UNIVERSITY OF CALIFORNIA, SAN DIEGO

A Combinatorial Comparison of Elliptic Curves and Critical Groups of Graphs

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Mathematics

by

Gregg Joseph Musiker

Committee in charge:

Professor Adriano Garsia, Chair
Professor Ronald Graham
Professor Russell Impagliazzo
Professor Harold Stark
Professor Nolan Wallach

2007

The dissertation of Gregg Joseph Musiker is approved, and it is acceptable in quality and form for publication on microfilm:

Chair

University of California, San Diego

2007

To the memory of my Grandparents Bette and Philip Rosenthal
who continue to inspire me.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

First, with much appreciation I thank my advisor, Adriano Garsia, for his guidance, continual enthusiasm, and unwavering dedication over the last five years. Adriano has taught me a variety of beautiful mathematical topics, and his passion for mathematics, the beach, and life in general has been quite contagious. I am also indebted to Nolan Wallach for aiding me in my studies of representation theory and algebraic geometry, and appreciate his invaluable feedback during my graduate school. Many other professors have helped me along my journey, and while I can't name them all, I wanted to especially thank Wee Tak Gan, Allen Knutson, Jim Propp, Christophe Reutenauer, Harold Stark, William Stein, and Richard Stanley. I am thankful to Sam Buss and Jim Lin for their dedication as Chair and Vice-Chair and their work initiating new opportunities and support for graduate students; as well as the excellent staff of the UCSD Mathematics Department, especially Lois Stewart, Wilson Cheung, and Yi Ling Ng. Additionally, I would like to thank the San Diego Chapter of the ARCS Foundation for their financial support during my graduate school.

My colleagues and friends, Jason Bandlow, Arthur Berg, Dave Clark, Mark Colarusso, Eric Tressler, Jake Wildstrom, Aaron Wong, Scott Cohen, Rick Capella, Andrew Cosand, German Eichberger, Emmi Olson, Jeff Gold, Emily Anderson, Lee Lovejoy, and countless others, thank you all for enriching my graduate school experience with technical and moral support, as well as frisbee, poker, and many good meals. Most of all, I would like to thank my parents Brian and Lori Musiker for always inspiring me to learn. This thesis is only possible because of their continual love, advice, and support.

Much of the material in Chapter 4 and 5 has been submitted for publication in the paper "Combinatorial Aspects of Elliptic Curves" by Gregg Musiker. The dissertation author is the primary investigator and author of this paper.

VITA

1980	Born, Philadelphia, Pennsylvania
2002	B. A., *magna cum laude*, Harvard University
2002–2006	Teaching assistant, Department of Mathematics, University of California San Diego
2004	M. A., University of California San Diego
2006	Associate Instructor, Department of Mathematics, University of California San Diego
2007	Ph. D., University of California San Diego

PUBLICATIONS

J. Bandlow and G. Musiker. Quasi-invariants of S_3. *J. Combin. Theory Ser. A* **109** (2005), no 2, 281-298.

G. Musiker and J. Propp. Combinatorial Interpretations for Rank-Two Cluster Algebras of Affine Type. *Electronic Journal of Combinatorics.* **14** (2007), no R15, 1-23.

A. Garsia and G. Musiker. *Basics on Hyperelliptic Curves over Finite Fields.* Mongraphies du LaCIM. (To appear.)

G. Musiker. Combinatorial Aspects of Elliptic Curves. (Submitted.)

ABSTRACT OF THE DISSERTATION

A Combinatorial Comparison of Elliptic Curves and Critical Groups of Graphs

by

Gregg Joseph Musiker
Doctor of Philosophy in Mathematics

University of California San Diego, 2007

Professor Adriano Garsia, Chair

In this thesis, we explore elliptic curves from a combinatorial viewpoint. Given an elliptic curve E, we study here $N_k = \#E(\mathbb{F}_{q^k})$, the number of points of E over the finite field $\mathbb{F}_{q^k}$. This sequence of numbers, as k runs over positive integers, has numerous remarkable properties of a combinatorial flavor in addition to the usual number theoretical interpretations. In particular we prove that $N_k = -\mathcal{W}_k(q,t)|_{t=-N_1}$ where $\mathcal{W}_k(q,t)$ is a (q,t)-analogue for the number of spanning trees of the wheel graph. Additionally we develop a determinantal formula for N_k where the eigenvalues can be explicitly written in terms of q, N_1, and roots of unity. We also discuss here a new sequence of bivariate polynomials related to the factorization of N_k, which we refer to as elliptic cyclotomic polynomials because of their various properties.

The above formula for N_k in terms of $\mathcal{W}_k$ motivates a closer examination of the relationship between points on an elliptic curve E over $\mathbb{F}_{q^k}$ and spanning trees on the wheel graph W_k. An elliptic curve E has an abelian group structure, and indeed the set of spanning trees of a graph also has an abelian group structure. Here we study one isomorphic to the critical group of the graph, which has ties to the theory of chip-firing games and abelian sandpile models of dynamical systems. While we first focus on the relationship between the integer sequences $\{N_k\}$ and $\{\mathcal{W}_k(q,N_1)\}$, we also compare these two group structures, illustrating that the

connections between elliptic curves and spanning trees run even deeper. Numerous theorems which are true for elliptic curve groups have analogues in terms of critical groups of the (q, t)-wheel graph.

Additionally, the theory of critical groups will also allow us to re-interpret the group elements as the set of admissible words for a primitive circuit in a specific deterministic finite automaton. As an application, we will then compare the zeta function of an elliptic curve and the zeta function of the corresponding cyclic language.

1 Introduction

An interesting problem at the cross-roads between combinatorics, number theory, and algebraic geometry is that of counting the number of points on an algebraic curve over a finite field. Over a finite field, the locus of solutions to an algebraic equation is a discrete subset, but since they satisfy a certain type of algebraic equation this imposes a lot of extra structure below the surface. One of the ways to detect this additional structure is by observing that considering field extensions, the infinite sequence of cardinalities is only dependent on a finite set of data. Specifically we let $\mathbb{F}_q$ denote the unique finite field, up to isomorphism, which has q elements. Since q is the size of this field, q must be a power of a prime, e.g. p^ℓ, and finite algebraic extensions of this field will result in fields with $q^k = p^{\ell k}$ elements. In the case of a genus g algebraic curve, the number of points over $\mathbb{F}_q$, $\mathbb{F}_{q^2}$, ..., and $\mathbb{F}_{q^g}$ will be sufficient data to determine the number of points over any other algebraic field extension.

This observation motivates the question of how the points over higher field extensions relate to points over the first g extensions. In this thesis we explore this question from a combinatorial point of view. We begin with background on algebraic curves which includes standard algebraic geometric terminology. This will include a definition of the zeta function, which is an exponential generating function defined by considering the sequence of numbers given by the cardinalities over various extension fields. We will then switch gears, and in Chapter 2 discuss a more combinatorial way to approach this problem and include connections to the theory of symmetric functions.

Afterwards, we will analyze in depth the case of elliptic curves, providing background in Chapter 3. We will utilize combinatorial methods with an eye towards future research for higher genus examples, such as the hyperelliptic case; and other algebraic varieties. However, while spelunking in the elliptic case during the course of my graduate school, many gems have been uncovered which have led to additional research directions with connections to critical groups of graph theory and dynamical systems. It will be this topic with which this thesis will be principally concerned, as Chapters 4-6 will illuminate. We close with connections to the zeta functions of rational languages, and in particular cyclic languages.

1.1 Background on algebraic curves

Unless otherwise specified, we will work over the finite field $\mathbb{F}_q$ in this section. We also will assume that we have taken C to be a nonsingular projective curve of genus g. (If not, our curve of interest is isomorphic to such a curve). Thus we can embed our curve into $\mathbb{P}^2$ and write its defining equation using the variables X, Y, and Z (or on a standard affine patch $\hat{C}$ with equation $f_{\hat{C}}$ in variables $x = X/Z$ and $y = Y/Z$). Note that the defining equation for C, f_C, will be homogeneous. We say that curve C is **defined over** $\mathbb{F}_q$ (or more generally defined over field k) if the coefficients of f_C lie in field $\mathbb{F}_q$ (resp. k). We note that the background material of these first few sections (except for Section 1.2) are common to numerous sources, for example [Ful89], [Lan82, Ch. 1], [Mil06], [Sil92].

Definition 1.1. The **coordinate ring** for affine curve $\hat{C}$ is defined as $\mathbb{F}_q[x,y] \big/ (f_{\hat{C}})$. We will sometimes denote this as $\mathbb{F}_q[\hat{C}]$.

Note that $\hat{C}$ being a variety implies that $f_{\hat{C}}$ is irreducible and this coordinate ring is an integral domain. Thus the notion of prime ideal is sensible. There is in fact a one-to-one correspondence between prime ideals and irreducible subvarieties of C. In particular, over an algebraically closed field k, the only prime ideals in $k[x,y] \big/ (f_{\hat{C}})$ are maximal ones, which correspond to points on C. For example in the **hyperelliptic case**, where $f_{\hat{C}}$ can be expressed as $y^2 = f_0(x)$, the prime ideals will either look like $(g(x),\ y - h(x))$ with $g(x), h(x) \in \mathbb{F}_q[x]$, or will be principal.

The entire curve C can be broken into two affine patches, so by considering the coordinate ring of both patches, we can catalogue all prime ideals of projective curve C. For example, if C is a nonsingular hyperelliptic curve of odd degree, i.e.

$$f_C = Y^2 Z^{2g-1} - X^{2g+1} - a_{2g} X^{2g} Z - \cdots - a_0 Z^{2g+1},$$

then the points at infinity correspond to those with $Z = 0$, for which $(0:1:0)$ is the only such projective point. Thus the list of prime ideals consist of the primes in the coordinate ring of $\hat{C}$ plus one additional prime, namely $(X/Y - 0, Z/Y - 0)$ on the affine patch $Y = 1$, which corresponds to the ideal which vanishes strictly on the one point at infinity. In particular, we take such a hyperelliptic curve to correspond to an affine curve $\hat{C}$ (on the standard affine patch) of the form $y^2 = f_0(x)$, with $f_0(x) \in \mathbb{F}_q[x]$, a polynomial of odd degree with distinct roots.

Definition 1.2. A **divisor** on curve C is a formal linear combination $D = \sum r_i \mathfrak{p}_i$ with $r_i \in \mathbb{Z}$, $\mathfrak{p}_i$ a nonzero prime ideal, and only finitely many of the r_i's are nonzero.

A divisor is **positive** if $r_i \geq 0$ for all i. This is also frequently called **effective** in algebraic geometric literature. The degree of $\mathfrak{p}$ is the degree of the extension $[\mathbb{F}_q[C]/\mathfrak{p} : \mathbb{F}_q]$. The **degree** of a divisor is given by $\deg D = \sum r_i \deg \mathfrak{p}_i$.

We let $\mathbb{F}_q(\hat{C})$ signify the ring of meromorphic functions on the affine curve $\hat{C}$, which is the fraction field of the coordinate ring. If $f \neq 0 \in \mathbb{F}_q(\hat{C})$, then we can define the order of f with respect to prime $\mathfrak{p}$, denoted $\mathrm{ord}_{\mathfrak{p}}(f)$.

Definition 1.3. We first observe that for $\mathfrak{p}$, a prime ideal in $\mathbb{F}_q[\hat{C}]$, we can define the **localization** with respect to $\mathfrak{p}$ as

$$\mathbb{F}_q[\hat{C}]_{\mathfrak{p}} = \left\{ \frac{g}{h} : g, h \in f_q[\hat{C}], \quad h \notin \mathfrak{p} \right\}.$$

Here, we really mean this set modulo equivalence of equal fractions. In other words, prime ideal $\mathfrak{p}$ signifies a collection of affine points of C since $\mathbb{F}_q$ is not algebraically closed, and $\mathbb{F}_q[\hat{C}]_{\mathfrak{p}}$ equals the set of rational functions, up to equivalence, which do not have a pole on the set corresponding to $\mathfrak{p}$. $\mathbb{F}_q[\hat{C}]_{\mathfrak{p}}$ is a local ring, which means that there is a unique nonzero prime ideal, namely $\mathfrak{p}$. Thus, any $f \in \mathbb{F}_q(\hat{C})$ can be written as a Laurent series in terms of t, a generator of $\mathfrak{p}$, which is referred to as a **local parameter**. A Laurent series is simply a power series which might start with a negative exponent. Furthermore, the lowest power of t appearing in this

Laurent series is a well-defined integer which doesn't depend on the choice of t, only depends on $\mathfrak{p}$. We define $\mathrm{ord}_{\mathfrak{p}}(f)$ as this integer for expressing element f in terms of the local ring $\mathbb{F}_q[\hat{C}]_{\mathfrak{p}}$. Note that this order is ≥ 0 if $f \in \mathbb{F}_q[\hat{C}]_{\mathfrak{p}}$ and < 0 otherwise. This is known as a valuation of the **discrete valuation ring** $\mathbb{F}_q[\hat{C}]_{\mathfrak{p}}$.

Furthermore, for $f \in \mathbb{F}_q(\hat{C}), f \neq 0$, then we can define a corresponding divisor $(f) = \sum \mathrm{ord}_{\mathfrak{p}}(f) \cdot \mathfrak{p}$. We call such a divisor a **principal divisor**. Note that if $\mathfrak{p}$ is a prime ideal of degree one, e.g. $(X-a, Y-b)$ for $a, b \in \mathbb{F}_q$, then $\mathrm{ord}_{\mathfrak{p}}(f)$ is defined as the order of the zero or pole that rational function f has at the point (a, b). However, the nice thing about this definition in terms of primes, which generalizes the notion of the order of a function at a point, is that we gain information about all the extensions of $\mathbb{F}_q$ as well. A standard result regarding the divisor of a function is a restriction on its degree.

Proposition 1.4. *If f is a nonzero meromorphic function in $\mathbb{F}_q(\hat{C})$, then the degree of (f) is zero.*

Proof. See [Ful89, Ch. 8]. □

Now that we have a way of attaching a divisor to a rational function (with coordinates in $\mathbb{F}_q$), we are ready to state and use the Riemann-Roch Theorem to better understand what these divisors look like. Before discussing this theorem however, we take an interlude to discuss a combinatorialist's definition of prime divisor.

1.2 Combinatorial definition of primes

Recall that we defined a divisor on curve C over field k as a formal linear combination $D = \sum r_i \mathfrak{p}_i$ with $r_i \in \mathbb{Z}$, $\mathfrak{p}_i$ a nonzero prime ideal in $k[C]$, and only finitely many of the r_i's are nonzero. To get some intuition for this definition of prime ideals, we note that if k is an algebraically closed field instead of $\mathbb{F}_q$, then the only prime ideals on an affine curve would be the maximal ones, $(X-a, Y-b)$ s.t. $a, b \in k$. (The nonsingular projective curve always has exactly one extra prime ideal, namely the maximal ideal which vanishes solely at the point at infinity.)

Prime ideals exactly correspond to points on $C(k)$ when k is algebraically closed, and thus all primes are of degree one. Further divisors of such curves can be written as $D = \sum r_i \cdot P_i$ where P_i is a point of C over k. The degree of D is simply given as $\sum r_i$.

