n}$. But from (10), $S^{q+1} = (\beta^{q^2} - \beta)^{q+1} = \gamma^{q^n+1}$. So $S^{q+1} \in \mathbb{F}_{q^n}$. By fact 7, we can then see that $S^{q^j+1} \in \mathbb{F}_{q^n}$ for any odd j. That means $S^{q^{n+j}+q^n} = (S^{q^j+1})^{q^n} = S^{q^j+1}$.

Notice that n is odd, so any even integer greater than n can be written as $n + j$ for some positive odd integer j. Let $n = 2k + 1$. Now we can write

$$\begin{aligned}
0 &= S^{q^n+1} + S^{q^n+q^2} + S^{q^n+q^4} + ... + S^{q^n+q^{2(n-1)}} \\
&= S^{q^n+1} + S^{q^n+q^2} + S^{q^n+q^4} + ... + S^{q^n+q^{n-1}} + S^{1+q} + S^{1+q^3} + ... + S^{1+q^{n-2}} \\
&= (S^{q^n+q^2} + S^{q^n+q^4} + ... + S^{q^n+q^{n-1}}) + (S^{1+q} + S^{1+q^3} + ... + S^{1+q^{n-2}} + S^{1+q^n}) \\
&= \sum_{i=1}^{k} S^{q^n+q^{2i}} + \sum_{i=0}^{k} S^{1+q^{2i+1}}.
\end{aligned}$$

By fact 6,

$$\begin{aligned}
0 &= \beta^{q^{n+2}+q^{n+1}} - \beta^{q^{n+2}+q^2} - \beta^{q^{n+1}+q^n} + \beta^{q^n+q^2} + \beta^{q^{n+2}+q^2} - \beta^{q^{n+2}+1} - \beta^{q^2+q} + \beta^{q+1} \\
&= \beta^{q^{n+2}+q^{n+1}} - \beta^{q^{n+1}+q^n} + \beta^{q^n+q^2} - \beta^{q^{n+2}+1} - \beta^{q^2+q} + \beta^{q+1}.
\end{aligned}$$

Since $\beta \in \mathbb{F}_{q^{2n}}$, so $\beta^{q^{2n}} = \beta$, we can write $\beta^{q^{n+2}+1} = \beta^{q^{n+2}+q^{2n}}$. That means

$$\begin{aligned}
0 &= (\beta^{q^2+q} - \beta^{q+1} - \beta^{q^2+q^n})^{q^n} - (\beta^{q^2+q} - \beta^{q+1} - \beta^{q^2+q^n}) \\
&= ((\beta^{q+1})^q - \beta^{q+1} - (\beta^{q+1})^{\frac{q^2+q^n}{q+1}})^{q^n} - ((\beta^{q+1})^q - \beta^{q+1} - (\beta^{q+1})^{\frac{q^2+q^n}{q+1}}) \\
&= -(T_n(\beta^{q+1}))^{q^n} + T_n(\beta^{q+1}).
\end{aligned}$$

So $T_n(\beta^{q+1})$ is an element of $\mathbb{F}_{q^n}$.

Now consider that Q_n and P_n are polynomials in $S^{q+1} \in \mathbb{F}_{q^n}$ with coefficients in $\mathbb{F}_{q^n}$, and fact 6 tells us that

$$B_n = Q_n + T_n + P_n.$$

That means that $B_n \in \mathbb{F}_{q^n}$, so every $\mathbb{F}_{q^{2n}}$-rational affine point of $\mathcal{X}_n$ splits completely in $\mathcal{C}_n$. □

Claim 8. *The genus of* $\mathcal{C}_n$ *is* $\frac{(q-1)(q^{n+1}+q^n-q^2)}{2}$.

We discovered when proving proposition 1 that p divides the index of ramification at infinity in the covering $\mathcal{C}_n \to \mathcal{X}_n$. This is called wild ramification, and it greatly complicates our problem of calculating genus. Garcia, Guneri, and Stichtenoth do calculate the genus using a fairly specialized theorem from Stichtenoth [45]. However, if all ramification in a covering is non-wild (tame), we can use a simple form of the Riemann-Hurwitz to calculate the genus of one curve if we know the genus of the other and the index of ramification at each point. With this in mind, we can calculate the genus using the other covering map, $\pi : \mathcal{C}_n \to \mathcal{H}_q$, which is only tamely ramified.

Proof. Consider the following covering map:

$$\begin{array}{cc} \mathcal{C}_n & (x,y,z) \\ \pi \downarrow & \downarrow \\ \mathcal{H}_q & (x,y) \\ \phi \downarrow & \downarrow \\ \mathbb{P}_1 & y \end{array}$$

For $P_i \in \mathcal{C}_n$, let e_i be the index of ramification in the covering $\pi : \mathcal{C}_n \to \mathcal{H}_q$. The Riemann-Hurwitz formula states that

$$2(g(\mathcal{C}_n)) - 2 = \deg(\pi)(2g(\mathcal{H}_q) - 2) + \sum_{P_i \in \mathcal{H}_q} (e_i - 1).$$

First, note that $\deg(\pi) = \frac{q^n+1}{q+1}$, since (10) means that there are $\frac{q^n+1}{q+1}$ values of z corresponding to a single y value. Ramification occurs when the fibers above points in $\mathcal{H}_q$ have fewer points than this degree. For this covering, this happens at ∞ and when $y^{q^2} - y = 0$, so for all $y \in \mathbb{F}_{q^2}$. There are q values of x satisfying (9) for each of these y values, resulting in a total of q^3 points of ramification besides ∞. Each of these points (x, y) has the single point $(x, y, 0)$ lying above it in $\mathcal{C}_n$, so $e_i = \deg(\pi) = \frac{q^n+1}{q+1}$ for each of these points. The point ∞ on $\mathcal{H}_q$ must also be fully ramified since there is a single point above it on $\mathcal{C}_n$. Using this, and the genus of $\mathcal{H}_q$ determined earlier, we have:

$$\begin{aligned}
2g(\mathcal{C}_n) - 2 &= \frac{q^n+1}{q+1}(2\frac{q(q-1)}{2} - 2) + (q^3+1)(\frac{q^n+1}{q+1} - 1) \\
&= \frac{q^n+1}{q+1}(q^3+q^2-q-1) - (q^3+1) \\
2g(\mathcal{C}_n) &= \frac{q^n+1}{q+1}(q^3+q^2-q-1) - (q^3-1) \\
&= (q^n+1)(q^2-1) - (q^3-1) \\
&= q^{n+2} - q^n + q^2 - 1 - q^3 + 1 \\
&= (q-1)(q^{n+1} + q^n - q^2) \\
g(\mathcal{C}_n) &= \frac{(q-1)(q^{n+1}+q^n-q^2)}{2}.
\end{aligned}$$

□

These claims lead to the following theorem.

Theorem 10. *The curve $\mathcal{C}_n$ is maximal over $\mathbb{F}_{q^{2n}}$.*

Proof. Since $\mathcal{X}_n$ is maximal, we know that

$$\#\mathcal{X}_n(\mathbb{F}_{q^{2n}}) = q^{2n} + 1 + q^n(q-1)(q^n-1) = q^{2n+1} - q^{n+2} + q^{n+1} + 1.$$

Now $\mathcal{C}_n$ is maximal if and only if

$$\#\mathcal{C}_n(\mathbb{F}_{q^{2n}}) = q^{2n} + 1 + q^n(q-1)(q^{n+1}+q^n-q^2) = q^{2n+2} - q^{n+3} + q^{n+2} + 1.$$

This amounts to the equality

$$(\#\mathcal{C}_n(\mathbb{F}_{q^{2n}}) - 1) + 1 = q(\#\mathcal{X}_n(\mathbb{F}_{q^{2n}}) - 1) + 1.$$

So, since every $\mathbb{F}_{q^{2n}}$-point of $\mathcal{X}_n$ except the single point at infinity splits completely (has q preimages) in $\mathcal{C}_n$, and the point at infinity ramifies completely (has 1 preimage) in $\mathcal{C}_n$ by proposition 3, this equality holds and $\mathcal{C}_n$ is $\mathbb{F}_{q^{2n}}$-maximal. □

6.11 Ramification in coverings of quotient curves

We can draw the following diagram of quotient curves of $\mathcal{C}_n$:

$$\begin{array}{cccccc}
 & & q & & q^2 & \\
 & \mathcal{C}_n & \rightarrow_{(1)} & \mathcal{X}_n & \rightarrow_{(2)} & \mathbb{P}^1_z \\
\frac{q^n+1}{q+1} & \downarrow_{(3)} & & \downarrow_{(4)} & & \downarrow_{(5)} \\
 & \mathcal{H}_q & \rightarrow_{(6)} & \mathbb{P}^1_y & \rightarrow_{(7)} & \mathbb{P}^1_t \\
q+1 & \downarrow_{(8)} & & \downarrow_{(9)} & \not\square & \downarrow_{(10)} \\
 & \mathbb{P}^1_x & \rightarrow_{(11)} & \mathbb{P}^1_w & \rightsquigarrow_{(12)} & \mathbb{P}^1_s \\
q-1 & \downarrow_{(13)} & & & & \\
 & \mathbb{P}^1_u & & & &
\end{array}$$

The numbers above the right arrows and to the left of down arrows correspond to the degrees of the coverings. We can understand the ramification of the finite places in the extensions above by examining the equations of the curves in question. To describe each covering, we will consider the associated field extensions.

- (1) and (6) are degree q Artin-Schreier covers with field extension created by adjoining the variable x subject to the relation $x^q + x = y^{q+1}$. There is no ramification at finite places in these coverings.

- (2) is a degree q^2 Artin-Schreier cover with field extension created by adjoining the variable y subject to the relation $y^{q^2} - y = z^{\frac{q^n+1}{q+1}}$. There is no ramification at finite places in this covering.

- (11) is a degree q Artin-Schreier cover with field extension created by adjoining the variable x subject to the relation $x^q + x = w$. There is no ramification at finite places in this covering.

- (7) is a degree q^2 Artin-Schreier cover with field extension created by adjoining the variable y subject to the relation $y^{q^2} - y = t$. There is no ramification above finite places in this covering.

- (3) and (4) are degree $\frac{q^n+1}{q+1}$ Kummer covers with field extensions created by adjoining the variable z subject to the relation $y^{q^2} - y = z^{\frac{q^n+1}{q+1}}$). These covers are fully ramified above all $\alpha \in \mathbb{F}_{q^2}$ and at no other finite places.

- (8) is a degree $q+1$ Kummer cover with field extension created by adjoining the variable y subject to the relation $x^q + x = y^{q+1}$. This cover is fully ramified above $x \in \mathbb{F}_{q^2}$ with $\mathrm{Tr}_{\mathbb{F}_{q^2}/\mathbb{F}_q}(x) = 0$ and at no other finite places.

- (9) is a degree $q+1$ Kummer cover with field extension created by adjoining the variable y subject to the relation $y^{q+1} = w$. This is fully ramified above $w = 0$ and at no other finite places.

- (10) must be a cyclic cover of degree $q+1$, so is a Kummer cover and hence Galois. Here the field extension is created by adjoining the variable t subject to the relation $t^{q+1} - s$). This is fully ramified above $s = 0$ and at no other finite places.

- (12) is a non-Galois cover of degree q^2. The equation for this cover is a little trickier, but using the relations $s = t^{q+1}$, $y^{q^2} - y = t$, and $y^{q+1} = w$, we find that $w(w^{q-1}-1)^{q+1} = s$. So this cover has $L = K(w)/(w(w^{q-1}-1)^{q+1} - s)$. This map is ramified above $s = 0$ and unramified at other finite places. We can see this is true by letting $s = c \in \overline{\mathbb{F}_q}$. The number of places above c in the cover is the number of distinct roots of $w(w^{q-1}-1)^{q+1} - c$. Let $f_c(w) = w(w^{q-1}-1)^{q+1} - c$. Then $f_c'(w) = -w^{q^2-q} + 1$, which has roots at $w = \zeta_{q-1}^i$ for ζ_{q-1} a primitive $q-1$st root of unity and $1 \leq i \leq q-1$. These are roots of $f_c(w)$ if and only if $c = 0$. Thus $f_c(w)$ has repeated roots if and only if $c = 0$. There are q roots of $f_0(w)$, meaning q places lying above 0 in the cover. The ramification indices are $e(0|0) = 1$, $e(\zeta_{q-1}^i|0) = q+1$.

- (13) is a cyclic Galois cover of degree $q-1$. The field extension is $\mathbb{K}(x)/\mathbb{K}(t)$ where $u = x^{q-1}$.

Part II

Ihara Zeta Functions of Graphs

This work began at the Women In Numbers workshop at Banff in the fall of 2008. Audrey Terras and Winnie Li led a project group on Ihara zeta functions and introduced some open problems in the area. Michelle Manes and I worked together on the topic of understanding ramified coverings of graphs and whether a ramified covering of graphs might imply divisibility among the corresponding zeta functions. That work eventually grew into two papers, one in the Journal of Linear Algebra and its Applications, another in the Fields Proceedings Volume of Women In Numbers. This part of the dissertation gives background on Ihara zeta functions and describes my own individual work in the area, as well as the joint work on divisibility.