Even though we require k algebraically closed for the above definition of divisors in terms of points, rather than primes, we now can use this observation and adapt this definition so it works even when k is not algebraically closed, e.g. $k = \mathbb{F}_q$. For this, we define an important map from the curve back to itself. We define this map on the curve over an algebraic closure $\overline{\mathbb{F}_q} = \overline{\mathbb{F}_p}$ of $\mathbb{F}_q$ which contains all algebraic extensions of $\mathbb{F}_q$. (In particular $\overline{\mathbb{F}_q} \cong \bigcup_{k\geq 1} \mathbb{F}_{q^k}$.)

Definition 1.5. Given a projective curve C defined over $\mathbb{F}_q$, the **Frobenius map**

$$\pi : C(\overline{\mathbb{F}_q}) \to C(\overline{\mathbb{F}_q})$$

denotes the point obtained by raising each of the coordinates to the qth power. We can think of this action in terms of $\mathbb{P}^2$, i.e. $(X : Y : Z) \mapsto (X^q : Y^q : Z^q)$, noting that

$$(\lambda X : \lambda Y : \lambda Z) \mapsto (\lambda^q X^q : \lambda^q Y^q : \lambda^q Z^q) = (\lambda X^q : \lambda Y^q : \lambda Z^q)$$

for any scalar $\lambda \in \mathbb{F}_q$. Alternatively, it is clear that $\pi\Big((0 : 1 : 0)\Big) \mapsto (0 : 1 : 0)$, i.e. the point at infinity is a fixed point of π, and on the affine patch the Frobenius map acts as $\pi\Big((x, y)\Big) \mapsto (x^q, y^q)$.

Proposition 1.6. *The above definition is well defined, in particular, if $P \in C$, i.e. $P \in \mathbb{P}^2$ satisfies $f_C(P) = 0$ then $Q = \pi(P)$ also satisfies $f_C(Q) = 0$. Furthermore, $P \in C(\overline{\mathbb{F}_q})$ is a fixed point of the kth power of π if and only if $P \in C(\mathbb{F}_{q^k})$.*

Proof. Let $P = (X_0, Y_0, Z_0)$ be a point on $C(\overline{\mathbb{F}_q})$. For $\alpha, \beta \in \overline{\mathbb{F}_q}$ we have the property

$$(\alpha\beta)^q = \alpha^q\beta^q \quad \text{and} \quad (\alpha + \beta)^q = \alpha^q + \beta^q.$$

Thus a polynomial $f_C(x, y, z)$ satisfies $\Big(f_C(X_0, Y_0, Z_0)\Big)^q = f_C(X_0^q, Y_0^q, Z_0^q)$. In particular, if $f_C(P) = 0$, so does $f_C\Big(\pi(P)\Big)$. Additionally, $\alpha^{q^k} = \alpha$ if and only if $\alpha \in \mathbb{F}_{q^k}$ and thus $\pi^k(P) = (X_0^{q^k}, Y_0^{q^k}, Z_0^{q^k}) = (X_0, Y_0, Z_0) = P$ if and only if $P \in \mathbb{F}_{q^k}$. $\square$

As a consequence of this map, we can think of primes on a curve in a more combinatorial way as the primitive sets of $\overline{\mathbb{F}_q}$-points such that the set is invariant under the Frobenius map. Here, such a set S is **primitive** if there is no π-invariant nonempty proper subset of S. It is clear that if a point has coordinates in $\mathbb{F}_q$, it is fixed by the Frobenius map. This corresponds to the fact that the point is the geometric analogue of the maximal ideal $(x - a_x, y - a_y)$, or in the case of the point at infinity, $(0 : 1 : 0) \leftrightarrow (X - 0, Z - 0)$.

Otherwise, the collection of points $\{P_1, \ldots, P_k\}$ will be such that there exists a univariate $\mathbb{F}_q$-polynomial $g(x)$ whose roots correspond to the $x-$coordinates of points P_1 through P_k. In particular, we obtain the following.

Lemma 1.7. *If $S = \{P_1, P_2, \ldots, P_k\}$ is a π-invariant primitive set with $P_1 = (x_1, y_1), \ldots, P_k = (x_k, y_k)$ then $g(x) = (x - x_1)(x - x_2) \cdots (x - x_k)$ is an irreducible polynomial in $\mathbb{F}_q[x]$ on which P_1 through P_k vanish.*

Proof. It is clear that P_1 through P_k vanish on $g(x)$ by construction. Since the Frobenius map π leaves $S = \{P_1, P_2, \ldots, P_k\}$ invariant, it therefore induces a permutation σ of these points. In particular

$$\begin{aligned} g(x)^q &= (x^q - x_1^q)(x^q - x_2^q) \cdots (x^q - x_k^q) \\ &= (x^q - x_\sigma 1)(x^q - x_\sigma 2) \cdots (x^q - x_\sigma k) \\ &= (x^q - x_1)(x^q - x_2) \cdots (x^q - x_k) = g(x^q) \end{aligned}$$

and thus $g(x)$ has coefficients in $\mathbb{F}_q$. Furthermore, since set S was assumed to be primitive, polynomial $g(x)$ is irreducible. $\square$

Thus P_1 through P_k will both lie on the locus of f_C as well as $g(x)$. Notice however that $V\Big(g(x)\Big)$, the variety for ideal $(g(x))$, i.e. the set of points of C which vanish on $g(x)$ will not generally recover set S, but rather a superset of S. This is due to the fact that not all prime ideals are principal. However for any such S, there exist additional bivariate polynomials $h_1(x, y)$, $h_2(x, y), \ldots, h_r(x, y)$ such that S does in fact equal $V\Big(g(x), h_1(x,y), h_2(x,y), \ldots, h_r(x,y)\Big)$. For example, in the case $C = \mathbb{P}^1$, all primes correspond to irreducible polynomials in $\mathbb{F}_q[x]$ since $\mathbb{F}_q[x]$ is a principal ideal domain. On the other hand, in the hyperelliptic case, there are at most two points on $C(\overline{\mathbb{F}_q})$ with the same x-coordinates. Thus

$$\begin{aligned} V(g(x)) &= V\Big((x-x_1)(x-x_2)\cdots(x-x_k)\Big) \\ &= \Big\{(x_1,y_1),(x_1,-y_1),(x_2,y_2),(x_2,-y_2),\ldots,(x_k,y_k),(x_k,-y_k)\Big\}. \end{aligned}$$

Here we have abused notation, and have listed special points of the form $(x_i, 0)$ twice, even though they only appear once in $V(g(x))$.

Proposition 1.8. *In the hyperelliptic case (and in particular char $k \neq 2$), $V(g(x))$ is either a prime divisor or splits into exactly two prime divisors via*

$$\begin{aligned} V(g(x)) &= \{(x_1,y_1),(x_2,y_2),(x_3,y_3),\ldots,(x_k,y_k)\} \\ &\cup \{(x_1,-y_1),(x_2,-y_2),(x_3,-y_3),\ldots,(x_k,-y_k)\}. \end{aligned}$$

In particular all prime divisors of hyperelliptic curves (char $k \neq 2$) arise in this way.

Proof. Assume $S = \{(x_1,y_1),(x_2,y_2),(x_3,y_3),\ldots,(x_k,y_k)\}$ is a prime divisor, where we do not assume the x_i's are necessarily distinct. Since S is a primitive set, the point (x_i, y_i) does not appear twice in this list, and so even though the x_i's are not necessarily distinct, we cannot have i and j so that $x_i = x_j$ and $y_i = y_j$ simultaneously. Since a hyperelliptic curve has only at most two points with same x-coordinate, if successive application of the Frobenius map yields $x_i^{q^\ell} = x_i$ and $y_i^{q^\ell} \neq y_i$, this forces $(x_i^{q^{2\ell}}, x_i^{q^{2\ell}}) = (x_i, y_i)$. We thus have two cases:

- 1) $(x_1, y_1) \in \mathbb{F}_{q^k} \times \mathbb{F}_{q^k}$ and all the x_i and y_i are distinct. In this case $V(g(x)) = S \cup \overline{S}$ where $\overline{S}$ is the set by taking the negative of all the y-coordinates.
- 2) $k = 2\ell$ and $(x_1, y_1) \in \mathbb{F}_{q^\ell} \times \mathbb{F}_{q^{2\ell}}$. In this case $V(g(x)) = S$, and every x-coordinate appears twice.

Note that these are the only two cases because $x_1^{q^k} = x_1$ implies $x_1 \in \mathbb{F}_{q^k}$ and if $x_1 \in \mathbb{F}_{q^\ell}$ for $\ell < k/2$ then set S would contain a repeated a point. $\square$

So in particular if $P_1 = (x_1, y_1), \ldots, P_k = (x_k, y_k)$ with no two x-coordinates the same, then by Lagrange interpolation we have a polynomial $L(x)$ with the proper y-coordinates. Explicitly, the polynomial $L(x) = \sum_{j=1}^{k} y_j \prod_{\substack{i=1 \\ i \neq j}}^{k} \frac{x - x_i}{x_j - x_i}$ satisfies $L(x_i) = y_i$ for all $i \in \{1, \ldots, k\}$. Thus we let $h(x, y) = y - L(x)$ and note that in the case $(g(x)) = S \cup \overline{S}$, then depending on our choice of $L(x)$, we have $y - L(x)$ will vanish at either S or $\overline{S}$, but not both.

Thus the Frobenius cycle $\{P_1, \ldots, P_k\}$ is the algebraic set for an ideal of the form $\Big(g(x)\Big)$ or $\Big(g(x), h(x, y)\Big)$ for the hyperelliptic case.

Thus we will sometimes refer to these prime ideals as Frobenius cycles, and take away the algebraic scaffolding and think of primes as these primitive collections. We partition the set of all points on $C(\overline{\mathbb{F}_q})$ into an infinite collection of these primitive subsets. Since all elements $\alpha \in \overline{\mathbb{F}_q}$ are also an element of $\mathbb{F}_{q^k}$ for some k, we also obtain that any point $P \in C(\overline{\mathbb{F}_q})$ lies in $C(\mathbb{F}_{q^k})$ for some k. (Take for example the lowest common multiple of k_1 and k_2 where $P = (\alpha, \beta)$ and $\alpha \in \mathbb{F}_{q^{k_1}}$ and $\beta \in \mathbb{F}_{q^{k_2}}$.) Thus Frobenius cycles will always be of finite length. Thinking of the primes as Frobenius cycles, the degree of $\mathfrak{p} = S = \{P_1, \ldots, P_k\}$ is the number of points in the cycle, i.e. k in this case.

Map π therefore acts as a permutation of the infinite set $C(\overline{\mathbb{F}_q})$ which has fixed points given by the elements of $C(\mathbb{F}_q)$, 2-cycles given by the primes of degree 2, etc. We let I_k denote the number of primitive cycles/prime ideals of degree k. A divisor is a formal linear combination of such primes, and we still define the degree of a divisor, as $\deg D = \sum r_i \deg \mathfrak{p}_i$. However, we can now also view a positive divisor D as a π-invariant (not necessarily primitive) multiset of points in $C(\overline{\mathbb{F}_q})$. (A multiset is a set where repetitions are allowed.) In this case the degree of D is its cardinality as a multiset. We let H_k denote the number of positive divisors of degree k.

1.3 The Riemann-Roch theorem and rationality of the zeta function

We now return to the topic at hand, divisors of functions and zeta functions. Given a rational function $f = g/h$ in lowest terms, where g and h are polynomials in $\mathbb{F}_q[x, y]$, we define the order of point P with respect to f as follows. If P is a zero of f, then its order is the order of vanishing of g at P. If on the other hand, P is a pole of f, then its order is the negative of the order of vanishing of h at P. Otherwise, the order of P with respect to f is defined to be zero. By logic similar to that of Lemma 1.7, we observe if P is a point of order d (with respect to f) then so is $\pi(P)$. Thus using the viewpoint of the last section, the valuation at a prime $\mathfrak{p}$, i.e. Frobenius cycle S, can be defined as the order of any one of the representative points $P_i \in S$. This definition also agrees with $ord_{\mathfrak{p}}(f)$ using discrete valuations.

For any divisor D, we define the vector space $L(D)$ to be

$$\left\{ f \in \mathbb{F}_q(\hat{C}), f \neq 0 : (f) + D \text{ is positive} \right\} \cup \left\{ 0 \right\}.$$

Considering the case of genus g curves over a not necessarily algebraically closed field k, the Riemann-Roch Theorem states:

Theorem 1.9. *(Riemann-Roch) For any divisor D, $L(D)$ is a finite dimensional vector space over field k. Furthermore, if deg $D < 0$ then $\dim L(D) = 0$ and otherwise*

$$\dim L(D) = \deg(D) + 1 - g - \dim L(K - D)$$

where K is the divisor corresponding to the canonical class, which has degree $2g - 2$ in the case of a genus g curve. In particular, if deg $D > 2g - 2$, then

$$\dim L(D) = \deg(D) + 1 - g.$$

This theorem is proven several ways in the literature, either via adeles or as a corollary of Serre Duality. See for example [Har77, Ch. 3], or [Lan82, Ch. 1]. The upshot of the the Riemann-Roch theorem is that it is true regardless of the choice of field k, and in particular we can let $k = \mathbb{F}_q$ as we have been doing. Consequently, we can immediately translate a fact about the dimension of a vector space into a fact about the number of elements in such a space. Namely a $d-$dimensional space over $\mathbb{F}_q$ has q^d elements. This allows us to count the number of positive divisors of a certain degree by splitting up the problem by linear equivalence classes.

Let $P(D)$ denote the set of all positive divisors D' that are linearly equivalent to D, i.e. $D' = D + (f)$ for some meromorphic function f.

Lemma 1.10. *The set of positive divisors equivalent to D, also called the linear system of divisor D, is a projective space of dimension equal to dim $L(D) - 1$.*

Proof. Notice there is a surjective map $\phi_D : (L(D) - \{0\}) \to P(D)$ via $\phi(f) = (f) + D$. This map also has the property that $\phi(g) = \phi(h)$ if and only if there exists $c \in \mathbb{F}_q^\times$ such that $g = c \cdot h$, since $(g) = (h)$ only if $g = c \cdot h$. Thus

$$\overline{\phi_D} : (L(D) - \{0\}) \Big/ \mathbb{F}_q^\times \to P(D)$$

is a bijection. $\square$

Assuming $\dim L(D) = m \geq 1$, this bijection implies

$$|P(D)| = \frac{q^m - 1}{q - 1} = 1 + q + q^2 + \cdots + q^{m-1}.$$

Hence we obtain that

$$H_m = \sum_{\overline{D} \in Pic^m} \frac{q^{\dim L(D)} - 1}{q - 1} \tag{1.1}$$

where H_m equals the number of positive divisors of degree m, and the sum is taken over all linear equivalence classes of degree m. (Note that since a principal divisor, the divisor of a function, always has degree zero, it makes sense to discuss the degree of a linear equivalence class.) We let Pic denote the **divisor class group**, i.e. the quotient group all divisors modulo principal ones. Let Pic^m denote the set of all equivalence classes of degree m divisors, and let D be a representative of class $\overline{D}$. To understand this quantity H_m better, we construct an ordinary generating function for it, i.e. $\sum_{m\geq 0} H_m T^m$. We will shortly see that this generating function is in fact the zeta function $Z(C,T)$ of the curve C. The Riemann-Roch Theorem will be used to prove the rationality of this function.

Recall our definitions of primes and points on a curve. More precisely, I_k is the number of Frobenius cycles of C of length k, i.e. a collection of k distinct pairs in $\mathbb{F}_{q^k} \times \mathbb{F}_{q^k}$ of the form

$$\left\{(\alpha,\beta),(\alpha^q,\beta^q),\ldots,(\alpha^{q^{k-1}},\beta^{q^{k-1}})\right\} \quad \text{with} \quad f_C(\alpha,\beta)=0.$$

We will let N_k denote the number of points on the curve C, defined over $\mathbb{F}_q$, over finite field $\mathbb{F}_{q^k}$. These two quantities are actually related in a simple way.

Lemma 1.11. *For all $m, d \geq 1$ we have*

$$N_m = \sum_{d|m} d \cdot I_d.$$

Proof. We let $\{\mathfrak{p}\}$ be the collection of prime ideals in the function field $\mathbb{F}_q(C) = \mathbb{F}_q[X,Y,Z] \Big/ (f_C)$, where f_C is the defining equation of curve C over $\mathbb{P}^2$. Note that $P = (a:b:1) \in C$ is a point over $\mathbb{F}_{q^m}$ if and only if $\pi^m(P) = P$, where π is the Frobenius map. Consequently, $d|m, P \in \mathbb{P}(\mathbb{F}_{q^d})$ implies that P also in $\mathbb{F}_{q^m}$.

The points of purely degree m (whose coordinates are not contained in any smaller subfield) will be contained in some Frobenius cycle of length m, and in fact the Frobenius cycles of length m will partition the space of such points. Since each such cycle has m points on it, there are $m \cdot I_m$ purely $\mathbb{F}_{q^m}$ points on C where I_m is the number of $m-$cycles. By summing up the number of points of purely degree d for $d|m$, we obtain the desired identity. $\square$

Note that by Möbius Inversion, we get a formula for the I_m's in terms of N_d's as well:

$$I_m = \frac{1}{m}\sum_{d|m} \mu(m/d) N_d$$

where

$$\mu(n) = \begin{cases} 0 \text{ if } n \text{ contains a square} \\ (-1)^k \text{ if } n \text{ is squarefree with } k \text{ prime factors} \end{cases}.$$

Definition 1.12. The **zeta function,** or more precisely the Hasse-Weil zeta function for a nonsingular projective algebraic variety, is an exponential generating function for the sequence $\{N_m\}$ given by

$$Z(C,T) = \exp\left(\sum_{m=1}^{\infty} N_m \frac{T^m}{m}\right). \tag{1.2}$$

Theorem 1.13. *We can also express the zeta function is a number of equivalent ways.*

$$\begin{aligned} Z(C,T) &= \prod_{\mathfrak{p}} \frac{1}{1-T^{\deg \mathfrak{p}}}, \quad \mathfrak{p} \text{ is a prime} \\ &= \prod_{k\geq 1} \left(\frac{1}{1-T^k}\right)^{I_k} \\ &= \sum_{m=0}^{\infty} (\#\text{ positive divisors on } C \text{ of deg } m)\ T^m = \sum_{m=0}^{\infty} H_m T^m. \end{aligned}$$

Proof. By Lemma 1.11, $N_m = \sum_{d|m} d\cdot I_d$ where $d\cdot I_d$ equals the number of points on C over $\mathbb{F}_{q^d}$ which are not present over any smaller subfield. This allows us to rewrite $\sum_{m=1}^{\infty} N_m \frac{T^m}{m}$, using the notation $\chi(Expression)$, which equals 1 if $Expression$ is true and equals 0 otherwise.

$$\begin{aligned} \sum_{m=1}^{\infty} N_m \frac{T^m}{m} &= \sum_{m=1}^{\infty} \sum_{d|m} d\cdot I_d \frac{T^m}{m} = \sum_{d=1}^{\infty} d\cdot I_d \sum_{m=1}^{\infty} \frac{T^m}{m}\chi(d|m) \\ &= \sum_{d=1}^{\infty} d\cdot I_d \sum_{k=1}^{\infty} \frac{T^{dk}}{dk} = \sum_{d=1}^{\infty} I_d \cdot \sum_{k=1}^{\infty} \frac{T^{dk}}{k} \\ &= \sum_{d=1}^{\infty} \log\left(\frac{1}{(1-T^d)^{I_d}}\right) = \sum_{\mathfrak{p}} \log\left(\frac{1}{1-T^{\deg \mathfrak{p}}}\right). \end{aligned}$$

By taking the exponential of both sides we obtain

$$Z(C,T) = \prod_{k\geq 1} \left(\frac{1}{1-T^k}\right)^{I_k} = \prod_{\mathfrak{p}} \frac{1}{1-T^{\deg \mathfrak{p}}}, \quad \mathfrak{p} \text{ is a prime.}$$

Now, using the fact that

$$\frac{1}{1-T^{\deg \mathfrak{p}}} = (1 + T^{\deg \mathfrak{p}} + T^{2\deg \mathfrak{p}} + \dots),$$

we multiply out this generating function and write it as a sum, getting the terms corresponding to all possible nonnegative linear combinations of primes. Since each of these terms contributes T^m where m is the degree of the linear combination (i.e. divisor), this is exactly the generating function for the H_m's. More specifically,

$$Z(C,T) = \prod_{\mathfrak{p}} \frac{1}{1 - T^{\deg \mathfrak{p}}}$$

and so

$$Z(C,T)\Big|_{T^m} = \prod_{\mathfrak{p} \text{ of degree } \leq m} \frac{1}{1 - T^{\deg \mathfrak{p}}}\Big|_{T^m}.$$

There are a finite number of primes of degree at most m, and so enumerating these as $\mathfrak{p}_1, \mathfrak{p}_2, \ldots, \mathfrak{p}_N$, this expression gives

$$Z(C,T)\Big|_{T^m} = \sum_{n_1 \geq 0} \sum_{n_2 \geq 0} \cdots \sum_{n_N \geq 0} \chi\Big(n_1|\mathfrak{p}_1| + n_2|\mathfrak{p}_2| + \cdots + n_N|\mathfrak{p}_N| = m\Big) = H_m.$$

□

We now proceed to prove a result due to Weil [Wei48].

Theorem 1.14 (Rationality)**.**

$$Z(C,T) = \frac{(1-\alpha_1 T)(1-\alpha_2 T)\cdots(1-\alpha_{2g-1}T)(1-\alpha_{2g}T)}{(1-T)(1-qT)}$$

for complex numbers α_i's, where g is the genus of the curve C. Furthermore, the numerator of $Z(C,T)$, which we will denote as $L(C,T)$, has integer coefficients since the H_m's, have a combinatorial interpretation.

We have already seen, from (1.1), that we can also describe $Z(C,T) = \sum_{m=0}^{\infty} H_m$ as

$$\sum_{m=0}^{\infty} \sum_{\overline{D} \in Pic^m(C)} \left(\frac{q^{\dim L(D)} - 1}{q-1}\right) T^m.$$

Using this expression will allow us to apply Riemann-Roch to prove that $Z(C,T)$ is a rational expression. To get started, we need a couple auxiliary results.

Lemma 1.15. *Let divisor D of curve C over field k have degree d. If $d < 0$ then $L(D) = 0$. Otherwise, the dimension of $L(D)$ satisfies the bounds*

$$0 \leq \dim L(D) \leq d + 1.$$

Proof. We follow [Was03, Ch. 11]. Firstly, if degree $D < 0$ but $L(D) \neq 0$, then there exists a nonzero rational function f such that $(f) + D \geq 0$. However, since principal divisors have degree zero and degree is linear, this inequality implies deg $D =$ deg $((f) + D) \geq 0$, a contradiction. Thus we assume we are in the case of a divisor with nonnegative degree. We prove the bound by induction. If $D = 0$, then $L(D)$ is the vector space of rational functions which have no zeros or poles. As in [Ful89, Ch. 8], the only such functions are the constant functions. Thus $\dim L(0) = 1$.

Now assume temporarily that k is algebraically closed. We can obtain any divisor from the zero divisor by adding or subtracting a point at a time. For any point P we consider the quotient space

$$L(D + P) \Big/ L(D).$$

This vector space has dimension 0 or 1 by the following argument. Assume $f_1, f_2 \in L(D + P) \Big/ L(D)$ and let $-n$ be the multiplicity of point P in $D + P$. The fact that f_1 and $f_2 \in L(D + P)$ means that the order of P must be at least n for both f_1 and f_2, but since f_1 and $f_2 \notin L(D)$ by assumption, we must have equality, i.e. functions f_1 and f_2 must both have order exactly n at P. We let u be a local parameter at P which enables us to write

$$f_1 = u^n g_1 \qquad \text{and} \qquad f_2 = u^n g_2$$

such that g_1 and g_2 do not vanish or have a pole at P. Thus $g_1(P) = c_1$ and $g_2(P) = c_2$ are nonzero elements of k, and observe that function

$$c_2 f_1 - c_1 f_2 = u^n (c_2 g_1 - c_1 g_2)$$

vanishes at point P and so $c_2 f_1 - c_1 f_2$ has order greater than n at P, hence $c_2 f_1 - c_1 f_2 \in L(D)$ and so any two elements $f_1, f_2 \in L(D + P) \Big/ L(D)$ are linearly dependent. Thus every time we add (subtract) a point to divisor D, we increase (resp. decrease) the dimension of $L(D)$ by at most one. We now take away the restriction of algebraically closed by recalling that we can construct any divisor by subsequent additions (or subtractions) of prime divisors. However, adding a prime divisor of degree r is tantamount to adding r points, which can change the dimension by at most r, and so we get the desired bounds even when k is not algebraically closed. $\square$

In fact there is a stronger result in the literature, Clifford's Theorem [Har77, pg. 343], which states

$$\dim L(D) > d + 1 - g \Rightarrow \dim L(D) \leq \frac{1}{2}d + 1$$

(with equality if and only if $D = 0, K$, or C is hyperelliptic and D is a multiple of a class D_2 satisfying $\deg D_2 = 2, \dim D_2 = 2$), but Lemma 1.15 will actually be sufficient for our needs.

Lemma 1.16. *$\#Pic^m(C) = \#Pic^0(C)$ for all $m \in \mathbb{Z}$.*

Proof. Recall that two divisors D_1 and D_2 are equivalent if and only if for some $f \in \mathbb{F}_q[C]$ we have $D_2 = D_1 + (f)$. Now from the Riemann-Roch Theorem we derive that if $deg(D) = m > g$ then

$$\dim\ L(D) \geq m + 1 - g > 1,$$

and in particular there is an $f \in L(D)$ such that

$$D' \ = \ (f) + D \geq 0\,.$$

Thus in the equivalence class of D there is a positive divisor, and a trivial bound for $|Pic^m|$ in this case is H_m. Moreover, note that if the number of divisor classes varies with m, i.e. for $m \neq m'$ we have

$$Pic^m = \left\{D_1^{(m)}, D_2^{(m)}, \ldots, D_{r_m}^{(m)}\right\} \quad \text{and} \quad Pic^{m'}(C) = \left\{D_1^{(m)}, D_2^{(m')}, \ldots, D_{r_{m'}}^{(m')}\right\}$$

then denoting by P_∞ the point at infinity we have that

$$D_1^{(m)} + (m' - m)P_\infty,\ D_2^{(m)} + (m' - m)P_\infty,\ \ \ldots,\ D_{r_m}^{(m)} + (m' - m)P_\infty$$

are inequivalent divisors of degree m'. This gives

$$|Pic^m| \leq |Pic^{m'}|.$$

The reverse inequality is obtained by considering the divisors

$$D_1^{(m')} + (m - m')P_\infty,\ D_2^{(m')} + (m - m')P_\infty,\ \ldots,\ D_{r_{m'}}^{(m')} + (m' - m')P_\infty.$$

Thus the cardinality of Pic^m is finite and constant for all m, completing our argument.

$\square$

Proof of Theorem 1.14. Armed with Lemmas 1.15 and 1.16 , we let $A_{i,j}$ equal the number of divisor classes $\overline{D}$ which satisfy $\deg(\overline{D}) = i$ and $\dim L(\overline{D}) = j$. By Riemann-Roch,

$$A_{i,j} = 0 \text{ if } j < i + 1 - g.$$

Clearly, $\sum_{j \geq 0} A_{i,j} = Pic^i$, the number of classes of degree i, since the $A_{i,j}$'s are counting the divisor classes more finely. By Lemma 1.15,

$$A_{i,j} = 0 \text{ if } j > i + 1$$

and so we can write more specifically $\sum_{j=0}^{i+1} A_{i,j} = Pic^i$. We therefore derive via algebra:

$$\begin{aligned} Z(C,T) &= \sum_{m=0}^{g-1} \Big(A_{m,1} + A_{m,2}(q+1) + \cdots + A_{m,m+1}(q^m + q^{m-1} + \cdots + q + 1) \Big) T^m \\ &+ \sum_{m=g}^{2g-2} \left(A_{m,m+1-g} \left(\frac{q^{m+1-g} - 1}{q-1} \right) + \cdots + A_{m,m+1} \left(\frac{q^{m+1} - 1}{q-1} \right) \right) T^m \\ &+ \sum_{m=2g-1}^{\infty} |Pic^m| \cdot \left(\frac{q^{m+1-g} - 1}{q-1} \right) T^m. \end{aligned}$$

By the observation that $m + 1 - i \geq m + 1 - g$ for all $0 \leq i \leq g$, we can change the indices of the last summand and subtract its terms from that of the second summand. This operation reduces the expression to

$$\begin{aligned}
Z(C,T) &= \sum_{m=0}^{g-1}\Big(A_{m,1} + A_{m,2}(q+1) + \cdots + A_{m,m+1}(q^m + q^{m-1} + \cdots + q + 1)\Big)T^m \\
&+ \sum_{m=g}^{2g-2}\Big(A_{m,m+1-(g-1)}q^{m+1-g} + \cdots + A_{m,m+1}(q^{m+1-g} + q^{m+2-g} + \cdots + q^{m+1})\Big)T^m \\
&+ \sum_{m=g}^{\infty}|Pic^m|\cdot\left(\frac{q^{m+1-g}-1}{q-1}\right)T^m.
\end{aligned}$$

We can reduce this further via

$$A_{i,j} = A_{2g-2-i,j-i+g-1} \tag{1.3}$$

$$H_m = A_{m,1} + A_{m,2}(q+1) + \cdots + A_{m,m+1}(q^m + \cdots + q + 1) \tag{1.4}$$

The reciprocity (1.3) comes from the second statement of Riemann-Roch,

$$\dim L(D) = \deg(D) + 1 - g - \dim L(K-D),$$

and the fact that the canonical class K, satisfies $\deg L(K) = 2g-2$. The second identity, (1.4), comes directly from the definitions of H_m and $A_{m,i}$ along with the bounds of Lemma 1.15. Letting $n = 2g-2-m$, and applying equation (1.3) yields

$$\begin{aligned}
Z(C,T) &= \sum_{m=0}^{g-1} H_m T^m \\
&+ \sum_{n=0}^{g-2}\Big(A_{n,1}q^{g-1-n} + \cdots + A_{n,g}(q^{g-1-n} + q^{g-n} + \cdots + q^{2g-1-n})\Big)T^{2g-2-n} \\
&+ \sum_{m=g}^{\infty}|Pic^m|\cdot\left(\frac{q^{m-g+1}-1}{q-1}\right)T^m.
\end{aligned}$$

Since $A_{n,j} = 0$ for $j > n+1$ by Lemma 1.15, we reduce this to

$$\begin{aligned}
Z(C,T) &= \sum_{m=0}^{g-2} H_m\Big(T^m + q^{g-1-m}T^{2g-2-m}\Big) + H_{g-1}T^{g-1} \\
&+ \sum_{m=g}^{\infty}|Pic^m|\cdot\left(\frac{q^{m-g+1}-1}{q-1}\right)T^m.
\end{aligned}$$

To finish our analysis, we use Lemma 1.16 which describes the number of divisor classes of various degrees. Based on Lemma 1.16, we can actually replace the superscript m from Pic^m with zero since the number of divisor classes (of a certain degree) actually does not depend on the degree. Thus we can rewrite the zeta function as

$$\begin{aligned} Z(C,T) &= \sum_{m=0}^{g-2} H_m\Big(T^m + q^{g-1-m}T^{2g-2-m}\Big) + H_{g-1}T^{g-1} \\ &+ \frac{|Pic^0|\cdot T^g}{(1-T)(1-qT)} \end{aligned}$$

and have thus proven the rationality of the generating function $Z(C,T)$. Even better, we can write

$$Z(C,T) = W(T) + \frac{|Pic^0|\cdot T^g}{(1-T)(1-qT)}$$

where $W(T)$ equals $\sum_{m=0}^{g-1} H_m T^m + \sum_{m=g}^{2g-2} H_{2g-2-m}q^{m-g+1}T^m$, a polynomial of degree $2g-2$. Consequently $Z(C,T)$ is a rational function with the numerator and denominator as described by the theorem. $\square$

This method of proof also allows us to obtain an explicit expression for $|Pic^0|$ by taking the coefficient of T^g in the latest expression of $Z(C,T)$.