Chapter 1

Introduction

The Ihara zeta function of a graph was defined by Ihara in the 1960s [18]. It was modeled on other zeta functions in its form, an infinite product over primes, and has some analogous properties, for example convergence to a rational function. Ihara introduced the function in the context of groups, and Serre observed that the function could be interpreted in terms of graphs. This connection was first seen for the finite regular graphs that arose as quotients of regular trees. Sunada [22] studied the zeta functions of these regular graphs in the 1980s. Soon after, Hashimoto [17] further developed the theory of zeta functions, with an emphasis on bipartite graphs. Hashimoto also proved that the complexity, i.e. number of spanning trees, of a regular graph can be expressed in terms of the zeta function of the graph [17]. Northshield extended this result to include irregular graphs [31]. Bass extended earlier results to general, non-regular graphs. Terras and Stark have also studied Ihara zeta functions extensively [41] [42] [43] and explored how (unramified) maps between graphs and are reflected in the zeta functions of the graphs. Manes and I considered one notion of ramified coverings in [27].

The zeta functions of members of some well-known families of graphs have been determined. Regular graphs and bipartite graphs have been studied extensively. We will summarize some essential results on zeta functions and explore some examples.

A good background source is Norman Biggs' "Algebraic Graph Theory" [4].

1.1 Background and Definitions

Let H be a finite connected graph with edge set $E(H)$ and vertex set $V(H)$. We arbitrarily orient the edges so that we can refer to directions of travel along the edges, i.e. if e is an edge with an arbitrary orientation, let e^{-1} denote that edge with opposite orientation. Let $P = a_1 a_2 ... a_n$ be a path in H, where $a_i \in E(H)$ for each i. A path is closed if it starts and ends at the same vertex. A path is said to contain a backtrack if $a_i = a_{i+1}^{-1}$ for some i. A closed path is said to contain a tail if $a_1 = a_n^{-1}$. Paths are considered equivalent if they are the same cycle of edges with a different starting point. If C is a closed path in H, let $[C]$ denote the equivalence class of C in H. A prime path in H is a tailless, backtrackless, closed path C such that $C \neq D^s$ for any positive integer s and any path D in H. In other words, a prime is a closed path without any backtracking that is not simply another path traced several times. A prime in a graph is an equivalence class $[C]$ of prime paths in the graph. The length of a prime $[C]$ is the number of edges in any representative of the class and is denoted $\nu(C)$.

Definition 3. *Let $u \in \mathbb{C}$ with $|u|$ sufficiently small. The Ihara zeta function of a finite connected graph H is*

$$\zeta_H(u) := \prod_{[P] \text{ primes in } H} (1 - u^{\nu(P)})^{-1}.$$

To calculate the zeta function for anything but an extremely simple graph we will use the following, known as Ihara's formula. The adjacency matrix of a graph H with $|V(H)| = m$ is an $m \times m$ matrix where the $(i,j)^{\text{th}}$ entry is 1 if vertex i is adjacent to vertex j and 0 otherwise. Also, the rank of the fundamental group of the finite, connected graph H with no vertices of degree 1 is equal to $|E(H)| - |V(H)| + 1$.

Theorem 11. *Let A be the adjacency matrix of a finite connected graph H which has no vertices of degree 1. Let I be the $m \times m$ identity matrix. Let Q be the diagonal matrix with the $(j,j)^{th}$ entry equal to one less than the degree of vertex j. Let r be the rank of the fundamental group of H. Then*

$$\zeta_H^{-1}(u) = (1-u^2)^{r-1} det(I - Au + Qu^2). \tag{1.1}$$

This result is due to Ihara for regular graphs, Hashimoto for semi-regular bipartite graphs, and Bass for general finite, connected graphs. It provides a powerful tool for calculating zeta functions.

1.2 Example: The Platonic Solids

The platonic solids make pretty examples of graphs and their zeta functions. We create a graph G from each platonic solid P by letting the vertex set and edge set of G be the vertices and edges of P.

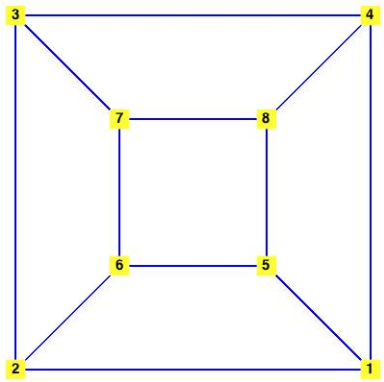
The graph of the tetrahedron is K_4, the complete graph on 4 vertices. It has adjacency matrix

$$A_{\text{tetra}} = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}.$$

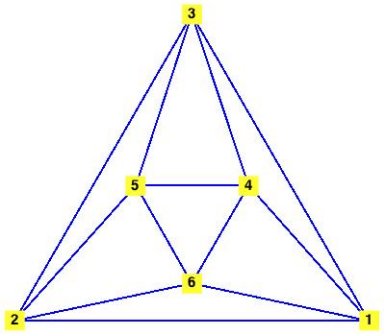
Every vertex has degree 3, so

$$Q_{\text{tetra}} = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}.$$

The rank r of the fundamental group for the tetrahedron is $r = 6 - 4 + 1 = 3$.



(a) The cube



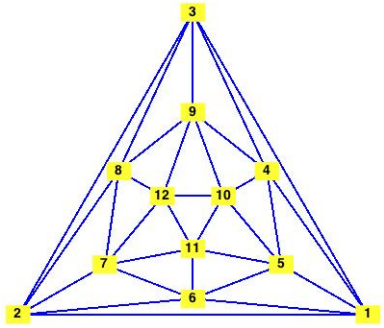
(b) The octahedron

Therefore

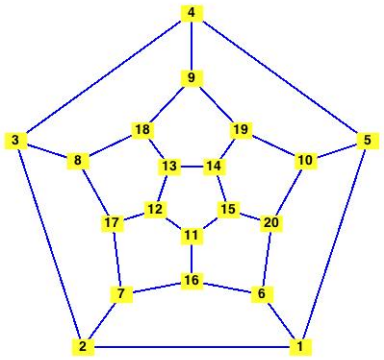
$$\zeta_{\text{tetra}}^{-1}(u) = (1-u^2)^2 \det \begin{pmatrix} 1+2u^2 & -u & -u & -u \\ -u & 1+2u^2 & -u & -u \\ -u & -u & 1+2u^2 & -u \\ -u & -u & -u & 1+2u^2 \end{pmatrix}$$
$$= (1-u^2)^2(u-1)\,(2u-1)\left(2u^2+u+1\right)^3.$$

The zeta functions for the cube, octahedron, icosahedron, and dodecahedron can be obtained similarly (matrices for the icosahedron and dodecahedron are omitted):

$$\zeta_{\text{cube}}^{-1} = (1-u^2)^4$$
$$\cdot \det \begin{pmatrix} 1+2u^2 & -u & 0 & -u & 0 & -u & 0 & 0 \\ -u & 1+2u^2 & -u & 0 & 0 & 0 & -u & 0 \\ 0 & -u & 1+2u^2 & -u & 0 & 0 & 0 & -u \\ -u & 0 & -u & 1+2u^2 & -u & 0 & 0 & 0 \\ 0 & 0 & 0 & -u & 1+2u^2 & -u & 0 & -u \\ -u & 0 & 0 & 0 & -u & 1+2u^2 & -u & 0 \\ 0 & -u & 0 & 0 & 0 & -u & 1+2u^2 & -u \\ 0 & 0 & -u & 0 & -u & 0 & -u & 1+2u^2 \end{pmatrix}$$
$$= (1-u^2)^5(2u+1)(2u-1)(2u^2-u+1)^3(2u^2+u+1)^3.$$



(a) The icosahedron



(b) The dodecahedron

$$\zeta_{\text{octa}}^{-1} = (1-u^2)^6 \det \begin{pmatrix} 1+3u^2 & -u & -u & -u & 0 & -u \\ -u & 1+3u^2 & -u & 0 & -u & -u \\ -u & -u & 1+3u^2 & -u & -u & 0 \\ -u & 0 & -u & 1+3u^2 & -u & -u \\ 0 & -u & -u & -u & 1+3u^2 & -u \\ -u & -u & 0 & -u & -u & 1+3u^2 \end{pmatrix}$$

$$= (1-u^2)^6(3u-1)(u-1)(3u^2+2u+1)^2(1+3u^2)^3.$$

$$\zeta_{\text{icos}}^{-1} = (1-u^2)^{18}(4u-1)(u-1)(1+3u^2+16u^4)^3(4u^2+u+1)^5.$$

$$\zeta_{\text{dod}}^{-1} = (1-u^2)^{10}(2u-1)(u-1)(1-u^2+4u^4)^3(1+2u^2)^4(2u^2+2u+1)^4(2u^2-u+1)^5.$$

Chapter 2

Regular Graphs

2.1 The General Case

In the case of regular graphs, Theorem 11 gives rise to an explicit formula for the zeta function of a graph H in terms of the eigenvalues of its adjacency matrix. For this, we need a lemma from linear algebra.

Lemma 5. *Let a and n be scalars. If M is a square matrix, with $\vec{v}$ and eigenvector of M with eigenvalue λ, then $\vec{v}$ is an eigenvector of*

$$aM + bI,$$

with eigenvalue $a\lambda + b$.

Proof. This is just a simple calculation:

$$(aM + bI)\vec{v} = aM\vec{v} + bI\vec{v} = (a\lambda + b)\vec{v}.$$

□

Proposition 15. *Let H be a regular graph with n vertices of adjacency k and adjacency matrix A. Let $\{\lambda_j \colon 1 \leq j \leq n\}$ be the multi-set of eigenvalues of A. Then*

$$\zeta_H^{-1}(u) = (1-u^2)^{r-1}\Pi_{j=1}^{n}(1 - \lambda_j u + (k-1)u^2).$$

Proof. For a regular graph of adjacency k,

$$Q = (k-1)I.$$

Therefore, if H is regular of adjacency k, we have

$$\zeta_H^{-1}(u) = (1-u^2)^{r-1}\det(I - Au + Qu^2) = (1-u^2)^{r-1}\det((1+(k-1)u^2)I - uA).$$

For any square matrix M, we know that $\det(M)$ is the product of the eigenvalues of M (with appropriate multiplicities). Lemma 5 implies that the eigenvalues of $((1+(k-1)u^2)I - uA)$ are $\{1 - \lambda_j u + (k-1)u^2 \colon\ 1 \leq j \leq n\}$, which proves the proposition.

□

This means that for any regular graph H, finding the zeta function of H is equivalent to finding the eigenvalues of its adjacency matrix. This allows us to calculate the zeta functions for many classes of regular graphs.

2.2 Example: Strongly Regular Graphs

Definition 4. *Let H be a regular graph of adjacency k, with v vertices. The graph H is strongly regular if there exist λ, μ integers such that each pair of adjacent vertices share exactly λ neighbors and each pair of nonadjacent vertices share exactly μ neighbors. Such a graph is called an srg(v, k, λ, μ).*

Properties of strongly regular graphs:

- Necessary: $\mu(v-k-1) = k(k-\lambda-1)$. However, this property is not sufficient for existence.
- Let J be the $v \times v$ matrix of all 1s. Then, for a strongly regular graph,

$$A^2 + (\mu - \lambda)A + (\mu - k)I = \mu J.$$

- A strongly regular graph has 3 distinct eigenvalues:

  1. $\lambda_1 = k$, multiplicity 1
  2. $\lambda_2 = \frac{(\lambda-\mu)+\sqrt{(\lambda-\mu)^2+4(k-\mu)}}{2}$, multiplicity $m_2 = \frac{1}{2}(v-1-\frac{2k+(v-1)(\lambda-\mu)}{\sqrt{(\lambda-\mu)^2+4(k-\mu)}})$
  3. $\lambda_3 = \frac{(\lambda-\mu)-\sqrt{(\lambda-\mu)^2+4(k-\mu)}}{2}$, multiplicity $m_3 = \frac{1}{2}(v-1+\frac{2k+(v-1)(\lambda-\mu)}{\sqrt{(\lambda-\mu)^2+4(k-\mu)}})$

  The last property, along with the proposition, implies that

$$\zeta_H^{-1}(u) = (1-u^2)^{r-1}(1-ku+(v-1)u^2)(1-\lambda_2 u+(v-1)u^2)^{m_2}(1-\lambda_3 u+(v-1)u^2)^{m_3}$$

  Examples of strongly regular graphs:

- Paley graphs. Let q be a prime power congruent to 1 modulo 4. Denote by $a_1, a_2, ..., a_q$ the elements of $\mathbb{F}_q$. Let G be the graph with vertex set $\{a_1, a_2, ...a_q\}$ and edge set $\{a_i, a_j\}$, where $i \neq j$ and $a_i - a_j$ is a square in $\mathbb{F}_q$. This is called the Paley graph for $\mathbb{F}_q$. Then G is an srg(q, $\frac{q-1}{2}$, $\frac{q-5}{4}$, $\frac{q-1}{4}$).

  This means that the adjacency matrix for G has eigenvalues $\frac{q-1}{2}$ with multiplicity 1, $\frac{-1+\sqrt{q}}{2}$ with multiplicity $\frac{q-1}{2}$, and $\frac{-1-\sqrt{q}}{2}$ with multiplicity $\frac{q-1}{2}$. So the field of definition of the eigenvalues of the adjacency matrix is $\mathbb{Q}(\sqrt{q})$. Also,

$$\begin{aligned}\zeta_G^{-1} = (1-u^2)^{\frac{q^2-5q}{4}}&(1-\frac{q-1}{2}u+(q-1)u^2)\\ &\cdot(1-\frac{-1+\sqrt{q}}{2}u+(q-1)u^2)^{\frac{q-1}{2}}(1-\frac{-1-\sqrt{q}}{2}u+(q-1)u^2)^{\frac{q-1}{2}}.\end{aligned}$$

  Consider the fields of definition of the eigenvalues of the adjacency matrix and the poles of the zeta function. Generally, we have

  1. $\mathbb{Q}(\sqrt{q^2-18q+17})$ for $1-\frac{q-1}{2}u+(q-1)u^2$.
  2. $\mathbb{Q}(\sqrt{17-2\sqrt{q}-15q})$ for $1-\frac{-1+\sqrt{q}}{2}u+(q-1)u^2$.
  3. $\mathbb{Q}(\sqrt{17+2\sqrt{q}-15q})$ for $1-\frac{-1-\sqrt{q}}{2}u+(q-1)u^2$.