Corollary 1.17.

$$|Pic^m| = H_g - qH_{g-2}$$

for all $m \geq 0$.

Proof. Since $Z(C,T)\Big|_{T^g} = H_g$ by definition of the H_k's, by comparing this quantity with the coefficient of T^g on the right-hand-side of (1.5) we obtain $H_g = qH_{g-2} + |Pic^m|$ and thus the corollary is proved. $\square$

In fact we can write $Z(C,T)$ in a nice compact form which highlights a functional equation satisfied by $Z(C,T)$.

Theorem 1.18.

$$\begin{aligned} Z(C,T) &= \sum_{m=0}^{g-2} H_m T^m + H_{g-1}T^{g-1} + \sum_{m=g}^{2g-2} H_{2g-2-m}q^{m-g+1}T^m \\ &+ \frac{(H_g - qH_{g-2})T^g}{(1-T)(1-qT)}. \end{aligned}$$

Furthermore,

$$Z(C,T) = q^{g-1}T^{2g-2}Z(C,1/qT).$$

Proof. We have

$$\begin{aligned} q^{g-1}T^{2g-2}Z(C,1/qT) &= \sum_{m=0}^{g-2} H_m q^{g-1-m}T^{2g-2-m} + H_{g-1}q^{(g-1)-(g-1)}T^{(2g-2)-(g-1)} \\ &+ \sum_{m=g}^{2g-2} H_{2g-2-m}q^{(m-g+1)+(g-1)-m}T^{2g-2-m} \\ &+ \frac{(H_g - qH_{g-2})q^{(g-1)-g}T^{(2g-2)-g}}{(1-\frac{1}{qT})(1-\frac{1}{T})}. \end{aligned}$$

The rational expression can be simplified by multiplying top and bottom by $(-qT)(-T)$ and after changing indices by letting $m' = 2g-2-m$, the two summands switch roles. Thus, we recover $Z(C,T)$, as was to be shown. ☐

The functional equation also tells us that the α_i's come in pairs that multiply to q.

Corollary 1.19. *Up to reordering of the α_i's, we have for $1 \leq i \leq g$, $\alpha_i\alpha_{g+i} = q$.*

Proof. By Theorems 1.14 and 1.18 we can write

$$Z(C,T) = \frac{(1-\alpha_1 T)\cdots(1-\alpha_{2g}T)}{(1-T)(1-qT)}$$

as $q^{g-1}T^{2g-2}Z(C,1/qT)$ which, after multiplying top and bottom by $(-qT)(-T)$, equals

$$q^g T^{2g}\frac{(1-\frac{\alpha_1}{qT})\cdots(1-\frac{\alpha_{2g}}{qT})}{(1-T)(1-qT)}.$$

Multiplying and dividing through by the product $\prod_{i=1}^{2g}\frac{-qT}{\alpha_i}$ we obtain

$$Z(C,T) = \frac{\prod_{i=1}^{2g}\alpha_i}{q^g}\cdot\frac{(1-\frac{q}{\alpha_1}T)\cdots(1-\frac{q}{\alpha_{2g}}T)}{(1-T)(1-qT)}. \tag{1.5}$$

Before finishing the proof of this corollary, we spend a moment discussing how we can derive an expression for the numerator of $Z(C,T)$, i.e. $L(C,T)$. Namely, by multiplying through the polynomial portion of the expression from Theorem 1.18 by the quantity $(1-T)(1-qT)$, we obtain

$$\begin{aligned} L(C,T) &= (1-T)(1-qT)\left(\sum_{m=0}^{g-2} H_m T^m + H_{g-1}T^{g-1} + \sum_{m=g}^{2g-2} H_{2g-2-m} q^{m-g+1} T^m\right) \\ &+ \quad (H_g - qH_{g-2})T^g. \end{aligned}$$

In particular, the highest term in $L(C,T)$ is $q^g T^{2g}$, which is the product of all the α_i's. Thus in equation (1.5), the constant in front is in fact one. It follows that the inverse roots have simply been re-ordered, and so for all $1 \leq i \leq 2g$, there exists $1 \leq j \leq 2g$ such that $\alpha_i = q/\alpha_j$. By permuting the α_i's appropriately we get they pair up as claimed. □

1.4 The Weil conjectures

The following four conjectures of Andre Weil [Wei48] (now theorems via Dwork [Dwo60] and Deligne's work [Del74]) were instrumental in the theory of algebraic varieties. In fact these four were proven by Weil for curves, and this work along with that on other examples, including Fermat hypersurfaces, provided him with evidence for the conjectures for varieties in general. Here they are without further adieu.

Theorem 1.20 (The Weil Conjectures). *Let V be a smooth projective variety of dimension n over field $\mathbb{F}_q$. Let $Z(V,T)$ denote the zeta function of V, defined by considering the exponential generating function for the N_k's as defined above for curves. Then*

- *Rationality. $Z(V,T)$ is a rational function of T, i.e. a quotient of polynomials with rational coefficients.*

- *Functional equation. Let E be the self-intersection number of the diagonal Δ of $V \times V$. Then $Z(V,T)$ satisfies a functional equation which will have the form*

$$Z(1/q^n T) = \pm q^{nE/2} T^E Z(V,T).$$

- *Riemann hypothesis.* *It is possible to write*

$$Z(V,T) = \frac{P_1(T)P_3(T)\cdots P_{2n-1}(T)}{P_0(T)P_2(T)\cdots P_{2n}(T)}$$

  where $P_0(T) = 1 - T$, $P_{2n}(T) = 1 - q^n$ *and each of the other* $P_i(T)$*'s are polynomials with integer coefficients which are usually written in factored form* $P_i(T) = \prod(1 - \alpha_{ij}T)$ *where the* α_{ij} *are algebraic integers satisfying* $|\alpha_{ij}| = \sqrt{q^i}$.

- *Betti numbers.* *Given the analogue of the Riemann hypothesis, define the* i*th Betti number* $B_i = B_i(V)$ *to be the degree of the polynomial* $P_i(T)$. *Then the quantity* E *arising in the functional equation satisfies* $E = \sum_{i=0}^{2n} B_i$. *Furthermore, if* V *is obtained from variety* W *defined over an algebraic number ring* R, *by reduction modulo a prime ideal of* R, *then the* $B_i(X)$*'s equal the usual Betti numbers of the topological space thinking of* W *over* $\mathbb{C}$.

An exposition of the proof of these is clearly beyond the scope of this thesis, as Deligne won a Field's Medal for this work. Nonetheless, observe that in the case of curves, we have in fact already written out all the details (except for the Riemann-Roch theorem) for the proof of three of these four conjectures. The remaining one, analogue of the Riemann hypothesis, is the hardest one and in fact is the conjecture that was proved last in the general variety case. While Weil's original proof of the Riemann Hypothesis for curves, i.e. the fact that the $\alpha_{1,j}$'s all satisfy $|\alpha_{1,j}| = \sqrt{q}$, uses intersection theory and the theory of correspondences, a more elementary proof was given by Bombieri [Bom74]. This proof uses only the Riemann-Roch theorem, properties of the Frobenius map, and a couple facts from Galois theory. If one is willing to restrict oneself to the case of hyperelliptic curves, which exist for all genus and include the case of elliptic curves, then one can even avoid the Galois theory. Such a proof is appealing since the Riemann-Roch theorem and Frobenius map can both be described in the combinatorial framework, i.e. as in Section 1.2. While this result will be used later on in Chapter 3, the details of the proof will not, and thus we refer the interested reader to [Bom74] or Chapter 8 of [GM]. For more on the history of the Weil conjectures, see [Har77, Appendix C].

Note that one of the key steps in proving the Weil conjectures was the development of étale cohomology, which provides a sequence of spaces of characteristic zero on which the Frobenius map acts. Given representations of this space, we can think of Frobenius as a linear map, and thus compute the characteristic polynomial

$$\frac{1}{det(I - Fr \cdot T)}. \tag{1.6}$$

In the case of a curve, we need to consider three cohomologies classes: H^0, H^1 and H^2. H^0 and H^2 are both one-dimensional in this case; and furthermore the Frobenius map acts trivially on H^0, and as multiplication by q on H^2. Additionally, for at least the elliptic curve case, H^1 can be thought of as the Tate Module, which is isomorphic to $\mathbb{Z}_\ell \times \mathbb{Z}_\ell$ when ℓ is a prime other than p and $\mathbb{Z}_\ell$ denotes the ℓ-adic integers. We will discuss an elementary formulation of this action in Chapter 3. Additionally, in Chapter 6, we discuss the theory of zeta functions for rational languages where expressions analogous to (1.6) arise, however in this case, they have combinatorial interpretations rather than cohomological ones.

1.5 Introduction to symmetric functions

In the next chapter, we will illustrate how the theory of symmetric functions can be used to analyze the zeta function of an algebraic curve for higher genuses, subsuming elliptic curves as a special case. Because the zeta function of a curve is in fact a rational generating function, and moreover one with quite a nice form, one can use the theory of symmetric functions to analyze coefficients which arise in this generating function. Before giving these applications, we provide the reader with a crash course in symmetric functions.

A symmetric polynomial P in the variables x_1 through x_k is a polynomial with the property that any permutation of the variables $\{x_1, x_2, \ldots, x_k\}$ maps polynomial P back to itself. There are special classes of symmetric polynomials which come up again and again. Since we wish to be able to formally define these expressions in an infinite number of variables or in the abstract, we will work with **symmetric functions** instead, which are these symmetric polynomials with the scaffolding of a specific alphabet taken away. The symmetric functions that we utilize most often in this thesis are the **power symmetric functions** p_k, the **complete homogeneous symmetric functions** h_k, and the **elementary symmetric functions** e_k. Given the alphabet $\{x_1, x_2, \ldots, x_n\}$, each of these can be written as

$$\begin{aligned} p_k &= x_1^k + x_2^k + \cdots + x_n^k, \\ h_k &= \sum_{\substack{0 \le i_1, i_2, \ldots, i_n \le k \\ i_1+i_2+\cdots+i_n = k}} x_1^{i_1} x_2^{i_2} \cdots x_n^{i_n}, \quad \text{and} \\ e_k &= \sum_{1 \le i_1 < i_2 < \cdots < i_k \le n} x_{i_1} x_{i_2} \cdots x_{i_k}. \end{aligned}$$

Theorem 1.21. *The space of symmetric functions in k variables, as a ring, is isomorphic to the polynomial ring $\mathbb{Z}[e_1, e_2, \ldots, e_k]$, $\mathbb{Z}[h_1, h_2, \ldots, h_k]$, or $\mathbb{Q}[p_1, p_2, \ldots, p_k]$.*

Proof. See [Sta99, Ch. 7]. The ring isomorphism between the symmetric functions and the polynomial ring in the e_k's is typically called the fundamental theorem of symmetric functions. However, as this theorem illustrates, there are other important bases for this ring. □

To begin, we will use the following well-known symmetric function identity

Lemma 1.22.

$$\begin{aligned} \prod_{k \in \mathcal{I}} \frac{1}{1 - t_k T} &= \exp\left(\sum_{n \ge 1} p_n \frac{T^n}{n} \right) \\ &= \sum_{n \ge 0} h_n T^n \\ &= \frac{1}{\sum_{n \ge 0} (-1)^n e_n T^n} \end{aligned}$$

where e_n is the nth elementary symmetric function in the variables $\{t_k\}_{k \in \mathcal{I}}$

Proof. See [Sta99, pg. 21, 296]. □

We will also find the techniques of plethysm useful for both motivating the significance of various identities as well as providing their proofs.

Definition 1.23. In general, a **plethystic** substitution of a formal power series $F(t_1, t_2, \dots)$ into a symmetric polynomial $A(x)$, denoted as $A[E]$, is obtained by setting

$$A[E] = Q_A(p_1, p_2, \dots)|_{p_k \to E(t_1^k, t_2^k, \dots)},$$

where $Q_A(p_1, p_2, \dots)$ gives the expansion of A in terms of the power sums basis $\{p_\alpha\}_\alpha$.

Some standard plethystic techniques we will use are given in the next lemma. Note that in this lemma we will utilize ring isomorphism ω which is an involution on the space of symmetric functions. Since an isomorphism is defined by where it sends its' basis elements, it suffices to define

$$\omega(e_i) = h_i, \quad \omega(h_i) = e_i, \quad \text{or equivalently} \;\; \omega(p_i) = (-1)^{i-1} p_i.$$

Lemma 1.24.

$$\begin{aligned} p_n[X+Y] &= p_n[X] + p_n[Y] & (1.7)\\ p_n[XY] &= p_n[X] \cdot p_n[Y] & (1.8)\\ e_n[X+Y] &= \sum_{k=0}^{n} e_k[X] e_{n-k}[Y] & (1.9)\\ h_n[X+Y] &= \sum_{k=0}^{n} h_k[X] h_{n-k}[Y] & (1.10) \end{aligned}$$

If f is a (homogeneous) symmetric function of degree d and u represents a single variable, then

$$\begin{aligned} f[Au] &= f[A]u^d & (1.11)\\ f[-X] &= (-1)^d (\omega f)[X] & (1.12)\\ e_n[X-Y] &= \sum_{k=0}^{n} (-1)^{n-k} e_k[X] h_{n-k}[Y] & (1.13)\\ h_n[X-Y] &= \sum_{k=0}^{n} (-1)^{n-k} h_k[X] e_{n-k}[Y]. & (1.14) \end{aligned}$$

Proof. For a proof, see [Mac95]. We note the (1.7) and (1.8) follow from the definition of plethystic substitution. The other identities are not as obvious, but

(1.9) and (1.10) are actually special cases of the plethystic rule for a basis of symmetric functions known as the Schur functions. We will not use these elsewhere in this dissertation, nonetheless for completeness, we provide the plethystic rule for them:

$$s_\lambda[X+Y] = \sum_{\mu \subseteq \lambda} s_\mu[X] s_{\lambda/\mu}[Y].$$

Also (1.13) and (1.14) are both special cases of (1.12). $\square$

2 The zeta function and symmetric functions

Using the fact that the zeta function of a curve C is defined to be the exponential generating function

$$Z(C,T) = \exp\left(\sum_{k\geq 1} N_k \frac{T^k}{k}\right)$$

which also can be expressed as

$$Z(C,T) = \frac{(1-\alpha_1 T)(1-\alpha_2 T)\cdots(1-\alpha_{2g}T)}{(1-T)(1-qT)}, \tag{2.1}$$

we now apply symmetric function theory to better understand this generating function. We first observe that (1.2) and (2.1) imply the following expression for N_k.