For $q = 13$, we have

$$\zeta_G^{-1} = (1-u^2)^{26}(1-6u+12u^2)(1-\frac{-1+\sqrt{13}}{2}u+12u^2)^6(1-\frac{-1-\sqrt{13}}{2}u+12u^2)^6.$$

The smallest field of definition for the eigenvalues of the adjacency matrix of G is $\mathbb{Q}(\sqrt{13})$. The splitting fields of the factors of ζ_G^{-1} are:

1. $\mathbb{Q}(\sqrt{-3})$ for $(1 - 6u + 12u^2)$.
2. $\mathbb{Q}(\sqrt{-178 - 2\sqrt{13}})$ for $(1 - \frac{-1+\sqrt{13}}{2}u + 12u^2)$.
3. $\mathbb{Q}(\sqrt{-178 + 2\sqrt{13}})$ for $(1 - \frac{-1-\sqrt{13}}{2}u + 12u^2)$.

Let $s_1 = \sqrt{-178 - 2\sqrt{13}}$ and $s_2 = \sqrt{-178 + 2\sqrt{13}})$. The minimal polynomial for s_1 and is $922 + 7043x^2 + 38x^4$ (calculated in Maple). This means that $|\mathbb{Q}(s_1) : \mathbb{Q}| = 4$. However s_1 is the root of a quadratic polynomial in $\mathbb{Q}(\sqrt{13})$, so it must be that $\mathbb{Q}(\sqrt{13}) \subset \mathbb{Q}(s_1)$, and $|\mathbb{Q}(s_1) : \mathbb{Q}(\sqrt{13})| = 2$. Similarly, $|\mathbb{Q}(s_2) : \mathbb{Q}(\sqrt{13})| = 2$. The minimal polynomials of s_1 and s_2 over $\mathbb{Q}(\sqrt{13})$ are distinct quadratics, so $8 = |\mathbb{Q}(s_1, s_2) : \mathbb{Q}|$.

In fact, we have that $\mathbb{Q}(s_1, s_2) : \mathbb{Q}(\sqrt{13})$ is a biquadratic extension, since $s_1 s_2 = \sqrt{3 \cdot 659}$, not a square in $\mathbb{Q}(\sqrt{13})$. We know that $\sqrt{-3} \not\in \mathbb{Q}(\sqrt{13})$ because $\sqrt{13}$ is real. Since biquadratic extensions have exactly three intermediate subfields, we can check that $\sqrt{-3}$ is not contained in $\mathbb{Q}(s_1, s_2)$ because it is not contained in any of these three subfields. So adjoining $\sqrt{-3}$ to $\mathbb{Q}(\sqrt{13})$ is a quadratic extension. We can therefore see that $\sqrt{-3}$ is not in $\mathbb{Q}(s_1)$ or $\mathbb{Q}(s_2)$ because these are also quadratic extensions of $\mathbb{Q}(\sqrt{13})$ distinct from $\mathbb{Q}(\sqrt{13})(\sqrt{-3})$. The only thing that we need to check is that $\sqrt{-3}$ is not in $\mathbb{Q}(s_1 s_2)$, which it is not. Therefore $\mathbb{Q}(\sqrt{-3}) \cap \mathbb{Q}(s_1, s_2) = \mathbb{Q}$, so $|\mathbb{Q}(s_1, s_2, \sqrt{-3}) : \mathbb{Q}| = 16$.

- The complete balanced bipartite graph $B_{n,n}$ is an $\mathrm{srg}(2n, n, 0, n)$. This means that its zeta function is

$$\zeta_{B_{n,n}}^{-1}(u) = (1-u^2)^{n^2-2n+1}(1-(n-1)^2u^2)(1+(n-1)u^2)^{2n-2}.$$

Chapter 3

Bipartite Graphs and Extensions

The easiest class of irregular graphs for which the zeta functions can be easily determined are bipartite graphs. Hashimoto determined the zeta functions for semi-regular bipartite graphs in terms of the eigenvalues of their adjacency matrices [17]. Sato determined the zeta functions and complexities for the line graphs of semi-regular bipartite graphs [36]. Here, we prove a special case of Hashimoto's work to give a sense of why bipartite graphs have accessible zeta functions.

3.1 Example: The unbalanced complete bipartite graph, $B_{m,n}$

Let $B_{m,n}$ be the bipartite graph with $m+n$ vertices, with vertices partitioned into $P_m = \{a_1, a_2, ...a_m\}$ and $P_n = \{b_1, b_2, ...b_n\}$, and edges $\{(a_i, b_j)$ for all $i, j \in \mathbb{Z}$, $1 \leq i \leq m$ and $1 \leq j \leq n\}$.

Proposition 16. *The zeta function for the complete bipartite graph is given by:*

$$\zeta^{-1}_{B_{m,n}}(u) = (1-u^2)^{mn-(m+n)+1}(1+(n-1)u^2)^{m-1}(1+(m-1)u^2)^{n-1}(1-(m-1)(n-1)u^2).$$

Proof. To create the matrices A and Q for $B_{m,n}$, order the vertices $a_1, a_2, ...a_m, b_1, b_2, ...b_n$. Let $J_{i,j}$ be the $i \times j$ matrix consisting of all ones. Let $\mathbf{0}$ be the appropriately sized

matrix of all zeros. We represent A and Q as block matrices:

$$A_{B_{m,n}} = \begin{pmatrix} \mathbf{0} & J_{m,n} \\ J_{n,m} & \mathbf{0} \end{pmatrix}, \qquad Q_{B_{m,n}} = \begin{pmatrix} (n-1)I_m & \mathbf{0} \\ \mathbf{0} & (m-1)I_n \end{pmatrix}.$$

So

$$\zeta_H^{-1}(u) = (1-u^2)^{r-1}\det(I_{|V|} - Au + Qu^2)$$

$$\zeta_{B_{m,n}}^{-1}(u) = (1-u^2)^{mn-(m+n)}\det(I_{m+n} - Au + Qu^2)$$

Let $Z_H := (I - Au + Qu^2)$. Then $Z_{B_{m,n}}$ is the block matrix $\begin{pmatrix} U & V \\ V^T & W \end{pmatrix}$, where

$$U = (1 + (n-1)u^2)I_m,$$

$$W = (1 + (m-1)u^2)I_n,$$

and V is the $m \times n$ matrix with each entry equal to $-u$.

For a square block matrix $M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ where A is invertible, it is known that

$$\det(M) = \det(A)\det(D - CA^{-1}B).$$

In our case, this means that

$$\det(Z_{B_{m,n}}) = \det(U)\det(W - V^T U^{-1} V)$$

Calculation shows that $W - V^T U^{-1} V$ is the $n \times n$ matrix with each diagonal entry equal to

$$t_1 = 1 + (m-1)u^2 - \frac{mu^2}{1 + (n-1)u^2}$$

and all other entries equal to

$$t_2 = -\frac{mu^2}{1 + (n-1)u^2}.$$

So

$$\begin{aligned}\zeta^{-1}_{B_{m,n}}(u) =&(1-u^2)^{mn-(m+n)}\det(Z_{B_{m,n}})\\ =&(1-u^2)^{mn-(m+n)}(1+(n-1)u^2)^m(\det((1+(m-1)u^2)I_n - t_2 J_{n,n}).\end{aligned}$$

Let $A_{K_n} = J_{n,n} - I_n$ be the adjacency matrix for the complete graph on n vertices.

$$\zeta^{-1}_{B_{m,n}}(u) = (1-u^2)^{mn-(m+n)}(1+(n-1)u^2)^m\det((1+(m-1)u^2)I_n - t_2(A_{K_n}+I_n)).$$

It is well known that the eigenvalues of A_{K_n} are $n-1$ with multiplicity 1 and -1 with multiplicity $n-1$. So the eigenvalues of $(1+(m-1)u^2)I_n - t_2(A_{K_n}+I_n)$ are

- $(1+(m-1)u^2) - t_2((n-1)+1)$ with multiplicity 1
- $(1+(m-1)u^2) - t_2(-1+1) = (1+(m-1)u^2)$ with multiplicity $n-1$.

Therefore

$$\begin{aligned}\zeta^{-1}_{B_{m,n}}(u) =&(1-u^2)^{mn-(m+n)}(1+(n-1)u^2)^m\\ &((1+(m-1)u^2) - \frac{mu^2}{1+(n-1)u^2})(n)(1+(m-1)u^2)^{n-1}\\ =&(1-u^2)^{mn-(m+n)+1}(1+(n-1)u^2)^{m-1}(1+(m-1)u^2)^{n-1}(1-(m-1)(n-1)u^2).\end{aligned}$$

□

We get another class of examples by extending this method. When finding our determinant, we notice that it can be rewritten using the adjacency matrix of $B_{n,n}$ instead of K_n.

3.2 Example: The partially balanced complete tripartite graph, $T_{m,n,n}$

Let $T_{m,n,n}$ be a graph with $m+2n$ vertices, partitioned into 3 parts, $A = \{a_1, a_2, ..., a_m\}$, $B = \{b_1, b_2, ..., b_n\}$, and $B' = \{b'_1, b'_2, ..., b'_n\}$. The edges of $T_{m,n,n}$ are $\{(a_i, b_j), (a_i, b'_k), (b_j, b'_k)\} : 1 \leq i \leq m,\ 1 \leq j, k \leq n\}$. This is an irregular, non-bipartite graph, so the general form of the zeta function for regular graphs and Hashimoto's work on semi-regular bipartite graphs can't be directly applied. With some linear algebra and graph theory, however, we can determine the general form for the zeta function of this type of graph.

Proposition 17. *The zeta function for the complete partially balanced tripartite graph is given by*

$$\begin{aligned}\zeta^{-1}_{T_{m,n,n}} =&(1-u^2)^{n^2+2mn-2m-n}(1+(2n-1)u^2)^{m-1}(1-nu+(m+n-1)u^2)(1+(2n-1)u^2)-2nmu\\ &(1+nu+(m+n+1)u^2)(1+(m+n-1)u^2)^{2n-2}.\end{aligned}$$

Proof. The adjacency matrix of $T_{m,n,n}$ is the block matrix

$$A_{T_{m,n,n}} = \begin{pmatrix} 0 & J_{m,n} & J_{m,n} \\ J_{n,m} & 0 & J_{n,n} \\ J_{n,m} & J_{n,n} & 0 \end{pmatrix}$$

The matrix Q is:

$$Q_{T_{m,n,n}} = \begin{pmatrix} (2n-1)I_m & 0 & 0 \\ 0 & (m+n-1)I_n & 0 \\ 0 & 0 & (m+n-1)I_n \end{pmatrix}.$$

So we have that

$$\begin{aligned}\zeta^{-1}_{T_{m,n,n}}(u) &= (1-u^2)^{r-1}\det(I - Au + Qu^2) \\ &= (1-u^2)^{n^2+2mn-2m-n} \\ &\quad \det\begin{pmatrix} (1+(2n-1)u^2)I_m & -uJ_{m,n} & -uJ_{m,n} \\ -uJ_{n,m} & (1+(m+n-1)u^2)I_n & -uJ_{n,n} \\ -uJ_{n,n} & -uJ_{n,n} & (1+(m+n-1)u^2)I_n \end{pmatrix}.\end{aligned}$$

Let

$$Z_{T_{m,n,n}} = \begin{pmatrix} (1+(2n-1)u^2)I_m & -uJ_{m,n} & -uJ_{m,n} \\ -uJ_{n,m} & (1+(m+n-1)u^2)I_n & -uJ_{n,n} \\ -uJ_{n,n} & -uJ_{n,n} & (1+(m+n-1)u^2)I_n \end{pmatrix}.$$

Apply the same result on the determinants of block matrices used above, that is, that

$$\det\begin{pmatrix} U & V \\ V^T & W \end{pmatrix} = \det(U)\det(W - V^T U^{-1} V).$$

Let $U = (1+(2n-1)u^2)I_m$, from which the definitions of W and V follow. So

$$\begin{aligned}\det(Z_{T_{m,n,n}}) = (1+(2n-1)u^2)^m \det(&\begin{pmatrix} (1+(m+n-1)u^2)I_n & -uJ_{n,n} \\ -uJ_{n,n} & (1+(m+n-1)u^2)I_n \end{pmatrix} \\ &- (\frac{mu^2}{1+(2n-1)u^2} J_{2n,2n})).\end{aligned}$$

Let $A_{B_{n,n}} = \begin{pmatrix} 0 & J_{n,n} \\ J_{n,n} & 0 \end{pmatrix}$, the adjacency matrix of a complete balanced bipartite graph. The eigenvalues of $A_{B_{n,n}}$ are n with multiplicity 1, $-n$ with multiplicity 1, and 0 with multiplicity $2n-2$. Notice that $A^2_{B_{n,n}} + nA_{B_{n,n}} = nJ_{2n,2n}$, so $J_{2n,2n} = \frac{1}{n}A^2_{B_{n,n}} + A_{B_{n,n}}$. That means we can rewrite the above:

$$
\begin{aligned}
\det(Z_{T_{m,n,n}}) =&(1+(2n-1)u^2)^m \det((1+(m+n-1)u^2)I_{2n} \\
&\quad - uA_{B_{n,n}} - \frac{mu^2}{1+(2n-1)u^2}(\frac{1}{n}A_{B_{n,n}}^2 + A_{B_{n,n}})) \\
=&(1+(2n-1)u^2)^m \\
&\quad \cdot (1+(m+n-1)u^2 - (u+\frac{mu^2}{1+(2n-1)u^2})n - \frac{mu^2}{1+(2n-1)u^2}(\frac{1}{n})n^2) \\
&\quad \cdot (1+(m+n-1)u^2 + (u+\frac{mu^2}{1+(2n-1)u^2})n - \frac{mu^2}{1+(2n-1)u^2}(\frac{1}{n})n^2) \\
&\quad\quad \cdot (1+(m+n-1)u^2)^{2n-2} \\
=&(1+(2n-1)u^2)^m (1-nu+(m+n-1)u^2 - 2n(\frac{mu^2}{1+(2n-1)u^2})) \\
&\quad \cdot (1+nu+(m+n+1)u^2)(1+(m+n-1)u^2)^{2n-2} \\
=&(1+(2n-1)u^2)^{m-1}(1-nu+(m+n-1)u^2)(1+(2n-1)u^2) - 2nmu^2) \\
&\quad \cdot (1+nu+(m+n+1)u^2)(1+(m+n-1)u^2)^{2n-2}.
\end{aligned}
$$

This gives the above result.