Proposition 2.1. *For all $k \geq 1$ and for any curve C of genus g,*

$$N_k = 1 + q^k - \alpha_1^k - \alpha_2^k - \cdots - \alpha_{2g}^k. \tag{2.2}$$

Proof. Taking the logarithmic derivative of both sides of (2.1) with respect to T, we obtain

$$\begin{aligned}
\frac{\partial}{\partial T}\left(\sum_{k\geq 1} N_k \frac{T^k}{k}\right) &= \frac{\partial}{\partial T}\left(\sum_{i=1}^{2g} \log(1-\alpha_i T) - \log(1-qT) - \log(1-T)\right) = \\
\sum_{k\geq 1} N_k T^{k-1} &= \sum_{i=1}^{2g} \frac{-\alpha_i}{1-\alpha_i T} + \frac{1}{1-T} + \frac{q}{1-qT} \\
&= \sum_{k\geq 1} (1 + q^k - \alpha_1^k - \alpha_2^k - \cdots - \alpha_{2g}^k) T^{k-1}.
\end{aligned}$$

□

We note that expressions (2.2) can be written in plethystic notation as

$$p_k[1 + q - \alpha_1 - \alpha_2 - \cdots - \alpha_{2g}],$$

i.e. the N_k's are an analogue of the power symmetric functions.

2.1 Rewriting the zeta function via plethysm

We now illustrate further applications of this plethystic view of the zeta function. Namely, we observe $Z(C,T) = \exp(\sum_{k\geq 1} \frac{p_k[(1+q-\alpha_1-\alpha_2-\cdots-\alpha_{2g})T]}{k})$ and so using Lemma 1.22, we observe $Z(C,T)$ also equals $\sum_{k=0}^{\infty} h_k[(1+q-\alpha_1-\alpha_2-\cdots-\alpha_{2g})T]$. Comparing with the original definition of $Z(C,T)$ as an ordinary generating function we obtain

Proposition 2.2. *For $m \geq 0$, the number of positive divisors of degree m on genus g curve C satisfies*

$$H_m = h_m[1 + q - \alpha_1 - \alpha_2 - \cdots - \alpha_{2g}].$$

(Note that $H_0 = h_0 = 1$ since the divisor $D = 0$ is considered effective or positive.)

Another useful set of coefficients come from considering the sequence of E_k's obtained by writing the zeta function as a signed reciprocal.

Proposition 2.3. *The sequence of E_k's defined by*

$$Z(C,T) = \frac{1}{\sum_{k=0}^{\infty}(-1)^k E_k T^k}$$

satisfy $E_k = e_k[1 + q - \alpha_1 - \alpha_2 - \cdots - \alpha_{2g}]$.

Just like the N_k's and H_k's, the E_k's also have an algebraic geometric interpretation.

Proposition 2.4. *E_k corresponds to the* signed *number of sets (i.e. without repeats) of prime cycles such that the total number of points is k. Here a set of m different cycles is given weight $(-1)^{m+k}$ in this count. We can also think of this as the* signed *number of positive divisors D of degree k on curve C* such that no prime divisor, or equivalently no point, appears more than once in *D.*

Proof. We write

$$\frac{1}{\sum_{k\geq 0}(-1)^k E_k T^k} = Z(C,T) = \prod_{\mathfrak{p}} \frac{1}{1-T^{\deg \mathfrak{p}}}$$

thus

$$(-1)^k E_k = \prod_{\mathfrak{p}} (1-T^{\deg \mathfrak{p}})\Big|_{T^k} = \sum_{S=\{\mathfrak{p}_1,\ldots,\mathfrak{p}_m\},\quad \deg(\mathfrak{p}_1+\cdots+\mathfrak{p}_m)=k} (-1)^m.$$

Here the right-hand sum is over all sets (not multi-sets) S of prime cycles with total number of points equaling k. Multiplying the left- and right-hand sides by $(-1)^k$ completes the proof. ☐

Remark 2.5. This result is a manifestation of the fact that the reciprocity between h_k's and e_k's is analogous to the reciprocity between *choose* and *multi-choose*, i.e. choice with replacement.

We describe a more specific combinatorial interpretation of the E_k's for the case of elliptic curves in Section 4.2 of Chapter 4. We also note that the generating function methods from [Sta97, Sec. 4.7] to analyze monoids can be adapted to describe the relationship between the generating functions for the p_k's and h_k's.

2.2 Plethysm with a different alphabet

Another way for analogues of the elementary symmetric functions to appear is if we consider the numerator

$$L(C,T) = (1-\alpha_1)(1-\alpha_2)\cdots(1-\alpha_{2g}) = \sum_{i=1}^{2g}(-1)^i e_i[\alpha_1+\cdots+\alpha_{2g}]T^i.$$

We use $\tilde{E}_i$ to denote $e_i[\alpha_1+\cdots+\alpha_{2g}]$ for $0\leq i\leq 2g$, which also denote the elementary symmetric functions in the variables α_1 through α_{2g}.

Proposition 2.6. *The $\tilde{E}_k$'s satisfy initial conditions $\tilde{E}_0 = H_0 = 1, \tilde{E}_1 = H_1 - (q+1)$, and recursions*

$$\tilde{E}_k = H_k - (1+q)H_{k-1} + qH_{k-2} \text{ for } 2\leq k\leq g \text{ and} \tag{2.3}$$

$$\tilde{E}_{g+k} = q^k \tilde{E}_{g-k} \text{ for } 0\leq k\leq g. \tag{2.4}$$

Proof. We have $Z(C,T)\Big|_{T^0} = L(C,T)\Big|_{T^0}$, so $H_0 = 1 = \tilde{E}_0$. Also

$$Z(C,T)\Big|_{T^1} = L(C,T)(1+T)(1+qT)\Big|_{T^1}$$

so $H_1 = \tilde{E}_0(1+q) + \tilde{E}_1$ which proves the other initial condition. In fact in general we can rewrite $\frac{1}{(1-T)(1-qT)}$ as the infinite positive sum $(1+T+T^2+\dots)(1+qT+q^2T^2+\dots) = \sum_{0\le i\le j} q^iT^j$ which can be truncated when we try to find a single coefficient of $L(C,T)\cdot\frac{1}{(1-T)(1-qT)}$. To prove the recursion we instead use plethysm:

$$\begin{aligned}
\tilde{E}_k &= e_k[\alpha_1+\cdots+\alpha_{2g}] = e_k[(1+q)-(1+q-\alpha_1-\cdots-\alpha_{2g})] \\
&= \sum_{j=0}^{k}(-1)^{k-j}e_j[1+q]h_{k-j}[1+q-\alpha_1-\cdots-\alpha_{2g}] \\
&= e_0(1,q)H_k - e_1(1,q)H_{k-1} + e_2(1,q)H_{k-2}
\end{aligned}$$

which is the desired recursion. (We note that this recurrence has depth 2 because the denominator of $Z(C,T)$ has degree 2.)

To obtain (2.4), we use the fact that the α_i's come in pairs whose product is q, and the fact that e_{g+k} must contain at least k such pairs, by the pigeon-hole principle. After replacing each of these pairs by q and factoring them out of each term, we are left with q^k times a sum of terms which are a symmetric collection of products of distinct monomials. Thus we have obtained elementary symmetric functions in the same variables, but in a smaller degree, and so $\tilde{E}_{g+k} = q^k\tilde{E}_{g-k}$ for $0 \le k \le g$. □

The duality between the h_k's and e_k's allow us to present a dual to this proposition, or more specifically a dual to (2.3).

Proposition 2.7. *For $m \ge 0$,*

$$\begin{aligned}
H_m &= \tilde{E}_0(1+q+\cdots+q^m) - \tilde{E}_1(1+q+\cdots+q^{m-1} \\
&+ \tilde{E}_2(1+q+\cdots+q^{m-2}) - +\cdots+(-1)^{m-1}\tilde{E}_{m-1}(1+q)+(-1)^m\tilde{E}_m.
\end{aligned}$$

We can simplify such expressions by keeping in mind that $\tilde{E}_m = q^{m-g}\tilde{E}_{2g-m}$ if $g+1 \le m \le 2g$ and $\tilde{E}_m = 0$ for $m > 2g$.

Proof. We use the identity

$$h_m[1+q-(\alpha_1+\cdots+\alpha_{2g})] = \sum_{k=0}^{m}(-1)^k e_k(\alpha_1,\ldots,\alpha_{2g})h_{m-k}(1,q).$$

□

Subtracting H_{m-1} from H_m cancels most terms on the right-hand side, and so we get as an application

Corollary 2.8.

$$H_m - H_{m-1} = \tilde{E}_m + q\tilde{E}_{m-1} + \ldots q^{m-1}\tilde{E}_1 + q^m$$

for $m \geq 1$.

We also get analogous identities for writing the $\tilde{H}_k = h_k[\alpha_1+\cdots+\alpha_{2g}]$'s in terms of the E_k's and vice-versa.

Proposition 2.9. *For* $m \geq 0$,

$$\begin{aligned}
\tilde{H}_m &= E_0(1+q+\cdots+q^m) - E_1(1+q+\cdots+q^{m-1} \\
&+ E_2(1+q+\cdots+q^{m-2}) - + \ldots \\
&+ (-1)^{m-1}E_{m-1}(1+q) + (-1)^m E_m \quad \text{and} \\
\tilde{H}_m - \tilde{H}_{m-1} &= E_m + qE_{m-1} + \ldots q^{m-1}E_1 + q^m \quad \text{for} \quad m \geq 0
\end{aligned}$$

Similarly, $E_0 = 1$*,* $E_1 = 1+q-N_1$*, and*

$$E_k = \tilde{H}_k - (1+q)\tilde{H}_{k-1} + q\tilde{H}_{k-2}$$

for $k \geq 2$.

Proof. We use $h_m[\alpha_1+\cdots+\alpha_{2g}] = \sum_{k=0}^{m}(-1)^k e_k[1+q-(\alpha_1,\ldots,\alpha_{2g})]h_{m-k}(1,q)$ and $e_k[1+q-(\alpha_1-\cdots-\alpha_{2g})] = \sum_{j=0}^{k}(-1)^{k-j}e_j[1+q]h_{k-j}[1+q-(1+q-\alpha_1-\cdots-\alpha_{2g})]$. □

We summarize the relationship between coefficients of $Z(C,T)$ and symmetric functions in the following table. Hence, another application is a formula for writing

Table 2.1: Correspondence between algebraic geometric quantities and symmetric functions.

$$\begin{aligned}
N_k &\leftrightarrow p_k[1+q-\alpha_1-\cdots-\alpha_{2g}] \\
1+q^k-N_k &\leftrightarrow p_k[\alpha_1+\cdots+\alpha_{2g}] \\
E_k &\leftrightarrow e_k[1+q-\alpha_1-\cdots-\alpha_{2g}] \\
\tilde{E}_k &\leftrightarrow e_k[\alpha_1+\cdots+\alpha_{2g}] \\
H_k &\leftrightarrow h_k[1+q-\alpha_1-\cdots-\alpha_{2g}] \\
\tilde{H}_k &\leftrightarrow h_k[\alpha_1+\cdots+\alpha_{2g}].
\end{aligned}$$

N_k in terms of the H_m's via

$$N_k = p_k = \sum_{\lambda\vdash k} c_\lambda h_{\lambda_1}\cdots h_{\lambda_r} = \sum_{\lambda\vdash k} c_\lambda H_{\lambda_1}\cdots H_{\lambda_r} \tag{2.5}$$

where $c_\lambda = (-1)^{l(\lambda)-1} w(B_{\lambda,\mu})$, the weighted number of brick-tabloids [ER91] as in Eğecioğlu and Remmel 1990. (We use this identity more explicitly in Chapter 4 when we discuss elliptic curves.)

Remark 2.10. We can write the coefficients of $L(C,T)$, i.e. each of the $\tilde{E}_k$'s as a polynomial in $\{N_1, N_2, \ldots, N_k\}$ since one can write the elementary symmetric functions in terms of the power symmetric functions. Furthermore, since all the $\tilde{E}_k$'s can be expressed in terms of q and $\tilde{E}_1$ through $\tilde{E}_g$, by (2.4), we obtain $Z(C,T)$ only depends on q and N_1 through N_g, as claimed in the introduction.

2.3 Eğecioğlu and Remmel's combinatorial interpretation of formula (2.5)

The coefficients c_λ can be written down concisely as

$$c_\lambda = (-1)^{l(\lambda)-1} \frac{k}{l(\lambda)} \binom{l(\lambda)}{d_1, d_2, \ldots, d_k}$$

where $l(\lambda)$ denotes the length of λ, which is a partition of k with type $1^{d_1}2^{d_2}\cdots k^{d_k}$. We give one proof of this using Remmel's interpretation using weighted brick-tabloids, which can be derived by an equivalent combinatorial interpretation using *circular brick tabloids*. (Note that the individual terms in these weighted counts will differ, even though the weighted sums themselves are identical.) In Chapter 4 we will give an alternative proof simply using generating functions.

We present the definition of brick tabloids as in [Eğecioğlu, Remmel]. A Brick Tabloid of type $\lambda = 1^{d_1}2^{d_2}\cdots k^{d_k}$ and shape μ is a filling of the Ferrers' Diagram μ with bricks of various sizes, d_1 which are 1×1, d_2 which are 2×1, d_3 which are 3×1, etc. The weight of a brick tabloid is the product of the lengths of all bricks at the end of the rows of the Ferrers' Diagram. Let $w(B_{\lambda,\mu})$ denote the weighted-number of brick tabloids of type λ and shape μ, where each tabloid is counted with multiplicity according to its weight.

Proposition 2.11 (Eğecioğlu, Remmel)**.**

$$p_\mu = \sum_\lambda (-1)^{l(\lambda)-l(\mu)} w(B_{\lambda,\mu})$$

and in particular

$$p_k = \sum_\lambda (-1)^{l(\lambda)-1} w(B_{\lambda,(k)}).$$

Brick-Tabloids of type λ and shape (k) are simply fillings of the $k \times 1$ board with bricks as specified by λ. Thus if we divide these tabloid into classes based on the size of the last brick we obtain, by counting the number of rearrangements, that there are

$$\binom{l(\lambda)-1}{d_1, \ldots, d_i - 1, \ldots, d_k}$$

brick-tabloids of type (k) and shape $\lambda = 1^{d_1}2^{d_2}\cdots k^{d_k}$ which have a last brick of length i.