□

Chapter 4

Biregular graphs and graphs with three eigenvalues

Propositions 16 and 17 use the same determinant decomposition for block matrices and rely on the transformed matrix having a form that can be expressed as a polynomial function of some other matrix for which the eigenvalues are already known. Both $B_{m,n}$ and $T_{m,n,n}$ are examples of biregular graphs, i.e. graphs with exactly two valencies. This technique could be useful for other biregular graphs, because their vertices can be ordered so that the matrix $I-uA+u^2Q$ has a natural block structure. I would like to study which biregular graphs also give rise to decompositions in terms of powers of well-known matrices, and determine the zeta functions for these graphs.

Let H be a graph with vertex set $V = \{h_1, h_2, ..., h_n\}$ and edge set E. The cone over H is the graph with vertex set $V \cup \{v\}$ and edge set $E \cup \{(h_1, v), (h_2, v), ..., (h_n, v)\}$. One could generalize this construction by considering the graph with vertex set $V \cup \{v_1, v_2, ..., v_k\}$ and edge set $E \cup_{i=1}^{k} \{(h_1, v_i), (h_2, v_i), ..., (h_n, v_i)\}$. The graph $T_{m,n,n}$ is the generalized cone created by adjoining m vertices to $B_{n,n}$. Graphs created in this way also give a natural block structure to $I - uA + u^2Q$. If H is a regular graph, the generalized cone over H is a biregular graph with a particularly simple structure. These graphs would be good first candidates to study.

Another family of graphs which could be amenable to study are irregular graphs with 3 eigenvalues. A regular graph is strongly regular exactly when it has 3 eigenvalues. The simplest example of an irregular graph with 3 eigenvalues is the complete bipartite graph $B_{m,n}$. Other irregular graphs with 3 eigenvalues can be created from some strongly regular graphs by the process of "switching" (see [30]). Let A_H be the adjacency matrix for the graph H. The adjacency algebra for H is the algebra of polynomials in A_H. Graphs with 3 eigenvalues have adjacency algebras of dimension 3, meaning all higher powers of the adjacency matrix can be expressed as linear combinations of I, A_H, and A_H^2. It is possible that this could be exploited in finding the zeta function of H, or of a graph related to H.

Muzychuk and Klin considered when a biregular graph has exactly 3 eigenvalues [30]. The cone over a strongly regular graph is biregular, and Bridges and Mena showed that the cone over a strongly regular graph with certain parameters has 3 eigenvalues [5]. These graphs are members of both of the above families, so could be good starting examples.

Chapter 5

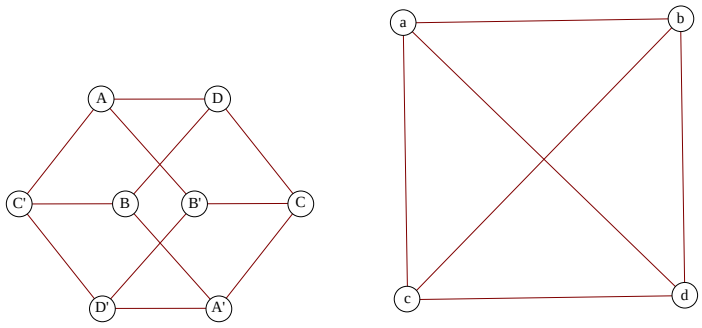
Graph Coverings

Much of the material in this section can be found in similar form in [26], including proofs in [27].

Terras and Stark considered the notion of an unramified covering of graphs. Intuitively, a covering is a surjective map $f : H \to G$ with $V(H) \mapsto V(G)$ and $E(H) \mapsto E(G)$ which respects the structure of the graph. For $a, b \in V(H)$ adjacent, let (a, b) indicate the edge from a to b. A covering map requires that if $(a, b) \in E(H)$, then $(f(a), f(b)) \in E(G)$, and $f(a, b) = (f(a), f(b))$. An unramified covering also has the property that the fiber above each edge and vertex of G has the same number of elements, known as the degree of the covering. Interesting examples of unramified coverings exist, for instance the map from the cube to the tetrahedron (see Figure 5.1).

If such an unramified covering map exists, then $\zeta_G^{-1} \mid \zeta_H^{-1}$. Ihara zeta functions of graphs have some parallels with Artin zeta functions of curves, in which we have zeta function divisibility for ramified covers as well. We wanted to know if the divisibility relation for Ihara zeta functions in an unramified cover also held for ramified coverings of graphs. We used a definition of ramified covering adopted from Urakawa (via Baker and Norine) as outlined below.

Definition 5. *Let H and G be graphs. A function $\phi : V(H) \cup E(H) \to V(G) \cup E(G)$*



(a) The cube.

(b) The complete graph on 4 vertices, also known as the tetrahedron.

Figure 5.1: The cube graph is an unramified degree 2 cover of the tetrahedron.

is called a morphism if

- $\phi(x) \in V(G)$ *for all* $x \in V(H)$*, and*
- *for every* $x \in V(H)$ *and* $e \in E(H)$ *such that* $x \in e$*, we have either* $\phi(e) \in E(G)$ *and* $\phi(x) \in \phi(e)$*, or* $\phi(e) = \phi(x)$*.*

A morphism $\phi : H \to G$ *is said to be* harmonic *if for all* $x \in V(H)$ *and* $e' \in E(G)$*, the number of edges* e *in* $E(H)$ *such that* $x \in e$ *with* $\phi(e) = e'$ *is the same for each* e' *which contains* $\phi(x)$*, i.e.* ϕ *is harmonic if for all* $x \in V(H)$*, the quantity* $|\{e \in E(H) : x \in e, \phi(e) = e'\}|$ *is constant for all* $e' \in E(G)$ *with* $\phi(x) \in e'$*.*

Notice that harmonic morphisms encompass the unramified coverings of graphs studied by Terras and Stark, but allows for the size of fibers to vary (as in ramified covers of curves). Edges in H may also collapse, i.e. be mapped to vertices of G. Urakawa [49] also proved that the following is a well-defined notion of the *degree* of a harmonic morphism.

Definition 6. *If* $\phi : G \to H$ *is a harmonic morphism of graphs with* $x \in V(G)$*, the*

vertical multiplicity of ϕ at x is given by

$$v_\phi(x) = |e \in E(G) : \phi(e) = \phi(x)| .$$

The horizontal multiplicity of ϕ at x is given by

$$m_\phi(x) = |e \in E(G) : x \in e, \phi(e) = e'|$$

for any edge $e' \in E(G)$ with $\phi(x) \in e'$. For ϕ as above, and any vertex $y \in V(H)$, the degree *of ϕ is given by*

$$deg(\phi) = \sum_{x \in V(G):\phi(x)=y} m_\phi(x).$$

Based on these definitions of multiplicity and degree, Baker and Norine [3] proved a graph-theoretic analogue of the Riemann-Hurwitz formula, as well as several other results. This indicates that harmonic morphisms could be a good choice for generalizing the idea of covering maps of graphs to include ramification. We consider the simplest version of a ramified covering, namely $k \geq 1$ copies of a fixed graph X, with a single vertex on each graph identified. See Figure 5.2 for an example. Note that this fits the definition of ramified covering given in [3].

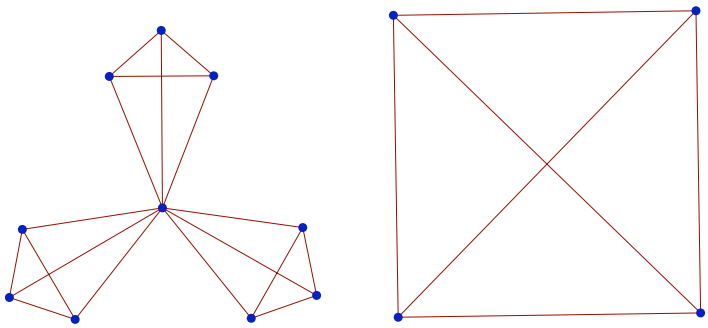
Manes and Malmskog proved a general result for the above mentioned type of covering of regular graphs [27] [26].

Theorem 12. *Let X be a finite $(q+1)$-regular graph with vertices labeled v_1, v_2, ..., v_n. Let A_X be the adjacency matrix for X, and suppose that A_X has d distinct eigenvalues. Let Y_k be the ramified cover of X created by identifying vertex v_1 on k copies of X. Then there exists a polynomial $P_{k,q}(u) \in \mathbb{Z}[u]$, with coefficients depending on k and q (and on the original graph X) and with $\deg_u P_{k,q} \leq 2d$, such that for $l \geq k$,*

$$\zeta_{Y_k}^{-1}(u)P_{l,q}(u) \text{ divides } \zeta_{Y_l}^{-1}(u)P_{k,q}(u). \tag{5.1}$$

In particular, since $Y_1 = X$, we have that for all $l \geq 1$

$$\zeta_X^{-1}(u) \text{ divides } \zeta_{Y_l}^{-1}(u)P_{1,q}(u).$$



(a) Three copies of K_4 (b) The complete graph on 4 vertices.

Figure 5.2: K_4 and three copies of the graph around a single identified vertex.

Here, the degree of the extra term $P_{1,q}(u)$ needed for divisibility is independent of both n and k, so the failure of true divisibility is controlled.

The main tools of the proof come from linear algebra. We use Ihara's formula to change the calculation of the zeta function of a graph into a determinant calculation.

In the case where the base graph X is K_n, the complete graph on n vertices, the adjacency matrices are simple enough that we can say even more. See Figure 5.2 for the $n = 4$ case.

Proposition 18. *The Ihara zeta function for K_n is*

$$\zeta_{K_n}^{-1}(u) = (1-u^2)^{n(n-3)/2}(u-1)\left((n-2)u-1\right)\left((n-2)u^2+u+1\right)^{n-1}. \tag{5.2}$$

Hashimoto also calculated the zeta function of K_n [17]. However, the following covers of K_n form a new family of examples [27] [26].

Proposition 19. *Let $X_{n,k}$ be the graph consisting of k copies of K_n identified at a single identified vertex. Then the Ihara zeta function for $X_{n,k}$ is*

$$\begin{aligned}\zeta_{X_{n,k}}^{-1}(u) = (1-u^2)^{r-1}(u-1)\left((n-2)(nk-k-1)u^3+(n-3)u-1\right)\\ \left((n-2)u^2+u+1\right)^{k(n-2)}\left((n-2)u^2-(n-2)u+1\right)^{k-1},\end{aligned} \tag{5.3}$$

where $r = k(n-1)(n-2)/2$ is the rank of the fundamental group of $X_{n,k}$.

Corollary 2. *Let $X_{n,k}$ be the graph consisting of k copies of K_n with a single vertex from each copy identified. If $i \leq j$, then $\zeta^{-1}_{X_{n,i}}(u)\left((n-2)(nj-j-1)u^3+(n-3)u-1\right)$ divides $\zeta^{-1}_{X_{n,j}}(u)\left((n-2)(ni-i-1)u^3+(n-3)u-1\right)$. In particular, for all $k \geq 1$*

$$\zeta^{-1}_{K_n}(u) \text{ divides } \zeta^{-1}_{X_{n,k}}(u)\left((n-2)^2u^3+(n-3)u-1\right).$$

Here we have a specific example of a graph K_n and a ramified cover of the graph $X_{n,k}$, where $\zeta^{-1}_{K_n}(u) \nmid \zeta^{-1}_{X_{k,n}}(u)$. We see that each term of (5.2) divides some term of equation (5.3), except for the term

$$(n-2)^2u^3+(n-3)u-1,$$

which will never divide

$$(n-2)(nk-k-1)u^3+(n-3)u-1$$

(nor will it divide any other term of $\zeta^{-1}_{X_{k,n}}(u)$ when $n > 3$).

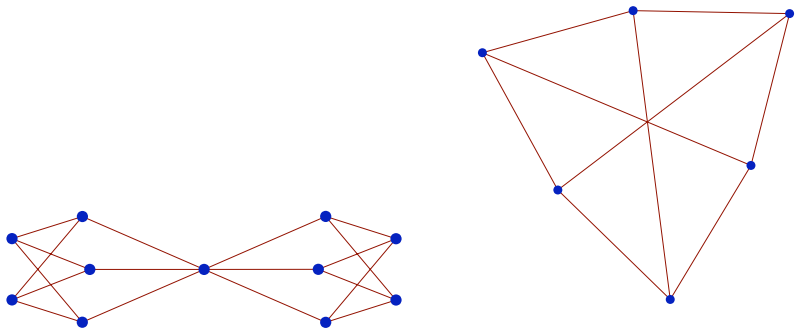
In the special case $n = 3$ we do actually have $\zeta^{-1}_{K_3}(u) \mid \zeta^{-1}_{X_{3,k}}(u)$ for every $k \geq 1$. The u^3-1 term divides into an "extra" copy of $(u-1)(u^2+u+1)$, where the first term arises from $(1-u^2)^{r-1}$, and the other arises from $(u^2+u+1)^k$.

We have already addressed the complete balanced bipartite graph, $B_{n,n}$. See Figure 5.3 for an example. Again, this was originally calculated by Hashimoto as the simplest example of a semi-regular bipartite graph, however, the zeta functions of the following covers had not previously been described [27].

Proposition 20. *Let $Y_{n,k}$ be the graph consisting of k copies of $B_{n,n}$ with a single vertex on each of the k copies identified. Then the Ihara zeta function for $Y_{n,k}$ satisfies*

$$\zeta^{-1}_{Y_{n,k}}(u) = (1-u^2)^{k(n-1)^2-1}\left((n-1)u^2+1\right)^{k(2n-3)} \left((n-1)^2u^4-(n-1)(n-2)u+1\right)^{k-1}(u^2-1)P_{k,n}(u), \text{ where} \tag{5.4}$$

$$P_{k,n}(u) = (n-1)^2(nk-1)u^4+(n-1)(n-2)u^2-1. \tag{5.5}$$



(a) Two copies of $B_{3,3}$

(b) The (3,3)-complete bipartite graph.

Figure 5.3: $B_{3,3}$ and two copies of the graph around a single identified vertex.

The complete bipartite covering again shows that "almost divisibility" is the best one may achieve for zeta functions of these special ramified covers of graphs. The polynomial $P_{k,n}(u)$ will not divide any term of $\zeta_{Y_{n,l}}(u)^{-1}$ when $k < l$. So either the parallel of graph zeta functions and curve zeta functions does not extend to include ramified coverings, or a different notion of ramified coverings must be considered. As explained briefly in section 1, I would like to investigate whether graph coverings with constant size fibers above each vertex and edge (as in the case of the cube covering the tetrahedron), for which divisibility does hold, could be understood as incorporating ramification above primes. One test for the validity of such an understanding would be to see if an analogue of the Riemann Hurwitz formula holds in this type of covering, as Baker and Norine proved for harmonic morphisms [3].

Chapter 6

Future Research Directions

- Calculate zeta functions for families of graphs. Given sufficient computing power, the zeta function for any particular finite graph can be computed directly using Ihara's determinant formula. However, finding the zeta functions for families of graphs is a different, theoretical problem. Ihara, Sunada, Hashimoto, Bass, Terras and Stark, and other current researchers have calculated zeta functions for many families of graphs; however, many have not been done.

  For regular graphs, knowledge of the zeta function is equivalent to knowledge of the eigenvalues of the graph's adjacency matrix, so these have been studied extensively. Semi-regular bipartite graphs' zeta functions have also been well studied. However, irregular graphs have not been addressed as fully. In these cases, the zeta function seems to contain information beyond the eigenvalues of the adjacency matrix of the graph. Michelle Manes and I calculated the zeta functions for two infinite families of irregular graphs [27] [26], and I have calculated the zeta function for a third, new infinite family. This work uses methods and results from linear algebra and graph theory. I am currently working to extend the technique that I used for the third family to find zeta functions for other families. Two families of particular interest (and a next step in my research) are biregular graphs and irregular graphs with 3 eigenvalues. For

references see [30], [7], [50], and [5]. A regular graph whose adjacency matrix has 3 eigenvalues is a strongly regular graph, and so well understood. The 3 eigenvalue irregular graphs have some properties in common with strongly regular graphs, so may be amenable to study.

- Determine the field of definition of poles of zeta functions of graphs. The magnitudes of the poles of graph zeta functions have been well studied. Regular graphs which obey an analogue of the Riemann Hypothesis have been shown to be Ramanujan [48]. Ramanujan graphs have applications in network theory. However, the field of definition has not been carefully considered for the poles of zeta functions of graphs. We have examples where this field of definition is a cyclotomic, quadratic, or biquadratic field (these appear when considering cycle graphs, complete graphs, and semi-regular bipartite graphs). I plan to compile data on these fields and perhaps determine what properties of the graph are linked to this field extension. What properties of a graph will guarantee an abelian extension, or that the extension has a 2-group as its Galois group?

  A related goal is to analyze the divisibility properties of the reciprocals of zeta functions under covers. If $\mathcal{Y} \to \mathcal{X}$ is a (possibly ramified) covering of curves defined over a finite field, then the zeta function of $\mathcal{X}$ divides the zeta function of $\mathcal{Y}$. Terras and Stark proved an analogous result for unramified graph coverings. Terras posed the question of what a ramified graph covering might be, and asked whether such a divisibility result might hold for such graph coverings [48]. Motivated by [3]Michelle Manes and I considered the harmonic morphism as a notion of ramified covering, and found that divisibility is not true for such coverings [27] [26]. This raises the question of whether divisibility would result if a different definition of ramified covering were adopted. Some preliminary experiments have indicated that the simplest ideas, such as ramification along an edge, are not promising. However, many other kinds of coverings could be

considered. It also seems possible that Terras and Stark's coverings could be reinterpreted as being ramified. Ramification occurs in a covering of curves when some point in the base curve has fewer points in the fiber above it than the degree of the covering. The covers that Terras and Stark considered are unramified in the sense that the fibers over each vertex and edge of the covered graph are of constant size; however, fibers above primes in the covered graph are not of constant size. I would like to investigate the possibility that this could be interpreted as ramification. The parallel between graph zeta functions and curve zeta functions would then be preserved. A good test of this reinterpretation would be whether this notion of ramification is compatible with the Riemann-Hurwitz formula.

Let ζ_G^{-1} be the reciprocal of the zeta function of a graph G. The two topics above are related because if G and H are graphs, and ζ_G^{-1} divides that of ζ_H^{-1}, then the field of definition of the poles of ζ_G is contained in the field of definition of the poles of ζ_H. This allows us to apply techniques of field theory to the problem of divisibility. Once the question of field of definition of poles is better understood, it may allow us to better understand the divisibility of reciprocal zeta functions. It is possible that ramification of graphs in some covers could be understood in terms of ramification in these field extensions.

Part III

The Zeta Function of Gauss' Curve

This paper was begun by Jeremy Muskat and submitted to the Rocky Mountain Journal of Mathematics. The referee suggested that finding the global zeta function of the singular curve would add greatly to the paper. I revised and streamlined the original paper and added Section 3 on the global zeta function.

Chapter 1

Introduction

The last entry in Gauss's mathematical diary [13] is the following conjecture.

Conjecture 1. *Suppose $p \equiv 1$ mod 4, and $a + bi \equiv 1$ mod $(2 + 2i)$ is such that $p = a^2 + b^2$. Then the number of solutions to $x^2 + y^2 + x^2y^2 = 1$ over $\mathbb{F}_p$ is $p + 1 - 2a$.*

Gauss's conjecture accounts for four points at infinity. It is interesting to note that Gauss was thinking of the curve projectively and counting the points birationally. Counting the points geometrically yields two points at infinity. Using Gauss's insight, and counting points geometrically, led to the following theorem.

Theorem 13. *[19, Chapter 11.5] Consider the curve $C : x^2t^2 + y^2t^2 + x^2y^2 - t^4 = 0$ in $\mathbb{P}^2$ defined over $\mathbb{F}_p$ where $p \equiv 1$ mod 4. Write $p = a^2 + b^2$ with b even and with $a \equiv (-1)^{b/2}$ mod $(2 + 2i)$. Then the number of points in $C(\mathbb{F}_p)$ is $N_1 = p - 1 - 2a$. Furthermore*

$$Z_C(u) = \frac{(1 - 2au + pu^2)(1 - u)}{1 - pu}.$$

The focus of this paper is an analogue for Theorem 13 for the case when $p \equiv 3$ mod 4. We give a proof that when $p \equiv 3$ mod 4 the number of points in $C(\mathbb{F}_{p^s})$ is

$$N_s(C) = \begin{cases} p^s + 3 & \text{if } 2 \nmid s; \\ p^s - 2(i\sqrt{p})^s - 1 & \text{if } 2|s. \end{cases}$$

As shown in Theorem 14, this yields the zeta function

$$Z_C(u) = \frac{(1+pu^2)(1+u)^2}{(1-pu)(1-u)}.$$

This result appears in [6], but its proof does not appear in the given reference [19, chapter 11.5].

If X is a smooth projective curve, then the Weil conjectures imply that the complex absolute value of the roots of $Z_X(u)$ is $\sqrt{p}$. Notice, for $p \equiv 3 \bmod 4$ the zeta function of C has roots with complex absolute value 1. Therefore $Z_C(u)$ does not satisfy the conclusion of the Weil conjectures, as expected since C is singular at infinity.

The method we use to determine the zeta function of C is to find a correspondence between the solutions of $x^2t^2 + y^2t^2 + x^2y^2 - t^4 = 0$ and the solutions of two other equations. The solutions to these other equations can be counted using Jacobi sums and the Weil conjectures.

In Theorem 15 we determine the global zeta function of C and relate this to the Hecke L-function associated to the normalization $\tilde{C}$ of Gauss's curve.

We would like to express our gratitude to our advisor Rachel Pries, whose technical and editorial advice was essential for the completion of this paper, and to the referee for very helpful comments.

Chapter 2

The Zeta Function of C

Definition 7. *Consider a projective plane curve X defined over $\mathbb{F}_p$. The zeta function of X is the series given by*

$$Z_X(u) = \exp\Big(\sum_{\mathrm{n=1}}^{\infty} \frac{\mathrm{N_n(X)u^n}}{\mathrm{n}}\Big) \text{ where } \mathrm{N_n(X)} \text{ denotes the size of } \mathrm{X}(\mathbb{F}_{\mathrm{p^n}}).$$

Therefore the sequence $N_n(X)$ determines the zeta function $Z_X(u)$. The converse is often true; the following explains how to reverse the process.

Fact 8. *[19, Chapter 11.1] If the zeta function of a projective plane curve X is rational, meaning $Z_X(u) = \Pi_i(1-a_iu)\Pi_j(1-b_ju)^{-1}$ for some $a_i, b_j \in \mathbb{C}$, then $N_n(X) = \sum_j b_j^n - \sum_i a_i^n$.*

We will use the following notation in this section. Let X denote the curve in $\mathbb{P}^2$ given by the zero locus of a homogeneous polynomial $F \in \mathbb{F}_p[x, y, t]$. Let X_0 represent the affine curve given by the zero locus of the polynomial $f(x, y) = F(x, y, 1)$. Let $N_n(X_0)$ denote the size of $X_0(\mathbb{F}_{p^n})$. Let p be a prime with $p \equiv 3 \bmod 4$. Let $\zeta_8 = \sqrt{2}/2 + \sqrt{2}i/2$, then $\zeta_8 \in \mathbb{F}_{p^n}$ if and only if n is even.

2.1 Near Bijections

Here, we define maps between curves which will relate the number of points on the curves. Two additional curves with similar notation will be used. We will consider

the projective curve E as well as the corresponding affine model E_0:

$$E : y^2t - x^3 + 4xt^2 = 0, \qquad E_0 : y^2 - x^3 + 4x = 0,$$

We also use the affine curve G_0:

$$G_0 : z^2 + w^4 - 1 = 0.$$

Proposition 21. *Consider the curves* $C_0 : x^2+y^2+x^2y^2-1 = 0$ *and* $G_0 : z^2+w^4-1 = 0$ *over* $\mathbb{F}_p$. *Then*

$$N_n(C_0) = \begin{cases} N_n(G_0) & \text{if } 2 \nmid n; \\ N_n(G_0) - 2 & \text{if } 2|n. \end{cases}$$

Proof. Consider the map

$$\mu : G_0(\mathbb{F}_{p^n}) \to C_0(\mathbb{F}_{p^n}) \quad \text{where} \quad (w, z) \mapsto \Big(w, \frac{z}{1+w^2}\Big)$$

The map μ is defined for all $(w, z) \in G_0(\mathbb{F}_{p^n})$ such that $w^2 \not\equiv -1 \bmod p$. Notice that if $x^2 + y^2 + x^2y^2 - 1 = 0$ then $((1 + x^2)y)^2 = 1 - x^4$. Define $\tilde{\mu} : C_0(\mathbb{F}_{p^n}) \to G_0(\mathbb{F}_{p^n})$ by $\tilde{\mu}(x, y) = (x, (1 + x^2)y)$. The maps μ and $\tilde{\mu}$ are inverses of each other. Therefore μ is a bijection for n odd and a bijection away from the points $(0, \pm\sqrt{-1}) \in G_0(\mathbb{F}_{p^n})$ for n even. Hence $N_n(C_0) = N_n(G_0)$ for n odd and $N_n(C_0) = N_n(G_0) - 2$ for n even. □

Proposition 22. *Consider the affine curve* $E_0 : y^2 - x^3 + 4x = 0$ *defined over* $\mathbb{F}_p$. *Then* $N_n(C_0) = N_n(E_0) - 3$ *for* n *even.*

Proof. Consider the following map defined over $\mathbb{F}_{p^n}$ for n even:

$$\alpha : E_0(\mathbb{F}_{p^n}) \to G_0(\mathbb{F}_{p^n}) \quad \text{where} \quad (x, y) \mapsto \Big(\frac{\zeta_8 y}{2x}, \frac{y^2 + 8x}{4x^2}\Big).$$

The map α is well defined away from $(0, 0) \in E_0(\mathbb{F}_{p^n})$ since

$$((y^2 + 8x)/4x^2)^2 - (\zeta_8 y/2x)^4 - 1 = 0.$$

Consider the following map defined over $\mathbb{F}_{p^n}$ for n even:

$$\tilde{\alpha} : G_0(\mathbb{F}_{p^n}) \to E_0(\mathbb{F}_{p^n}) \quad \text{where} \quad (w, z) \mapsto \Big(\frac{2}{z + iw^2}, \frac{4\zeta_8^7 w}{z + iw^2}\Big).$$

The map $\tilde{\alpha}$ is well defined for all points of $G_0(\mathbb{F}_{p^n})$ since there is no point (w, z) in $G_0(\mathbb{F}_{p^n})$ such that $z + iw^2 = 0$. Also

$$(4\zeta_8^7 w/(z+iw^2))^2 - (2/(z+iw^2))^3 + 4(2/(z+iw^2)) = 0.$$

The maps α and $\tilde{\alpha}$ are inverses. Therefore $\alpha : E_0(\mathbb{F}_{p^n}) - \{(0,0)\} \to G_0(\mathbb{F}_{p^n})$ is a bijection and $N_n(G_0) = N_n(E_0) - 1$ for n even. Proposition 21 proves the proposition.