Since each of these tabloids has weight i, summing up over all possible i, we get that

$$
\begin{aligned}
w(B_{\lambda,(k)}) &= \sum_{i=0}^{k} i \cdot \binom{l(\lambda)-1}{d_1,\ldots,d_i-1,\ldots,d_k} \\
&= \left(\sum_{i=0}^{k} i d_i\right) \cdot \binom{l(\lambda)-1}{d_1,\ldots,d_i,\ldots,d_k} \\
&= k \cdot \binom{l(\lambda)-1}{d_1,d_2,\ldots,d_k} = \frac{k}{l(\lambda)} \cdot \binom{l(\lambda)}{d_1,d_2,\ldots,d_k}
\end{aligned}
$$

Note that the formula for c_λ also appears elsewhere such as [Mac95]. Thus after comparing signs, we obtain that c_λ equals exactly the desired expression. Since these formulas include terms with negative signs, we unfortunately cannot decompose the set of points on curve C directly using these summands. Nonetheless, in Section 2.5, we provide an interpretation of the c_λ's using inclusion-exclusion.

2.4 Alternative to plethysm

In many of the results involving identities of the N_k's, H_k's, and E_k's we have used the technique of plethystic substitution. In fact, lurking below many of these proofs is the standard symmetric function identity that we have been using again and again:

$$\sum_{n=0}^{\infty} h_n T^n = \prod_{k\in\mathcal{I}} \frac{1}{1-t_k T} = \exp\left(\sum_{n=1} p_n \frac{T^n}{n}\right)$$

where h_n and p_n are symmetric functions in the variables in $\mathcal{I}$.

So far we have just thought of $Z(C,T)$ as equal to this expression by letting h_n and p_n be defined plethystically in the "alphabet" $[1+q-\alpha_1-\cdots-\alpha_{2g}]$. While this is internally consistent and shows why the ordinary generating function of the H_k's is equal to an exponential generating function of the N_k's, it leaves less clear why these expressions are both equal to

$$\prod_{\mathfrak{p} \text{ a prime or Frobenius Cycle}} \frac{1}{1-T^{\deg \mathfrak{p}}}.$$

To see this more directly, we use cyclotomic polynomials. These polynomials will be used again in Chapter 5 so this introduction provides a good warm-up.

The dth **cyclotomic polynomial** in variable x is defined as the unique irreducible polynomial of degree $\phi(d)$ in the factorization of $(x^k - 1)$ for any k, a multiple of d. Here $\phi(d)$ is the number Euler Totient function which counts the number of elements in $\{1, 2, \ldots, d\}$ which are relatively prime to d. Alternatively, we can use Möbius inversion to compute

$$Cyc_d(x) = \prod_{m|d} (x^n - 1)^{\mu(d/m)}.$$

Using these, we note that

$$(1 - T^{\deg \mathfrak{p}}) = \prod_{j=1}^{\deg \mathfrak{p}} (1 - t_j T)$$

by using the cyclotomic polynomial decomposition. Thus we let each of the t_j's to be the $(\deg \mathfrak{p})$th roots of unity. In other words, let $\mathcal{I}$ be the natural numbers $\mathbb{N}$ and let the alphabet $\mathcal{A}$ of variables be such that there are I_1 copies of 1, I_2 copies of 1 and -1, I_3 copies of $1, \omega$, and ω^2 $(\omega^3 = 1)$, I_4 copies of $1, i, -1, -i$, etc. Here I_k equals the number of prime divisors of degree k.

Because of the cancelations that occur when adding roots of unity or powers of roots of unity, we get correctly that $N_1 = h_1(\mathcal{A}) = p_1(\mathcal{A}) = I_1$ for instance. Namely, $1 + \omega + \omega^2 + \cdots + \omega^{k-1} = 0$ when ω is a primitive kth root of unity. Additional examples also result in surprisingly finite expressions for these symmetric functions in an infinite alphabet.

Using this interpretation we can again derive that the combinatorial interpretation of $e_k[1 + q - \alpha_1 - \cdots - \alpha_{2g}]$ should be the alternating sum of the number of sets of Frobenius cycles (consisting of a total of k points) where sets of different cardinalities are given positive or negative signs according to a simple rule, e.g. positive if $k - (\#sets)$ is even and negative if $k - (\#sets)$ is odd. The proof hinges on the algebraic fact that

$$\prod_{i=0}^{k-1} \omega^i = \omega^{\binom{k}{2}} \equiv \begin{cases} \omega^{k/2} = -1 \text{ if k even} \\ \omega^0 = 1 \text{ if k odd.} \end{cases}$$

Similar techniques recover the other identities discussed when we first used plethysm to get identities for the H_k's and E_k's.

Table 2.2: Cyclotomic polynomials $Cyc_d(x)$ for selected d.

$$
\begin{aligned}
Cyc_1(x) &= -1 + x \\
Cyc_2(x) &= 1 + x \\
Cyc_3(x) &= 1 + x + x^2 \\
Cyc_4(x) &= 1 + x^2 \\
Cyc_5(x) &= 1 + x + x^2 + x^3 + x^4 \\
Cyc_6(x) &= 1 - x + x^2 \\
Cyc_8(x) &= 1 + x^4 \\
Cyc_{10}(x) &= 1 - x + x^2 - x^3 + x^4 \\
Cyc_{12}(x) &= 1 - x^2 + x^4 \\
Cyc_{16}(x) &= 1 + x^8 \\
Cyc_{18}(x) &= 1 - x^3 + x^6 \\
Cyc_{22}(x) &= 1 - x + x^2 - x^3 + x^4 - x^5 + x^6 - x^7 + x^8 - x^9 + x^{10} \\
Cyc_{28}(x) &= 1 - x^2 + x^4 - x^6 + x^8 - x^{10} + x^{12} \\
Cyc_{30}(x) &= 1 + x - x^3 - x^4 - x^5 + x^7 + x^8 \\
Cyc_{36}(x) &= 1 - x^6 + x^{12} \\
Cyc_{40}(x) &= 1 - x^4 + x^8 - x^{12} + x^{16} \\
Cyc_{42}(x) &= 1 + x - x^3 - x^4 + x^6 - x^8 - x^9 + x^{11} + x^{12}
\end{aligned}
$$

2.5 An inclusion-exclusion interpretation for (2.5)

We now describe the alternating formulas $N_k = \sum_{\lambda \vdash k} c_\lambda H_{\lambda_1} H_{\lambda_2} \cdots H_{\lambda_{\ell(\lambda)}}$ by counting the number of points via inclusion-exclusion on the number of divisors. As a first example, consider the expression $N_2 = 2H_2 - H_1$. We can understand this equality by double-counting all positive divisors of degree two. Such divisors come in two forms

$$\begin{aligned} D_1 &= P_1 + P_2, \quad \text{where } P_1 \text{ and } P_2 \text{ are degree one points,} \\ D_2 &= \Pi = Q_1 + Q_2, \quad \text{where } Q_1 \text{ and } Q_2 \text{ are degree two points.} \end{aligned}$$

Let $|D_1|$ denote the number of divisors of type D_1 and $|D_2|$ denote the number of type D_2. Consequently, $2H_2 = 2|D_1| + 2|D_2| = 2|D_1| + 2I_2$, where we recall I_2 equals the number of prime divisors of degree 2 and $2I_2$ also equals the number of points in $C(\mathbb{F}_{q^2})$ of degree 2. Thus we really want to count $N_2 = N_1 + 2I_2$ but $2|D_1| > N_1$, i.e. we have over-counted. To describe more fully how much we have over-counted, we note a divisor of type D_1 either looks like $2P_1$ or $P_1 + P_2$ with $P_1 \neq P_2$. There is a map between ordered pairs (P_1, P_2) of points in $C(\mathbb{F}_q)$ and degree two divisors of type D_1 by letting $(P_1, P_2) \mapsto P_1 + P_2$. This map is 1-to-1 when $P_1 = P_2$ and 2-to-1 otherwise. Thus N_1^2, which counts the number of such ordered pairs, equals $N_1 + |D_1|$, and so we subtract N_1^2, which is H_1^2, and obtain the desired identity.

In fact we can repeat this same argument for higher cases and get in particular

$$\begin{aligned} H_1 &= I_1 \\ H_2 &= I_2 + \left(\!\!\binom{I_1}{2}\!\!\right) \\ H_3 &= I_3 + I_2 I_1 + \left(\!\!\binom{I_1}{3}\!\!\right) \\ H_4 &= I_4 + I_3 I_1 + \left(\!\!\binom{I_2}{2}\!\!\right) + I_2\left(\!\!\binom{I_1}{2}\!\!\right) + \left(\!\!\binom{I_1}{4}\!\!\right), \quad \text{etc.} \end{aligned}$$

Here we are decomposing the number of positive divisors, of degree k, into types of collections of multi-sets according to the possible partitions of k. Additionally,

$$N_k = \sum_{d|k} d \cdot I_d.$$

Thus combining these relations, we get formulas for the N_k's which illustrate the above inclusion-exclusion pattern. We will give more explicit details for the elliptic case in Chapter 4.

As a final comment, we note the resemblance between the above formulas for H_k and N_k's in terms of the I_k's and a class of symmetric functions introduced by Reutenauer, which are related to Witt vectors and the free Lie algebra. In [Reu95], he discusses a family of symmetric functions defined by

$$\prod_{n\geq 1} \frac{1}{1 - q_n t^n} = \sum_{n\geq 0} h_n t^n$$

which also implies that $p_i = \sum_{i=nk} n q_n^k$. In such a formula, the power symmetric functions are called the ghost components of these q_n's.

3 Elliptic curves

The theory of elliptic curves is quite rich, arising in both the areas of complex analysis and number theory. Such curves can be given a group structure using the tangent-chord method or the divisor class group of algebraic geometry. This property makes them not only geometric but also algebraic objects and allows them to be used for cryptographic purposes. Because of their appearance in such a varied number of subjects, we now will devote the rest of this thesis to this special case. In this chapter we present the necessary background material and provide details of some of the amazing facts that are true for the elliptic case. In particular, we will discuss (1) the group structure on elliptic curves, (2) the theory of division polynomials, and (3) how these can be used to prove a characteristic equation for the Frobenius map. We follow sources such as [Gan], [Sil92], and [Was03] for the material of this chapter.

3.1 Weierstraß form and group law

We recall from Chapter 1 that the Riemann-Roch Theorem tells us that a genus g curve has $L(D)$ of dimension given by

$$dimL(D) - dimL(K-D) = deg\ D + 1 - g$$

where K is the canonical divisor, which has degree $2g-2$. In the case of genus one, this gives an explicit description of such curves. Firstly, we have that K is a divisor of degree 0 in the $g=1$ case, and that for a divisor D_0 of degree zero, that $L(D_0)$ has dimension equal to the dimension of $L(K-D_0)$.

Proposition 3.1. *For genus one curves, the canonical class contains the zero divisor. Thus we set $K = 0$, up to class representative.*

Proof. Recall by Lemma 1.15 that $dimL(D) \leq \deg D+1$ and so in particular, if D_0 has degree zero, $L(D_0)$ has dimension 0 or 1. Also during the course of the proof of this lemma we noted that $L(0)$ has dimension one since the constant functions have no zeros or poles. Now assume there exists another D' of degree zero such that $L(D')$ also has nonzero dimension. Then there exists positive divisor D'' and rational function f such that $D' = D'' + (f)$. However, since D' is of degree zero, so is D''. However, we conclude $D'' = 0$ since the only positive divisor of degree zero is the zero divisor. Thus there is a unique class, the ones corresponding to principal divisors, of degree zero divisors D with dim $L(D) = 1$. Finally, since $L(0)$ has the same dimension as $L(K - 0)$ by Riemann-Roch, K must be in this unique class, i.e. the same divisor class as 0. □

Any degree zero divisor D_0 besides those equivalent to 0 will have dim $L(D_0) = 0$, and dim $L(0) = 1$. Since the constant functions have divisor 0, we obtain for degree zero D_0

$$L(D_0) = \begin{cases} \{0\} \text{ if } D_0 \not\equiv 0 \\ k \quad \text{if } D_0 \equiv 0. \end{cases}$$

For divisors D of degree greater than 0, we have that deg $(K - D) < 0$ thus dim $L(D) = deg\ D$. Using this dimension count, we can verify the following bases for the below vector spaces:

$$\begin{aligned} L(P_\infty) &= \{1\} \\ L(2P_\infty) &= \{1, x\} \\ L(3P_\infty) &= \{1, x, y\} \\ L(4P_\infty) &= \{1, x, y, x^2\} \\ L(5P_\infty) &= \{1, x, y, x^2, xy\} \\ L(6P_\infty) &= \{1, x, y, x^2, xy, x^3 = y^2\} \end{aligned}$$

The upshot is that the quotient space $L(6P_\infty) \big/ L(5P_\infty)$ has dimension one but spanning set $\{x^3, y^2\}$. Thus with respect to the genus one curve, we have the relation

$$y^2 - x^3 = A_1xy + A_2x^2 + A_3y + A_4x + A_6.$$

Theorem 3.2. *Any genus* 1 *curve is in fact a hyperelliptic curve. We call such curves elliptic curves. If the characteristic is not* 2 *or* 3 *the equation for the curve can be written as*

$$y^2 = x^3 + Ax + B$$

up to isomorphism. This is called the Weierstraß form of the curve. We call genus 1 *curves* **elliptic** *curves.*

Proof. We have done the heart of the proof above, we need only note that in characteristic $\neq 2, 3$ we can algebraically manipulate, using techniques such as completing the square, and choose $x' = \alpha_1 x + \beta_1$ and $y' = \alpha_2 y + \beta_2 x + \gamma_2$ such that

$$y'^2 = x'^3 + Ax' + B.$$

$\square$

Remark 3.3. Notice that the fact that $L(P_\infty)$ is spanned by $\{1\}$ also implies that there is no nonconstant function which has a pole at exactly one point on an elliptic curve. Thus, there are N_1 degree one positive divisors and they are all inequivalent.

Another amazing fact about the special case of elliptic curves is the existence of a group law. Thereby, the curve is not only a geometric object, but also an algebraic object.

Definition 3.4. If C, over an arbitrary field k, is defined by equation

$$y^2 = x^3 + Ax + B$$

and $P_1 = (x_1, y_1)$, $P_2 = (x_2, y_2)$, then

$$P_1 \oplus P_2 = P_3 = (x_3, y_3)$$

where

1) If $x_1 \neq x_2$ then

$$x_3 = m^2 - x_1 - x_2 \quad \text{and} \quad y_3 = m(x_1 - x_3) - y_1 \quad \text{with} \quad m = \frac{y_2 - y_1}{x_2 - x_1}.$$

2) If $x_1 = x_2$ but ($y_1 \neq y_2$, or $y_1 = 0 = y_2$) then $P_3 = P_\infty$.

3) If $P_1 = P_2$ and $y_1 \neq 0$, then

$$x_3 = m^2 - 2x_1 \quad \text{and} \quad y_3 = m(x_1 - x_3) - y_1 \quad \text{with} \quad m = \frac{3x_1^2 + A}{2y_1}.$$

4) The point at infinity, P_∞, acts as the identity element in this addition.

Lemma 3.5. *Definition 3.4 yields an associative abelian group on the set of points on C, including P_∞.*

We note that since the group law is defined explicitly, the associativity can be directly verified, though one needs to be careful to include all of the cases. However, since we have previously proven the Riemann-Roch Theorem, we instead give a shorter proof using this result. Before proceeding, we need the following lemma.

As we saw above, there exists a divisor class of degree one for all points on the curve. In fact we have the stronger result

Lemma 3.6. *Any degree m divisor is equivalent to a divisor of the form*

$$D = P + mP_\infty$$

where P is a point on the curve, possibly P_∞.