□

2.2 Jacobi Sums

We use Jacobi sums as a tool to count points on G_0. The multiplicative characters of $\mathbb{F}_{p^n}^*$ form a cyclic group of order $p^n - 1$. Let $S_{m,n}$ be the set of multiplicative characters of $\mathbb{F}_{p^n}^*$ of order m. Therefore, for each $m|(p^n - 1)$ the size of $S_{m,n}$ is $\phi(m)$. Let $\chi_{m,n}$ denote one of the multiplicative characters of order m on $\mathbb{F}_{p^n}^*$. Extend $\chi_{m,n}$ to $\mathbb{F}_{p^n}$ by defining $\chi_{m,n}(0) = 0$ for $m \neq 1$ and $\chi_{1,n}(0) = 1$. For the remainder of the paper, we drop the word multiplicative and refer to $\chi_{m,n}$ as a character of $\mathbb{F}_{p^n}$.

Proposition 23. *[19, Chapter 8.2] For $a \in \mathbb{F}_{p^n}$, let $N_n(x^n = a)$ denote the number of solutions to the equation $x^n = a$ over $\mathbb{F}_{p^n}$. Then*

$$N_n(x^n = a) = \sum_{m|n} \sum_{\chi \in S_{m,n}} \chi(a)$$

where the sum is over all characters of order m dividing n.

Definition 8. *For any two characters $\chi_{m,n}$ and $\chi_{l,n}$ of $\mathbb{F}_{p^n}$, set*

$$J(\chi_{m,n}, \chi_{l,n}) = \sum_{\substack{a,b \in \mathbb{F}_{p^n} \\ a+b=1}} \chi_{m,n}(a)\chi_{l,n}(b).$$

Then we call $J(\chi_{m,n}, \chi_{l,n})$ a Jacobi sum.

Proposition 24. *[19, Chapter 8.2] $J(\chi_{1,n}, \chi_{1,n}) = p^n$, and for $m \neq 1$, $J(\chi_{m,n}, \chi_{1,n}) = 0$. For $p \equiv 3 \bmod 4$, $J(\chi_{2,n}, \chi_{2,n}) = -(\chi_{2,n}) = -(-1)^n$.*

Notice that there is only one character of order 2.

Lemma 6. *Recall that G_0 is the affine curve with equation $z^2 + w^4 - 1 = 0$ defined over $\mathbb{F}_p$. Then $N_n(G_0) = p^n + 1$ when n is odd*

Proof. For odd values of n, $p^n \equiv 3 \bmod 4$. Hence the group of characters on $\mathbb{F}_{p^n}$ does not contain a character of order 4. Proposition 23 implies that $N(x^4 = b) = N(x^2 = b)$. Therefore

$$N_n(G_0) = \sum_{\substack{a,b \in \mathbb{F}_{p^n} \\ a+b=1}} N(x^2 = a)N(x^4 = b) = \sum_{\substack{a,b \in \mathbb{F}_{p^n} \\ a+b=1}} N(x^2 = a)N(x^2 = b).$$

Using Proposition 23, 24, and Definition 8 we can simplify the above sum as follows:

$$N_n(G_0) = \sum_{\substack{a,b \in \mathbb{F}_{p^n} \\ a+b=1}} (1 + \chi_{2,n}(a))(1 + \chi_{2,n}(b)) = J(\chi_{1,n}, \chi_{1,n}) + J(\chi_{2,n}, \chi_{2,n}).$$

Thus $N_n(G_0) = p^n - (\chi_{2,n}) = p^n + 1$. □

2.3 $\mathbf{E_0 : y^2 - x^3 + 4x^2}$

Our goal for this section is to determine $N_n(E_0)$. We will use the Weil conjectures [19, Chapter 11.4] to achieve this.

The Weil conjectures imply that

$$Z_E(u) = \frac{(1 - \alpha_p u + pu^2)}{(1 - u)(1 - pu)} \quad \text{where} \quad N_1(E) = p + 1 - \alpha_p.$$

In order to completely determine $Z_E(u)$, we just need to determine $N_1(E)$.

Lemma 7. *Recall that E_0 is the affine curve with the equation $y^2t - x^3 + 4xt^2 = 0$ over $\mathbb{F}_p$. Then $N_n(E_0) = p^n - 2(i\sqrt{p})^n$ when n is even.*

Proof. The elliptic curve E has only one point $[0, 1, 0]$ at infinity. Therefore $N_1(E) = 1 + N_1(E_0)$.

By [19, Theorem 5, page 307] $N_1(E) = p + 1$ and $\alpha_p = 0$. It follows that

$$Z_E(u) = \frac{(1 + pu^2)}{(1 - u)(1 - pu)} = \frac{(1 + i\sqrt{p}u)(1 - i\sqrt{p}u)}{(1 - u)(1 - pu)}.$$

Fact 8 implies

$$N_n(E) = (1^n + p^n) - ((i\sqrt{p})^n + (-i\sqrt{p})^n).$$

Therefore when n is even, $N_n(E) = p^n - 2(i\sqrt{p})^n + 1$ and $N_n(E_0) = p^n - 2(i\sqrt{p})^n$. $\square$

2.4 The Zeta Function for C

In this section we find the zeta function of Gauss's curve for the case when $p \equiv 3$ mod 4. Gauss's curve $x^2t^2 + y^2t^2 + x^2y^2 - t^4 = 0$ contains two ordinary double points at infinity. Therefore C does not satisfy the hypothesis of the Weil Conjectures. The zeta function $Z_C(u)$ has a different form than the zeta function of a smooth projective plane curve of similar degree.

Theorem 14. *Consider the curve $C : x^2t^2 + y^2t^2 + x^2y^2 - t^4$ over $\mathbb{F}_p$ where $p \equiv 3$ mod 4. Then*

$$N_n(C) = \begin{cases} p^n + 3 & \text{if } 2 \nmid n; \\ p^n - 2(i\sqrt{p})^n - 1 & \text{if } 2|n \end{cases}$$

and

$$Z_C(u) = \frac{(1 + u)^2(1 + pu^2)}{(1 - u)(1 - pu)}.$$

Proof. Recall from Lemma 6 that $N_n(G_0) = p^n + 1$ for odd n, and from Lemma 7 that $N_n(E_0) = p^n - 2(i\sqrt{p})^n$ for even n. Putting this together with Proposition 21 and 22 we have that

$$N_n(C_0) = \begin{cases} p^n + 1 & \text{if } 2 \nmid n; \\ p^n - 2(i\sqrt{p})^n - 3 & \text{if } 2|n. \end{cases}$$

The curve $C : x^2t^2 + x^2y^2 + y^2 - t^4$ has the two points $P_1 = [1, 0, 0]$ and $P_2 = [0, 1, 0]$ at infinity. Therefore

$$N_n(C) = \begin{cases} p^n + 3 & \text{if } 2 \nmid n; \\ p^n - 2(i\sqrt{p})^n - 1 & \text{if } 2|n. \end{cases}$$

In order to calculate the zeta function, notice that $N_n(C)$ can be rewritten for any value of n as

$$N_n(C) = p^n + 1 - (i\sqrt{p})^n - (-i\sqrt{p})^n - 2(-1)^n.$$

Therefore

$$Z_C(u) = \exp\Big(\sum_{n=1}^{\infty} \frac{(p^n + 1 - (i\sqrt{p})^n - (-i\sqrt{p})^n - 2(-1)^n)u^n}{n}\Big).$$

Using the identity $\sum_{n=1}^{\infty} w^n n^{-1} = -\ln(1-w)$ we get the desired result

$$Z_C(u) = \frac{(1+u)^2(1+pu^2)}{(1-u)(1-pu)}.$$

□

2.5 Normalization of Singular Curves

The relationship between the zeta function of a singular curve and its normalization has been studied in [46] and [51]. Gauss's curve C is an example of a projective plane curve with singularities. It has two ordinary double points at $P_1 = [1,0,0]$ and $P_2 = [0,1,0]$. By [15, Chapter 17], there exists a nonsingular projective curve $\tilde{C}$ along with a normalization map $\nu : \tilde{C} \to C$. For every nonsingular point P of C, the preimage $\nu^{-1}(P)$ consists of only one point.

Another approach to determining $Z_C(u)$ is to identify $\tilde{C}$ and its zeta function $Z_{\tilde{C}}(u)$. Then $N_n(C)$ can be calculated by comparing it to $N_n(\tilde{C})$ while considering the size and field of definition of $\nu^{-1}(P_1)$ and $\nu^{-1}(P_2)$. This is essentially what we have done in Sections 2-2.4 with $\tilde{C} = E$ and $\nu = \mu \circ \alpha$.

Let C_{sing} represent the set of singular points of C. Let $Q|P$ denote the set of points $q \in \tilde{C}$ such that $\nu(Q) = P$. Also let $\deg(P) = \dim(\tilde{\mathcal{O}}_P/\mathcal{O}_P)$ where $\tilde{\mathcal{O}}_P$ is the integral closure of $\mathcal{O}_P$. The following proposition explains how the zeta function of a singular curve is related to the zeta function of its normalization. It is a consequence of the Euler product representation of the zeta function [23, Chapter 8.4].

Proposition 25. *(See, e.g., [6, Section 2]) Let X be a complete irreducible algebraic projective curve with normalization $\tilde{X}$. Then*

$$\frac{Z_X(u)}{Z_{\tilde{X}}(u)} = \Pi_{P \in X_{sing}} \frac{\Pi_{Q|P}(1 - u^{\deg(Q)})}{1 - u^{\deg(P)}}.$$

For p any odd prime, C has two degree one singular points $P_1 = [1, 0, 0]$ and $P_2 = [0, 1, 0]$. If $p \equiv 3 \bmod 4$, there is one point of degree 2 on E for each of these, hence $Z_C(u)/Z_E(u) = (1 + u)^2$. When $p \equiv 1 \bmod 4$, there are two points of degree 1 on E for each of these, yielding $Z_C(u)/Z_E(u) = (1 - u)^2$.

Chapter 3

The Global Zeta Function of C

Let X be a non-singular elliptic curve defined over $\mathbb{Z}$ with discriminant Δ, let $X_p = X \times \mathbb{F}_p$, and let $\mathcal{S} = \{p \text{ prime: } p|\Delta\}$ be the set of primes of bad reduction for X. Then, the above defined function $Z_{X_p}(u)$ exists for all primes $p \notin \mathcal{S}$. Via the change in variables $u = p^{-s}$, we can define

$$Z_{X_p}(p^{-s}) = \zeta_{X_p}(s)$$

to be the local zeta function of X at p. See [19, chapter 18.2] for reference.

Since we will now be considering the same equations, but viewed over $\mathbb{F}_p$ for varying primes p, we will change notation slightly in what follows. Consider $C : x^2t^2 + y^2t^2 + x^2y^2 - t^4 = 0$ to be defined over $\mathbb{Z}$ and let $C_p : C \times \mathbb{F}_p$. Define E and E_p similarly. For any curve X, let $N_p(X) = |X_p(\mathbb{F}_p)|$ and let $\alpha_p = p + 1 - N_p$. We then have that, for $p \notin \mathcal{S}$,

$$\zeta_{X_p}(s) = \frac{1 - \alpha_p p^{-s} + p^{1-2s}}{(1 - p^s)(1 - p^{1-s})}.$$

For $p \in \mathcal{S}$, define

$$\zeta_{X_p}(s) = \frac{1}{(1 - p^{-s})(1 - p^{1-s})}.$$

The global zeta function of X is defined to be the product of the local zeta functions:

$$\zeta_X(s) = \prod_p \zeta_{X_p}(s).$$

Let $L_X(s) = \prod_{p|\Delta}(1 - \alpha_p p^{-s} + p^{1-2s})^{-1}$, called the L-function of X. Taking the product over all p, we have

$$\zeta_X(s) = \frac{\zeta(s)\zeta(s-1)}{L_X(s)},$$

where $\zeta(s)$ is the Riemann zeta function. Determining the global zeta function of X is equivalent to determining its L-function.