Proof. By Rieman-Roch the divisor of a line, which is a rational function, is a degree zero divisor. Bezout's Theorem [Har77] tells us that the number of points on the intersection of a degree three rational function, $y^2 = x^3 + Ax + B$, and a degree one rational function, $ay + bx + c = 0$ is $3 \cdot 1 = 3$ counting multiplicities. Thus the divisor of a line on a curve is equal to

$$P + Q + R - 3P_\infty$$

with P, Q, R, P_∞ not necessarily distinct. Thus given divisor

$$D = D_+ - D_-$$

where both D_+ and D_- are both positive divisors we can use various lines to reduce D_+ and D_- separately.

We have that for every P, Q (including $Q = P$) on the curve, their sum is equivalent to $-R + 3P_\infty$. Secondly, we have that the line $x = a$ contains both the points (a, b) and $(a, -b)$ (and P_∞ as the third point). This includes the case where line $x = a$ is tangent, multiplicity two, to the point $(a, 0)$. Thus the divisor $P_{(a,b)} + P_{(a,-b)} - 2P_\infty \equiv 0$ and we have that $-R + 3P_\infty \equiv \overline{R} + P_\infty$ where $\overline{R}$ is the conjugate point $(R_x, -R_y)$. By repeated application, we are left with a single point plus a multiple of the point at infinity. $\square$

Proof of Lemma 3.5. Thus we can define the group law, in fact it is inherited from the divisor class group, as

$$P \oplus Q = R \iff (P - P_\infty) + (Q - P_\infty) \equiv (R - P_\infty).$$

Associativity and commutativity thereby come for free. We only need to check this geometric description using lines is equivalent to the above algebraic description. By the fact that the three points $P_1 = (x_1, y_1)$, $P_2 = (x_2, y_2)$, and $-P_3 = (x_3, -y_3)$ lie on the same line, we have by similar triangles that

$$\frac{(-y_3) - y_1}{x_3 - x_1} = \frac{y_2 - y_1}{x_2 - x_1}.$$

Rearranging this equality gives us the formula for y_3. To get the expression for x_3 takes a little more work.

We first notice that for all (x, y) on the elliptic curve, $y = m(x - x_1) + y_1$ where m is the slope $\frac{y_2 - y_1}{x_2 - x_1}$. Since we have the equality $y^2 = x^2 + Ax + B$, we also obtain that

$$0 = x^3 - m^2x^2 + \ldots.$$

The three roots of this equation are exactly the three x-coordinates for the points in the intersection of line L through P_1 and P_2 and elliptic curve C. Consequently, since the coefficient of the quadratic term is the negative of the sum of the roots,

$$m^2 = x_1 + x_2 + x_3$$

and after rearrangement, we have our expression for x_3. The case of doubling a point using tangent lines is analogous. $\square$

3.2 Rational function representations of morphisms

We will define an **endomorphism** $\alpha : E \to E$ of an elliptic curve as a homomorphism, with respect to the group law, that can be represented as a pair of rational functions g_α and h_α. In other words, α fixes P_∞ and

$$\alpha(x, y) = \Big(g_\alpha(x, y),\ h_\alpha(x, y) \Big) \text{ and } (g_{\alpha+\beta}, h_{\alpha+\beta}) = (g_\alpha, h_\alpha) \oplus (g_\beta, h_\beta)$$

since

$$(\alpha + \beta)(P) = \alpha(P) \oplus \beta(P).$$

We will closely follow Section 2.8 of [Was03] in this subsection as we discuss further properties of endomorphisms.

Since α is a group homomorphism, it maps the identity P_∞ to itself. Borrowing from geometric language, an endomorphism is also sometimes referred to as an isogeny since it has such a fixed point. We will refer to α as the zero map if it sends every point of E to P_∞ and nontrivial otherwise.

We first note the following algebraic geometric fact concerning endomorphisms.

Theorem 3.7. *Let E be defined over $\overline{\mathbb{F}_q}$ (in fact any algebraically closed field). Then an endomorphism α is either surjective or the zero map.*

Proof. See [Gan], [Har77] for a proof, or [Was03, Thm 2.21] for a more elementary approach. $\square$

Lemma 3.8. *For elliptic curves, and more generally hyperelliptic curves, we can rationalize rational functions in $k(C)$ so that they are of the form $\frac{p_1(x)+p_2(x)y}{p_3(x)}$ where the p_i's are polynomials.*

Proof. If g is a rational function in $k(C)$ of the form $\frac{P(x,y)}{Q(x,y)}$, we have the relation $y^2 = f_0(x)$, e.g. $f_0(x) = x^3 + Ax + B$ in the elliptic case. Thus we can rewrite

$$\frac{P(x,y)}{Q(x,y)} = \frac{A(x) + yB(x)}{C(x) + yD(x)} = \frac{(A(x) + yB(x))(C(x) - yD(x))}{C(x)^2 - y^2 D(x)}$$

and the denominator can again be simplified so it is univariate in x. $\square$

In fact in the elliptic case, we can describe these rational functions even more precisely.

Lemma 3.9. *If* $\alpha(x,y) = \Big(g_\alpha(x,y),\ h_\alpha(x,y)\Big)$ *is an endomorphism of an elliptic curve, then*

$$g_\alpha \text{ is univariate in terms of } x \text{ and } h_\alpha = y\ \overline{h_\alpha}(x).$$

where $\overline{h_\alpha}(x)$ *is a univariate rational function.*

Proof. We obtain these last expressions by using the group law and the fact that α is a homomorphism to obtain

$$\alpha(x, -y) = \alpha(\ominus(x,y)) = \ominus\alpha(x,y).$$

Consequently, the x-coordinate of $\alpha(x,y)$, i.e $g_\alpha(x,y)$ satisfies $g_\alpha(x,y) = g_\alpha(x,-y)$ and analogously, $h_\alpha(x,y) = -h_\alpha(x,-y)$. Thus g_α has no y-coordinate and h_α has no x-coordinate. $\square$

Notational convention: if we wish to write these rational functions as polynomials we will write

$$g_\alpha \text{ as } n_\alpha(x)/d_\alpha(x) \text{ and } h_\alpha \text{ as } y\ \tilde{n}_\alpha(x)/\tilde{d}_\alpha(x)$$

such that both pairs n_α, d_α and $\tilde{n}_\alpha$, $\tilde{d}_\alpha$ have no common factors.

Note that since these are rational functions, as opposed to polynomials, there will exist choices of $x \in \overline{\mathbb{F}_q}$ such that the denominators are zero. A priori it might appear that it would be possible for one of $d_\alpha(x_0)$, $\tilde{d}_\alpha(x_0)$ to be zero and not the other but we will shortly find that we can consistently define $\alpha\Big((x_0, y_0)\Big) = P_\infty$ in this case by the following lemma.

Lemma 3.10. *For any $x_0 \in \overline{\mathbb{F}_q}$, either both $d_\alpha(x_0)$ and $\tilde{d}_\alpha(x_0) \neq 0$ or both $d_\alpha(x_0)$ and $\tilde{d}_\alpha(x_0) = 0$. Thus, in the former case we have $\alpha\Big((x_0, y_0)\Big) = (a, b) \in \overline{\mathbb{F}_q}^2 \cap E$, and the latter we have $\alpha\Big((x_0, y_0)\Big) = P_\infty$.*

Proof. First we note that the coordinates of $\alpha\Big((x, y)\Big)$, i.e. $(g_\alpha,\ h_\alpha) = (g_\alpha,\ y\ \overline{h}_\alpha)$ satisfy the defining equation

$$h_\alpha^2 = g_\alpha^3 + A\ g_\alpha + B.$$

Thus

$$h_\alpha^2 = y^2 \overline{h}_\alpha^{\ 2} = \frac{(x^3 + Ax + B)\ \tilde{n}_\alpha(x)^2}{\tilde{d}_\alpha(x)^2} = \frac{\tilde{\tilde{n}}_\alpha(x)}{d_\alpha(x)^3}$$

for polynomial $\tilde{\tilde{n}}_\alpha(x)$ with no common factor with $d_\alpha(x)$. More precisely, $\tilde{\tilde{n}}_\alpha(x) = n_\alpha^3(x) + A\ n_\alpha(x)\ d_\alpha(x)^2 + B\ d_\alpha(x)^3$ and $n_\alpha(x)$ has no factors in common with $d_\alpha(x)$.

If $d_\alpha(x_0) = 0$ then the denominator of the square of h_α is also zero hence $\tilde{d}_\alpha(x_0) = 0$. If, on the other hand, $\tilde{d}_\alpha(x_0) = 0$ then we might have that x_0 is a root of both $x^3 + A\ x + B$ and $\tilde{d}_\alpha(x)^2$, however the first expression has no multiple roots since $E(\mathbb{F}_q)$ was assumed to be a nonsingular curve, and the second has roots with multiplicities at least two. Thus the denominator will still be zero in this case, hence $d_\alpha(x_0) = 0$ as well. By the contrapositive, we have that one of these is nonzero if and only if the other is nonzero too. $\square$

Remark 3.11. We will see this relationship between g_α and h_α again when we study division polynomials in Section 3.3, namely, that there exists a polynomial $\Psi_\alpha(x)$ such that $\Psi_\alpha(x)^2 = d_\alpha(x)$ and $\Psi_\alpha(x)^3 = \tilde{d}_\alpha(x)$.

With this last lemma in mind, we note that the first coordinate alone determines whether or not $\alpha(P) = P_\infty$, and in fact only the denominator matters, which motivates the following definition. We define the **degree** of nontrivial endomorphism α to be

$$\deg(\alpha) = \mathrm{Max}\{\deg\ n_\alpha(x),\ \ \deg\ d_\alpha(x)\}.$$

The degree of the zero map is set to be 0. This quantity degree is important for several different reasons.

1. The $\deg(\alpha)$ serves as an upper bound for the size of the Ker α with equality in many cases. We will shortly make this rigorous.

2. A map α between curves E_1 and E_2 induces an contravariant injection α^* between funciton fields $k(E_2)$ and $k(E_1)$. In this context, the degree of map α is equal to the degree of the field extension $k(E_1)/k(\alpha(E_1))$.

3. We will see in Section 3.4 that the n-torsion subgroup (when $\gcd(n,q)=1$) of an elliptic curve is isomorphic to a lattice and thus endomorphisms can also be represented as $2-$by-2 matrices. In this context, the $\deg(\alpha)$ is equal to the determinant modulo n.

4. Using this $2-$by-2 matrix interpretation, or otherwise, we obtain that degree gives rise to a quadratic form on the space of endomorphisms; more precisely

   $$\deg(r\alpha + s\beta) = r^2\deg(\alpha) + s^2\deg(\beta) + rs\Big(\deg(\alpha+\beta) - \deg(\alpha) - \deg(\beta)\Big).$$

We now proceed to make precise the relationship between degree and the size of Ker α. We begin by calling a nontrivial endomorphism α **separable** if the derivative of rational function $g_\alpha(x)$ is not identically zero. Recall that g_α is the rational function corresponding to the x-coordinate of $\alpha((x,y))$.

Remark 3.12. One can also formulate the notion of separability using algebraic language, namely that α is separable if and only if it induces a separable extension on function fields. In other words,

$$\alpha : E_1 \to E_2$$

is separable if and only if

$$\alpha^* : k(E_2) : k(E_1)$$

induces

$$k(E_1)/\alpha^*(k(E_2)) \quad \text{a separable field extension.}$$

While this definition has its advantages, to be able to utilize it properly, we would have to discuss notions such as ramification degree that would take us away from our goal. One can find such an approach in [Sil92].

We see from the next Lemma, that one need not check separability at the rational function level, but that it suffices to check it for the corresponding polynomials.

Lemma 3.13. *Using our notation, $g_\alpha(x) = n_\alpha(x)/d_\alpha(x)$ for univariate polynomials n_α, d_α with no common factors, we have that α is separable if and only if at least one of $\frac{d}{dx}n_\alpha(x) = n'_\alpha(x)$ or $\frac{d}{dx}d_\alpha(x) = d'_\alpha(x)$ is not identically zero.*

Proof. $\frac{d}{dx}\left(\frac{n_\alpha(x)}{d_\alpha(x)}\right) = 0$ if and only if the numerator, using the quotient rule for derivation,

$$d_\alpha(x)n'_\alpha(x) - n_\alpha(x)d'_\alpha(x) = 0.$$

Since $d_\alpha(x)$ is assumed to be $\neq 0$, if we further assume that $d'_\alpha(x) \neq 0$, we get that

$$\frac{n_\alpha(x)}{d_\alpha(x)} = \frac{n'_\alpha(x)}{d'_\alpha(x)}$$

where both $n'_\alpha(x)$ and $d'_\alpha(x)$ have degrees smaller than $n_\alpha(x)$ and $d_\alpha(x)$, respectively. Since $n_\alpha(x)/d_\alpha(x)$ had been assumed to be in lowest terms we get a contradiction. Thus we must have $d'_\alpha(x)$ is identically zero, and hence $n'_\alpha(x) = 0$ also from the above equality. □

Now that we have reduced the notion of separability to considering polynomials, we can use the following observation to determine whether or not α is separable.

Lemma 3.14. *If the characteristic of our field is zero, then any nonconstant polynomial will have a nonzero derivative. If the characteristic is p, then any polynomial with zero derivative is of the form $P(x^p)$, or equivalently $P(x)^p$, for polynomial P.*

Proof. The derivative of a polynomial $a_nx^n + \cdots + a_1x + a_0$ is $na_nx^{n-1} + \cdots + 2a_2x + a_1$ which is the zero polynomial if and only if all the coefficients $ka_k \equiv 0 \mod p$. Thus the only terms with nonzero coefficients must be those with exponents a multiple of p. Since $(y^p + z^p) = (y+z)^p$ in characteristic p, we have the result. □

Proposition 3.15. *If $\alpha \neq 0$ is a separable endomorphism of elliptic curve E over $\overline{\mathbb{F}_q}$, or another algebraically closed field, then*

$$\deg(\alpha) = \#\ \mathrm{Ker}(\alpha).$$

If $\alpha \neq 0$ is not separable, then

$$\deg(\alpha) > \#\ \mathrm{Ker}(\alpha).$$

Proof. See [Was03, Ch. 2]. $\square$

3.3 Division polynomials and the multiplication by n map

This section is based on notes from [Cas91], [Lan78], [Was03, pg.77], and [Was03, Sec. 9.5]. To better understand the group structure of elliptic curves, we define a sequence of polynomials in $\mathbb{Z}[x, y, A, B]$ via the following initial conditions and recurrence equations:

$$\begin{aligned}
\psi_0 &= 0\\
\psi_1 &= 1\\
\psi_2 &= 2y\\
\psi_3 &= 3x^4 + 6Ax^2 + 12Bx - A^2\\
\psi_4 &= 4y(x^6 + 5Ax^4 + 20Bx^3 - 5A^2x^2 - 4ABx - 8B^2 - A^3)\\
&\cdots\\
\psi_{2m+1} &= \psi_{m+2}\psi_m^3 - \psi_{m-1}\psi_{m+1}^3 \text{ for } m \geq 2\\
\psi_{2m} &= (\frac{\psi_m}{2y}) \cdot (\psi_{m+2}\psi_{m-1}^2 - \psi_{m-2}\psi_{m+1}^2) \text{ for } m \geq 2
\end{aligned}$$

The polynomial ψ_n is known as the nth **division polynomial**. These polynomials turn out to have the remarkable property that all of the finite n-torsion points (x_0, y_0), i.e. elements of $\overline{E}[n] \setminus \{P_\infty\}$, satisfy $\psi_n^2(x_0, y_0) = 0$. Here $\overline{E}$ is shorthand for $E(\overline{\mathbb{F}_q})$ and $\overline{E}[n]$ signifies those points in $\overline{E}$ in the kernel of the multiplication by n map sending $P \mapsto P \oplus P \oplus \cdots \oplus P$. In fact we can describe this property more precisely.