Remark 2. *Let P be a prime of $\mathbb{Z}[i]$. Let $N(P)$ be the norm of P. For $A \in \mathbb{Z}[i]$, let $(\frac{A}{P})_4 \in \{0, \pm 1, \pm i\}$ be the quartic residue of A modulo P. That is,*

$$(\frac{A}{P})_4 = 0 \text{ if } P|A$$

and

$$(\frac{A}{P})_4 P = A^{\frac{N(P)-1}{4}} \text{ otherwise.}$$

Define a Hecke character χ on primes P of $\mathbb{Z}[i]$. If P divides 8 define $\chi(P) = 0$. If $N(P) = p^2$, then $p \equiv 3 \bmod 4$ and $P = p$, where p is inert in $\mathbb{Z}[i]$. In this case define $\chi(P) = -p$. If $N(P) = p$, i.e. (p) splits in $\mathbb{Z}[i]$ and $p \equiv 1 \bmod 4$, then $P = (\pi)$ for some $\pi \in \mathbb{Z}[i]$ with $\pi \equiv 1 \bmod (2+2i)$. Define $\chi(P) = \overline{(\frac{4}{(\pi)})_4}\pi$.

The Hecke L-function associated to χ is defined as

$$L(s, \chi) = \prod_{P \text{ prime of } \mathbb{Z}[i]} (1 - \chi(P)N(P)^{-s})^{-1}.$$

For the case of the elliptic curve $E = \tilde{C} : y^2t - x^3 + 4xt^2 = 0$, it is shown in [19, chapter 18.6] that $L_E(s) = L(s, \chi)$

We use a similar elementary definition for the global zeta function of a singular curve (contrast with [6] and [47]).

Definition 9. *Let Y be a singular curve with normalization $\tilde{Y}$, where $\tilde{Y}$ has singular set $\mathcal{S}$. Let $\zeta_{Y_p}(s) = \frac{1}{(1-p^{-s})(1-p^{1-s})}$ for $p \in \mathcal{S}$. Define the global zeta function of Y to*

be

$$\zeta_Y(s) = \prod_p \zeta_{Y_p}(s).$$

Definition 10. *Define a Dirichlet character* $\chi' : \mathbb{Z} \to \{0, \pm 1\}$*, where* $\chi'(n) = 0$ *if* n *is even,* $\chi'(n) = 1$ *if* $n \equiv 1 \bmod 4$*, and* $\chi'(n) = -1$ *if* $n \equiv 3 \bmod 4$*.*

The Dirichlet L-function associated to χ' is defined to be

$$L(s, \chi') = \prod_{p \text{ prime of } \mathbb{Z}} (1 - \chi'(p)p^{-s})^{-1}.$$

Theorem 15. *The global zeta function for* C *is given by*

$$\zeta_C(s) := \prod_p \zeta_{C_p}(s) = \frac{\zeta(s)\zeta(1-s)}{L_E(s)L(s,\chi')^2}.$$

Proof. Recall that $N_p(C) = p - 1 - 2a_p$ for $p \equiv 1 \bmod 4$, where a_p is the value a such that $a^2 + b^2 = p$ chosen in Conjecture 1. We then have

$$\begin{aligned} \zeta_C(s) &= \prod_p \zeta_{C_p}(s) \\ &= \frac{1}{(1-2^{-s})(1-2^{1-s})} \prod_{p \equiv 1(4)} \frac{(1 - 2a_p p^{-s} + p^{1-2s})(1-p^{-s})}{1-p^{1-s}} \prod_{p \equiv 3(4)} \frac{(1+p^{-s})^2(1+p^{1-2s})}{(1-p^{-s})(1-p^{1-s})}. \end{aligned}$$

A few simplifications yield the form:

$$\zeta_C(s) = \zeta(s)\zeta(1-s) \prod_{p \equiv 1(4)} (1 - 2a_p p^{-s} + p^{1-2s})(1-p^{-s})^2 \prod_{p \equiv 3(4)} (1+p^{-s})^2(1+p^{1-2s})$$

Now, consider the relationship between a_p and α_p, where $\alpha_p = p + 1 - N_p(E)$. When $p \equiv 1 \bmod 4$, the two singularities on C_p are double points, which in the normalization E_p yield two points each. That means that $N_p(E) = N_p(C) + 2$, giving

$$p + 1 - \alpha_p = p - 1 - 2a_p + 2,$$

so $\alpha_p = 2a_p$ for $p \equiv 1 \bmod 4$. When $p \equiv 3 \bmod 4$, we know that $N_p(E) = p + 1$, so $\alpha_p = 0$.

Therefore

$$\begin{aligned}\zeta_C(s) &= \zeta(s)\zeta(1-s)\prod_{p\neq 2}(1-\alpha_p^{-s}+p^{1-2s})(1-(-1)^{\frac{p-1}{2}}p^{-s})^2 \\ &= \zeta(s)\zeta(1-s)L_E(s)^{-1}L(s,\chi')^{-2}.\end{aligned}$$

□

Part IV

Digital Signatures from LWE over $\mathbb{Z}/q[x]/(f(x))$

This work started in the summer of 2010 during an internship with the cryptography group at Microsoft Research. It is joint work with Kristin Lauter, Michael Naehrig, and Vinod Vaikuntanathan. We have filed a patent on the scheme through Microsoft and are preparing a paper for submission to AsiaCrypt. This summary is intended to give background, outline the scheme, and give an idea of the security proof for one set of parameters. As written, the scheme is secure only for one-time use, i.e. new keys would need to be generated for each signature. However, Merkle trees provide a way to convert the one-time scheme into a reusable scheme with a loss of efficiency logarithmic in the number of signatures required [29].

Chapter 1

Overview: Learning With Errors over Polynomial Rings

1.1 Learning with Errors

Let q, n, and m be integers. Fix a vector $s \in (\mathbb{Z}/q)^n$. For $i \in [1, m]$, choose $a_i \in (\mathbb{Z}/q)^n$ uniformly at random and $e_i \in \mathbb{Z}/q$ from an error distribution concentrated near 0. The search learning with errors problem (LWE) is to find s, given access to m pairs $(a_i, b_i = a_i \cdot s + e_i)$. Each pair can be thought of as a noisy inner product of s with a random vector a_i. The decision LWE problem is to distinguish with some non-negligible probability the pairs of the form $(a_i, a_i \cdot s + e_i)$ from those of the form (a_i, u_i), where u_i is chosen independently and uniformly at random from $\mathbb{Z}/q$.

Solving these problems has been proven to be as hard as certain lattice problems. In particular, the approximate shortest vector problem (GapSVP) is to approximate the length of the shortest non-zero vector in a lattice within a polynomial factor. All known algorithms are exponential in n, the dimension of the lattice. In 2005, Regev [33] proved that given a quantum computer and a search-LWE oracle, can solve GapSVP in polynomial time with an approximation factor $O(\frac{n}{\alpha})$, where the error distribution is a discrete Gaussian scaled by αq. He also proved that if q

is prime and polynomial in n, search-LWE and decision-LWE are equivalent (even without a quantum computer). In 2007, Regev [34] proved that if q is the product of many small primes and the error distribution is Gaussian, search and decision are equivalent even for larger n. Peikert [32] proved in 2009 that if $q \geq 2^{\frac{n}{2}}$ and we have a search-LWE oracle, we can solve GapSVP in polynomial time with no quantum computer required, with the same approximation factor as Regev's 2005 result. Also, if q is polynomial in n, Peikert found a classical reduction to LWE from a variant of GapSVP known as ζ-to-γ-GapSVP.

Many cryptographic schemes with security based on LWE have been devised. However, large key sizes make these schemes less than ideal for practical use. One interesting way to increase efficiency is to work over certain polynomial rings instead of $\mathbb{Z}/q$.

1.2 Ring LWE

Let R_q be $\mathbb{Z}/q[x]/(f(x))$, where$f(x) \in \mathbb{Z}/q[x]$ is a polynomial of degree n. Elements of R_q can be described in terms of the power basis $\{1, x, x^2, ...x^{n-1}\}$. That is, for any $c \in R_q$, we can write $c = c_0 + c_1x + ... + c_{n-1}x^{n-1}$ for $c_i \in \mathbb{Z}/q$. Let s be a fixed ring element. The (search) ring LWE problem (R-LWE) is to find s given access to perturbed products of the form $(a_i, b_i = a_i * s + e_i)$ where the $a_i \in R_q$ are ring elements selected uniformly at random and the components of the $e_i \in R_q$ are chosen from some distributions concentrated near 0, and the operations $+$ and $*$ are addition and multiplication in the ring R_q. The decision R-LWE problem is to distinguish such pairs from pairs in which both members have been chosen uniformly at random from R_q.

As standard LWE problems can be related to problems in integer lattices, R-LWE problems relate to problems in ideal lattices inside rings. Lyubashevsky, Peikert, and Regev [25] and Stehle, Steinfeld, Tanaka, and Xagawa [44] both study the ideal

lattice/R-LWE connection. A comparison of their approaches can be found in Lyubashevsky, Peikert, and Regev's paper [25], and we in general follow their development here.

The main result of [25] can be broken into two parts. The first combines with results of Regev [34] to give a quantum reduction from GapSVP on ideal lattices in $\mathbb{Z}[x]/(f(x))$ to search R-LWE in R_q. The second part gives a non-quantum reduction from search R-LWE to the decision variant of R-LWE. An informal statement of their main theorem, paraphrased from [25], is: Suppose that it is hard for polynomial-time quantum algorithms to approximate the shortest vector problem in the worst case on ideal lattices in R to within a fixed polynomial-in-n factor. Then any polynomial-in-n number of R-LWE samples are pseudorandom to any polynomial time attacker. Note that approximating the shortest vector in an ideal lattice has not been found to be any easier than approximating the shortest vector in an integer lattice, despite many attempts in this direction. Therefore the hardness of R-LWE is a reasonable basis for cryptographic security.

Chapter 2

Specifics for R and R_q

For the purpose of this paper, we focus on the case $f(x) = x^n + 1$ with $n = 2^k$, where q is prime and $q \equiv 1 \bmod 2n$. In this case $R = \mathbb{Z}[x]/(f(x))$ is the ring of integers $\mathbb{Z}[\zeta_{2n}]$ inside the cyclotomic number field $K = \mathbb{Q}(\zeta_{2n})$. For simplicity, we will let $\zeta = \zeta_{2n}$. Since $|K : \mathbb{Q}| = n$, there are n different field homomorphisms of K into $\mathbb{C}$ which fix $\mathbb{Q}$. Let σ_i, for $i \in [1, n]$ be the i-th such embedding. Then R can be viewed as a lattice inside $\mathbb{C}^n$ via the canonical embedding map

$$\sigma(c) = (\sigma_1(c), \sigma_2(c), ..., \sigma_n(c)).$$

The Euclidean 2-norm on $\mathbb{C}^n$ induces a norm on the elements of R:

$$|c| = \sqrt{\sum_{i=1}^{n} |\sigma_i(c)|^2}.$$

The embeddings σ_i can be described by their action on x. It is most convenient here to index the embeddings by $i \in (\mathbb{Z}/(2n))^\times$, the group of units of $\mathbb{Z}/(2n)$. We say that $\sigma_i(x) = \zeta^i$ for $i \leq n$, and $\sigma_i(x) = \overline{\zeta^{i-n}}$ for $i \in [n+1, 2n]$. Consider $c = c_0 + c_1 x + c_2 x^2 + ... + c_{n-1} x^{n-1} \in R$. We can uniquely describe c by its coefficients with the column vector $\vec{c} = [c_0, c_1, ..., c_n - 1]$. Here, we drop the vector notation and refer to the ring element and the coefficient vector with the same symbol c. The

canonical embedding then acts on c by the matrix

$$\Sigma = \begin{pmatrix} 1 & \zeta & \zeta^2 & \dots & \zeta^{n-1} \\ 1 & \zeta^2 & \zeta^4 & \dots & \zeta^{2n-2} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & \zeta^{n+1} = \zeta^{-(n-1)} & \zeta^{2(n+1)} = \zeta^2 & \dots & \zeta^{(n+1)(n-1)} = \zeta^{-1} \end{pmatrix}$$

via $\sigma(c) = \Sigma c$.

The norm induced by this canonical embedding, while mathematically nice, can be less than intuitive. However in the specific case described here, all embeddings of x have $\sigma_i(x) = \sqrt{n}$, which leads to more intuitive behavior. In the situation described in this paper, i. e. for a cyclotomic number field $K = \mathbb{Q}(\zeta_{2n})$ where $n = 2^k$ is a power of 2, the norm $\mathrm{N}(a)$ of an element $a \in R$ induced by the 2-norm on $\mathbb{C}^n$ via the canonical embedding can be computed from the coefficients of a given in the basis $\{1, \zeta, \zeta^2, \dots, \zeta^n - 1\}$ of R. We have

$$\mathrm{N}(a) = \sqrt{n} \left(\sum_{i=0}^{n-1} a_i^2 \right)^{1/2}$$

for $a = \sum_{i=0}^{n-1} a_i \zeta^i$. This can be seen as follows. We first show that the canonical embeddings of the basis elements ζ^i are orthogonal. Let ζ^i and ζ^j be two such elements, i. e. $i, j \in [0, n-1]$. We consider the inner product

$$\langle \sigma(\zeta^i), \sigma(\zeta^j) \rangle = \langle \sigma(\zeta^i), \overline{\sigma(\zeta^{-j})} \rangle = \mathrm{Tr}(\zeta^{i-j}),$$

where the latter equality follows from [25, Section 2.3.3]. The trace can be computed as

$$\mathrm{Tr}(\zeta^{i-j}) = \sum_{l \in (\mathbb{Z}/(2n))^*} \sigma_l(\zeta^{i-j}) = \sum_{l \in (\mathbb{Z}/(2n))^*} \zeta^{(i-j)l}.$$

For $i = j$, we have $\mathrm{Tr}(\zeta^{i-j}) = n$. If $i \neq j$, for each $l \in (\mathbb{Z}/(2n))^*$ there exists $l' \in (\mathbb{Z}/(2n))^*$ such that $(i-j)l \equiv (i-j)l' + n \pmod{2n}$. This means that all terms in the trace cancel, so $\mathrm{Tr}(\zeta^{i-j}) = 0$ in this case. This shows that the embeddings of the basis elements are orthogonal to each other.

To compute the norm, we rewrite its square as a trace. We have

$$\mathrm{N}(a)^2 = \langle a, \overline{a} \rangle = \mathrm{Tr}(a \cdot \overline{a}) = \mathrm{Tr}(\sum_{i=0}^{n-1} \sum_{j=0}^{n-1} a_i a_j \zeta^{i-j}) = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} a_i a_j \mathrm{Tr}(\zeta^{i-j}).$$

Since we just computed the traces occuring in the sum, we conclude that $\mathrm{N}(a)^2 = n \sum_{i=0}^{n-1} a_i^2$.

2.1 Error Distributions

The R-LWE problem requires us to sample small random elements of a polynomial ring. The norm discussion above gives gives us an idea of what we mean by small. We begin by defining the standard generalization of the Gaussian distribution to higher dimensions.

Let $x \in \mathbb{R}^n$. The n-dimensional spherical Gaussian probability density function with parameter r is given by

$$\rho_r(x) = exp(-\pi(\frac{|x|}{r})^2).$$

This gives the probability of sampling a given value x according to the distribution D_r. Under the standard Euclidean norm of x, this spherical distribution chooses each coordinate of x using the same Gaussian distribution. If we were to use a different norm on $\mathbb{R}^n$, the distributions for each coordinate might be skewed, resulting in an elliptical Gaussian distribution in the standard norm.

To sample the error terms necessary for our signature scheme, we first consider the ring $K \otimes \mathbb{R}$, which essentially extends the field of constants of R to allow real coefficients. The canonical embedding of K extends to $K \otimes \mathbb{R}$. If $c' \in K \otimes \mathbb{R}$, then in the coordinates of $\vec{c'}$ are in $\mathbb{R}$ instead of $\mathbb{Q}$. This allows us to sample each coordinate of $\vec{c'}$ from a Gaussian distribution with parameter of our choosing.

When we embed c' using σ, the resulting distribution is in general an elliptical Gaussian distribution. However, for the specific case of R, the distribution remains

spherical (see [25] for a more complete discussion). Lyubashevsky, Peikert, and Regev proved the security reduction discussed above for error terms chosen from the family $\Psi_{\leq\alpha}$, defined to be the set of all elliptical Gaussian distributions over $K \otimes \mathbb{R}$ which result from choosing a Gaussian distribution with parameter $r_i \leq \alpha$ for each coordinate axis of $K \otimes \mathbb{R}$. These elements of $K \otimes \mathbb{R}$ are then rounded to yield elements of R, giving error terms from a discrete Gaussian probability distribution on R.

2.2 Working in R_q

Recall that we would like to work in a finite setting, i.e. $R_q = \mathbb{Z}/q[x]/(f(x))$. We chose $f(x) = x^n + 1$ with $n = 2^k$, which is irreducible over $\mathbb{Z}[x]$. However, this polynomial may factor in $\mathbb{Z}/q[x]$ depending on q. We chose $q \equiv 1 \bmod 2n$ because $x^n + 1$ factors completely in $\mathbb{Z}/q[x]$ for q of this form as $x^n + 1 = \Pi_{i=1}^{n}(x - a_i)$. This yields the ring isomorphism

$$R_q = \mathbb{Z}/q[x]/(f(x)) \cong \mathbb{Z}/q[x]/(x-a_1) \oplus \mathbb{Z}/q[x]/(x-a_2) \oplus ... \oplus \mathbb{Z}/q[x]/(x-a_n) \cong (\mathbb{Z}/q)^n.$$

This matches our concrete description of R_q above, where $c = c_0 + c_1 x + .. c_{n-1} x^{n-1} \in R_q$ with $c_i \in \mathbb{Z}/q$ corresponds to $c = [c_0, c_1, ..., c_{n-1}] \in R_q$. So a uniformly random element of R_q can be sampled as n uniformly random elements of $\mathbb{Z}/q$. As for the Gaussian distribution, q is chosen to be much larger than n so that an element from the Gaussian distribution described above can be sampled as n elements of $\mathbb{Z}$ from discrete Gaussian distributions centered at 0. These integers are then interpreted as elements of $\mathbb{Z}/q$ and the vector with these coefficients is interpreted as an element of R_q. The norm above is also valid for elements of R_q and is well defined when the representatives for the coefficients of ring elements are defined to have minimum absolute value modulo q. For example, the value $q - 1 \in \mathbb{Z}/q$ would be represented by -1.

Chapter 3

Digital Signatures from Ring LWE

3.1 Peikert, Lyubashevsky, and Regev's simple ring LWE scheme

The signature scheme we have devised is based on the simple public key cryptographic protocol designed by Peikert, Lyubashevsky, and Regev [25]

- Secret key : short $s \in R_q$.
- Public key : Choose random $a \in R_q$, key is $(a, b = as + e)$
- Encrypt $m \in \{0,1\}^n$: choose short $t \in R_q$. Output ciphertext

$$(c1, c2) = (at + e_1, bt + e_2 + m\lceil\frac{q}{2}\rceil) \approx (at, ast + m\lceil\frac{q}{2}\rceil)$$

- Decrypt: recover m from $c_2 - c_1 s$.

3.2 One-time signature scheme

We devised a signature scheme based on [25]. However, the details must be omitted here because the scheme is in the patent process.

3.2.1 Security for small q

We prove security based on the hardness of ring LWE for $q \approx n^{\log(n)}$. This results in large communication complexity and large keys. With an additional assumption, namely that ring LWE is still hard given some auxiliary inputs, the signature scheme is secure for q on the order of n^3. The second set of parameters results in a very efficient system and small keys. We are currently in the process of determining the weakest assumption under which we can prove security for smaller q. This would make the scheme feasible and much more efficient than lattice based signature schemes from the literature [24].

Bibliography

[1] Miriam Abdon, Juscelino Bezerra, and Lucianne Quoos. Further examples of maximal curves. 2007.

[2] Henning Stichtenoth Arnaldo Garcia and Chao-Ping Xing. On subfields of the Hermitian function field. *Compositio Mathematica*, 120, 2000.

[3] Matthew Baker and Serguei Norine. Harmonic morphisms and hyperelliptic graphs. *International Mathematics Research Notices*, 2009(15), 2009.

[4] Norman Biggs. *Algebraic Graph Theory, Second Edition*. Cambridge University Press, 1993.

[5] W. G. Bridges and R. A. Mena. Multiplicative cones - a family of three eigenvalue graphs. *Aequationes Mathematicae*, 22, 1981.

[6] F. N. Castro and C. J. Moreno. L-functions of singular curves over finite fields. *J. Number Theory*, 84(1):136–155, 2000.

[7] D. de Caen, E. R. van Dam, and E. Spence. A nonregular analogue of conference graphs. *Journal of Combinatorial Theory*, 88, 1999.

[8] David S. Dummit and Richard M. Foote. *Abstract Algebra*. John Wiley and Sons, 2004.

[9] Iwan Duursma and Kit-Ho Mak. Two families of maximal curves which are not galois subcovers of the hermitian curve. preprint, 2010.

[10] Rainer Fuhrmann and Fernando Torres. The genus of curves over finite fields with many rational points. *Manuscripta Mathematica*, 89, 1996.

[11] Arnaldo Garcia, Cem Guneri, and Henning Stichtenoth. A generalization of the Guilietti-Korchmaros maximal curve. *Advances in Geometry*, accepted.

[12] Massimo Giulietti and Gabor Korchmaros. A new family of maximal curves over a finite field. *Mathematische Annalen*, 343, January 2009.

[13] J. J. Gray. A commentary on Gauss's mathematical diary, 1796–1814, with an English translation. *Exposition. Math.*, 2(2):97–130, 1984.

[14] Larry C. Grove. *Classical Groups and Geometric Algebra*. AMS, 2002.

[15] J. Harris. *Algebraic geometry*, volume 133 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1995. A first course, Corrected reprint of the 1992 original.

[16] J.W.P. Hirschfeld, G. Korchmaros, and F. Torres. *Algebraic Curves over a Finite Field*. Princeton University Press, 2008.

[17] Ki ichiro Hashimoto. Zeta functions of finite graphs. In K. Hashimoto and Y. Namikawa, editors, *Automorphic Forms and Geometry of Arithmetic Varieties*. Academic Press, Harcourt Brace Jovanovich, 1989.

[18] Yasutaka Ihara. On discrete subgroups of the two by two projective linear group over p-adic fields. *J. Math. Soc. Japan*, 18:219–235, 1966.

[19] K. Ireland and M. Rosen. *A classical introduction to modern number theory*, volume 84 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, second edition, 1990.

[20] Neal Koblitz. *p-adic Numbers, p-adic Analysis, and Zeta-Functions*. Springer-Verlag, 1977.

[21] Gabor Korchmaros and Fernando Torres. Embedding a maximal curve in a hermitian variety. *Compositio Math*, 128, 2001.

[22] Motoko Kotani and Toshikazu Sunada. Zeta functions of finite graphs. *Journal of the Mathematical Society of University of Tokyo*, 7, 2000.

[23] D. Lorenzini. *An invitation to arithmetic geometry*, volume 9 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 1996.

[24] Vadim Lyubashevsky and Daniele Micciancio. Asymptotically efficient lattice-based digital signatures. In *TCC*, pages 37–54, 2008.

[25] Vadim Lyubashevsky, Chris Peikert, and Oded Regev. On ideal lattices and learning with errors over rings. In *EUROCRYPT*, pages 1–23, 2010.

[26] Beth Malmskog and Michelle Manes. Ramified covers of graphs and the ihara zeta function of certain ramified covers. In *WIN - Women In Numbers: Research Directions in Number Theory*, volume 60 of *Fields Institute Communications*.

[27] Beth Malmskog and Michelle Manes. Almost divisibility in the Ihara zeta functions of certain ramified covers of q+1-regular graphs. *Linear Algebra and its applications*, 432, 2010.

[28] Yuri Manin. The theory of commutative formal groups over fields of finite characteristic. *Russian Mathematical Surveys*, 18, 1963.

[29] Ralph Merkle. *Secrecy, authenication, and public key systems*. PhD thesis, Stanford University, 1979.

[30] Mikhail Muzychuk and Mikhail Klin. On graphs with three eigenvalues. *Discrete Mathematics*, 189, 1998.

[31] Sam Northshield. A note on the zeta function of a graph. *Journal of Combinatorial Theory*, 64, 1998.

[32] Chris Peikert. Public-key cryptosystems from the worst-case shortest vector problem: extended abstract. In *STOC*, pages 333–342, 2009.

[33] Oded Regev. On lattices, learning with errors, random linear codes, and cryptography. In *STOC*, pages 84–93, 2005.

[34] Oded Regev. On lattices, learning with errors, random linear codes, and cryptography. *J. ACM*, 56(6), 2009.

[35] H.G. Rück and Henning Stichtenoth. A characterization of Hermitian function fields over finite fields. *Journal für die reine und angewandte Mathematik*, 457, 1994.

[36] Iwao Sato. Zeta functions and complexities of a semiregular bipartite graph and its line graph. *Discrete Mathematics*, 307, 2007.

[37] Jean-Pierre Serre. *Algebraic Groups and Class Fields*. Springer, 1975, translation 1988.

[38] Jean-Pierre Serre. *Local Fields*. Springer, 1979.

[39] Igor Shafarevich. *Algebraic Geometry I: Varieties in Projective Space*. Springer-Verlag, 1977.

[40] Joseph Silverman. *The Arithmetic of Elliptic Curves*. Springer, 1986.

[41] Harold Stark and Audrey Terras. Zeta functions of finite graphs and coverings. *Adv. Math.*, 121(1):124–165, 1996.

[42] Harold Stark and Audrey Terras. Zeta functions of finite graphs and coverings II. *Adv. Math.*, 154(1):132–195, 2000.

[43] Harold Stark and Audrey Terras. Zeta functions of finite graphs and coverings III. *Adv. Math.*, 208(1):467–489, 2007.

[44] Damien Stehlé, Ron Steinfeld, Keisuke Tanaka, and Keita Xagawa. Efficient public key encryption based on ideal lattices. In *ASIACRYPT*, pages 617–635, 2009.

[45] Henning Stichtenoth. *Algebraic Function Fields and Codes*. Springer-Verlag, 1993.

[46] K.-O. Stöhr. On the poles of regular differentials of singular curves. *Bol. Soc. Brasil. Mat. (N.S.)*, 24(1):105–136, 1993.

[47] Karl-Otto Stöhr. Local and global zeta-functions of singular algebraic curves. *J. Number Theory*, 71(2):172–202, 1998.

[48] Audrey Terras. A stroll through the garden of graph zeta functions. Unpublished, available at http://math.ucsd.edu/%7Eaterras/newbook.pdf, 2007.

[49] Hajime Urakawa. A discrete analogue of the harmonic morphism. In Christopher Kum Anand, editor, *Harmonic Morphisms, Harmonic Maps and Related Topics*. CRC Press, 1999.

[50] Edwin R. van Dam. Nonregular graphs with three eigenvalues. *Journal of Combinatorial Theory*, 73, 1998.

[51] W. A. Zúñiga-Galindo. Zeta functions and Cartier divisors on singular curves over finite fields. *Manuscripta Math.*, 94(1):75–88, 1997.