Proposition 3.16. *For the ψ_n as defined above, we have the alternative definition that for $n \in \mathbb{Z}$, then $\psi_n(x, y)$ is defined as the unique rational function such that*

$$\psi_n(x,y)^2 = n^2 \cdot \prod_{P_i=(a_i,b_i)\in\overline{E}[n]\setminus\{P_\infty\}} (x - a_i)$$

and $\psi_n(x,y)$ has leading term $+n$.

Additionally, we can define the multiple of a point, $r\cdot(x,y)$, as a pair of rational functions in terms of x and y using the ψ_n's. In particular, we have the following:

Proposition 3.17. *Let $P = (x,y)$ be a point on the elliptic curve $y^2 = x^3+Ax+B$ over some field of characteristic $\neq 2$. Then for any positive integer n, $nP = P \oplus P \oplus P \oplus \cdots \oplus P$ is given by*

$$nP = \left(\frac{\phi_n(x)}{\psi_n^2(x)},\ \frac{\omega_n(x,y)}{\psi_n^3(x,y)}\right) = \left(x - \frac{\psi_{n-1}\psi_{n+1}}{\psi_n^2(x)},\ \frac{\psi_{2n}(x,y)}{2\psi_n^4(x)}\right).$$

$$-nP = \left(\frac{\phi_n(x)}{\psi_n^2(x)},\ \ -\frac{\omega_n(x,y)}{\psi_n^3(x,y)}\right) = \left(x - \frac{\psi_{n-1}\psi_{n+1}}{\psi_n^2(x)},\ \ -\frac{\psi_{2n}(x,y)}{2\psi_n^4(x)}\right)$$

where the polynomials ϕ_n and ω_n are defined as

$$\begin{aligned}\phi_m &= x\psi_m^2 - \psi_{m+1}\psi_{m-1}\\ \omega_m &= \frac{\psi_{m+2}\psi_{m-1}^2 - \psi_{m-2}\psi_{m+1}^2}{4y}.\end{aligned}$$

Proof. For the proofs of Propositions 3.16 and 3.17, see [Lan78] or [Was03, Ch. 9]. ☐

Note that by Proposition 3.16 or via the equivalence relation $y^2 \equiv x^3 + Ax + B$ and the recurrence relations for ψ_{2m} and ψ_{2m+1}, we can inductively prove that

$$\psi_n^2,\ \frac{\psi_{2n}}{y},\ \psi_{2n+1},\ \text{and } \phi_n \text{ are all functions in terms of } x.$$

As a corollary, the x-coordinate of nP is a rational function strictly in terms of x, and the y-coordinate has the form $y \cdot \Theta(x)$.

We can summarize these results as follows: ψ^2 is a function in x alone and has degree $n^2 - 1$, which equals the number of finite n-torsion points. The degree of ψ^2 is easily verified via the above recurrence relations. Furthermore, if n is odd and $(x_0, y_0) \in \overline{E} \setminus \{P_\infty\}$, then

$$\psi_n(x_0) = 0 \text{ if and only if } (x_0, y_0) \in \overline{E}[n].$$

If n is even, E is defined by equation $y^2 = (x-\alpha_1)(x-\alpha_2)(x-\alpha_3)$ over $\overline{\mathbb{F}_q}$, and $(x_0, y_0) \in \overline{E} \setminus \{P_\infty,\ (\alpha_1, 0),\ (\alpha_2, 0),\ (\alpha_3, 0)\}$, then

$$\frac{\psi_n}{y}(x_0) = 0 \text{ if and only if } (x_0, y_0) \in \overline{E}[n].$$

Corollary 3.18. *The degree of the endomorphism of multiplication by n has degree n^2.*

Proof. This is simply because the maximum of the degrees of $\phi_n(x)$ and $\psi_n^2(x)$, which in fact only depend on x, is n^2. □

Corollary 3.19. *If $gcd(n, p) = 1$ then $\alpha = [n]$ is a separable endomorphism, thus the $\#\mathrm{Ker}(\alpha) = \deg(\alpha) = n^2$.*

Proof. See [Sil92] or [Was03] for example, for the proof that $[n]$ is separable when $\gcd(n, p) = 1$.

In particular, when this morphism is separable, it has no multiple roots. Thus since the degree of the denominator is $n^2 - 1$, we have $n^2 - 1$ values of $\alpha \in \overline{\mathbb{F}_p}$ we can plug in to obtain a zero denominator, i.e. an x-coordinate of ∞.

Hence, if we let $P = P_\infty$ or (α, β) where α a zero of $\phi_n^2(x)$, we obtain $nP = P_\infty$. There are n^2 such possibilities, thus n^2 elements in the kernel of this separable morphism, and the multiplication by n map has degree n^2. □

Note in the case $gcd(n, p) > 1$ the multiplication map is not separable. The degree is still n^2, but the size of the kernel is smaller since there will be multiple roots.

Corollary 3.20. *If $gcd(n, p) = 1$ then the group $E[n] \cong \mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$.*

Proof. Based on [Gan]. We have just proven that the group $E[n]$ satisfies $\#E[n] = n^2$ in this case. By the Fundamental Theorem of Finite Abelian Groups, we have that

$$E[n] \cong (\mathbb{Z}/n_1\mathbb{Z})^{d_1} \times \cdots \times (\mathbb{Z}/n_k\mathbb{Z})^{d_k}$$

such that $n_1 | n_2 | \ldots | n_k$ and $n^2 = n_1^{d_1} \cdots n_k^{d_k}$.

Assume that $n_1 < n$. Then $E[n]$ contains a cyclic subgroup of order n_1 hence elements of order n_1. Thus $E[n]$ would have $E[n_1]$, the $[n_1]$-torsion points as a subgroup. $E[n_1]$ inherits its structure from $E[n]$ and since n_1 was assumed to be the smallest we have that $E[n_1] \cong (\mathbb{Z}/n_1\mathbb{Z})^{d_1}$ which implies that $d_1 = 2$. Furthermore, every generator of a cyclic subgroup of $E[n]$ would also be a generator for a cyclic subgroup of $E[n_1]$ since n_1 divides all thier orders. Thus the cyclic decomposition of $E[n_1]$ tells us that there at most two cyclic subgroup of $E[n]$, and we have that $E[n] \cong (\mathbb{Z}/n_1\mathbb{Z}) \times (\mathbb{Z}/n_1'\mathbb{Z})$, and since $n_1 n_1' = n^2$, we have $n_1 = n_1' = n$. □

Corollary 3.21. *The abelian group $E(\mathbb{F}_{q^k})$, for any elliptic curve E over finite field $\mathbb{F}_{q^k}$, can be decomposed as a product of at most two cyclic groups, i.e. of form*

$$E(\mathbb{F}_{q^k}) \cong \mathbb{Z}_{N_1} \times \mathbb{Z}_{N_2}$$

where $N_1 | N_2$.

Proof. Since $|E(\mathbb{F}_{q^k})|$ is finite, there exists an N such that $E(\mathbb{F}_{q^k}) \subset E(\overline{\mathbb{F}_q})[N]$. Thus $E(\mathbb{F}_{q^k})$ is a subgroup of $E(\overline{\mathbb{F}_q})[N] \cong \mathbb{Z}_N \times \mathbb{Z}_N$. Assume that $E(\overline{\mathbb{F}_q})[N]$ is generated by α and β, both of degree N. Then any subgroup of $E(\mathbb{F}_{q^k})$ will have at most two generators. Lastly, if $N_1 \not| N_2$ then $N_1 = ac$, $N_2 = bc$ with $gcd(a, b) = 1$ such that $gcd(a, c) = 1$ without loss of generality, and $a \neq 1$. Thus letting $N_1' = c$, $N_2' = abc$, we obtain $\mathbb{Z}_{N_1} \times \mathbb{Z}_{N_2} \cong \mathbb{Z}_{N_1'} \times \mathbb{Z}_{N_2'}$ with $N_1' | N_2'$. □

Remark 3.22. Division polynomials $\psi_n(x, y)$ are also an example of an *elliptic divisibility sequence* (EDS) [War48], which means

1) $\psi_n | \psi_m$ iff $n | m$.

2) The recurrence

$$\psi_{n+m}\psi_{n-m} = \psi_m^2 \psi_{n-1}\psi_{n+1} - \psi_{m-1}\psi_{m+1}\psi_n^2 \tag{3.1}$$

is satisfied. (Note that we proved recurrence (3.1) in the course of proving Proposition 3.16.)

3) Alternatively, we could let $m = 2$ and shift indices to see that the ψ_n's (or for that matter, any EDS) satisfy

$$\psi_n \psi_{n-4} = (\psi_2^2)\psi_{n-1}\psi_{n-3} + (-\psi_1\psi_3)\psi_{n-2}^2$$

This is a special case of the *Somos-4 sequence* [Pro] which in general looks like:

$$s_n s_{n-4} = \alpha s_{n-1} s_{n-3} + \beta s_{n-2}^2.$$

4) A proper EDS $\{s_n\}$ satisfies $s_0 = 0, s_1 = 1, s_2 | s_4$. Note that the division polynomials $\psi_n(x, y)$ satisfy this property.

There has been recent literature regarding this pattern, in particular for specific curves, the x-coordinates of the rational points form a Somos sequence. We invite the reader to read [VDPS06], [Pro], or [Swa] for more details. This sequence is a manifestation of the interplay between elliptic curves and combinatorics. We will discuss other connections of a different flavor starting in the next chapter.

3.4 Further properties of the Frobenius map

We now describe the remarkable properties of the Frobenius map in the special case of elliptic curves. One important property of the Frobenius map is its compatibility with the group law on elliptic curves over $\mathbb{F}_q$. In particular, we have the following:

Proposition 3.23. *If we let π signify the Frobenius map, then we have the relation*

$$\pi(P \oplus Q) = \pi(P) \oplus \pi(Q) \tag{3.2}$$

for points $P, Q \in C(\overline{\mathbb{F}_q})$.

Proof. This follows by explicit verification using the algebraic formulas for the group law, taking care to include the various cases. $\square$

Because of the reason that equation (3.2) resembles the distributive law, we sometimes refer to "acting by" the Frobenius map as multiplication by the Frobenius map. The Frobenius map allows to rephrase our main goal, namely calculating the order of $E(\mathbb{F}_{q^k})$, as the calculation of $\#\mathrm{Ker}(1 - \pi^k)$. We have that for $a \in \overline{\mathbb{F}_q}$, $\pi^k(a) = a$ if and only if $a \in \overline{\mathbb{F}_{q^k}}$.

Since $\pi\Big((x, y)\Big) = (x^q, y^q)$, we easily see that $\deg(\pi) = q$. However, $\frac{d}{dx}x^q = qx^{q-1} \equiv 0$ hence the Frobenius map is inseparable. Nonetheless we obtain

Lemma 3.24. *The endomorphism*

$$r\pi + s$$

$(r, s \in \mathbb{Z})$ is separable if and only if $gcd(s, q) = 1$. In particular, $1 - \pi$ is separable and

$$N_k = \#\mathrm{Ker}(1 - \pi^k) = \deg(1 - \pi^k).$$

Proof. See [Was03, Ch. 2]. □

Recall from Corollary 3.20 that $E(\overline{\mathbb{F}_q})[n] \cong \mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$ if $gcd(n, q) = 1$. Since π is a morphism which acts on $E(\overline{\mathbb{F}_q})[n]$ (since $\pi \circ [n] = [n] \circ \pi$ implies that $nP = P_\infty \Leftrightarrow n \circ \pi(P) = P_\infty$), we have that π's action on $E(\overline{\mathbb{F}_q})[n]$ can be represented by a 2×2 matrix with coefficients in $\mathbb{Z}/n\mathbb{Z}$.

As a consequence π satisfies a quadratic characteristic equation

$$\pi^2 - t_n\pi + d_n = 0$$

on $E(\overline{\mathbb{F}_q})[n]$, thus π satisfies

$$\pi^2 - t_n\pi + d_n \equiv 0 \pmod{n}.$$

Since we get such a quadratic characteristic equation for an infinite set of n satisfying $gcd(n, q) = 1$, we find a unique $t, d \in \mathbb{Z}$ such that

$$\pi^2 - t\pi + d = 0$$

on all points of $E(\overline{\mathbb{F}_q})$ with order relatively prime to q. There are an infinite number of such points.

Proposition 3.25. *For all points $P \in E(\overline{\mathbb{F}_q})$, we have the identity $\pi^2 - t\pi + d = 0$ where $t = 1 + q - N_1$ and $d = q$.*

Proof. See [Was03] for the details on why t and d are specifically $1 + q - N_1$ and q respectively. Once this is verified for all n such that $gcd(n, p) = 1$, we note that the expression $\pi^2 - t\pi + d$ is also a morphism which can be represented by a pair of rational functions (using the definition of the Frobenius map, division polynomials, composition, and the group law). Thus there can only be a finite number of elements in the kernel, unless it is the zero map. Thus we obtain

$$\pi^2 - t\pi + d = 0$$

on all of $E(\overline{\mathbb{F}_q})$. □

In fact, by considering the inverse limit of the sequence $\{E(\overline{\mathbb{F}_q})[\ell^k]\}$, where each term is isomorphic to $\mathbb{Z}/\ell^k\mathbb{Z} \times \mathbb{Z}/\ell^k\mathbb{Z}$, we recover a construction of the Tate Module, a two dimensional space on which the Frobenius endomorphism acts. See [Sil92] for more on the Tate Module. One of the surprising and important results of étale cohomology is that the choice of prime ℓ does not matter for this calculation, as long as $\ell \neq p$. In this respect, the value t is the trace of the Frobenius map, and d is the determinant of the Frobenius map under this 2-dimensional action.

4 Combinatorial aspects of elliptic curves

Recall that when E is an elliptic curve, $Z(E,T)$ can be expressed as

$$\frac{1-(\alpha_1+\alpha_2)T+\alpha_1\alpha_2T^2}{(1-T)(1-qT)}$$

and in particular we have

$$N_k = 1+q^k-\alpha_1^k-\alpha_2^k = p_k[1+q-\alpha_1-\alpha_2].$$

Plugging in $k=1$ the relation $\alpha_1+\alpha_2 = 1+q-N_1$ and we note that $\alpha_1\alpha_2 = q$ is a special case of the zeta function's functional equation we saw in Chapter 1.

Hence we can rewrite the zeta function $Z(E,T)$ totally in terms of q and N_1 and as a consequence, all the N_k's are actually dependent on these two quantities. This data gives rise to the following observation of Adriano Garsia.

Table 4.1: N_k's as polynomials for small k.

$$\begin{aligned}
N_2 &= (2+2q)N_1 - N_1^2\\
N_3 &= (3+3q+3q^2)N_1 - (3+3q)N_1^2 + N_1^3\\
N_4 &= (4+4q+4q^2+4q^3)N_1 - (6+8q+6q^2)N_1^2 + (4+4q)N_1^3 - N_1^4\\
N_5 &= (5+5q+5q^2+5q^3+5q^4)N_1 - (10+15q+15q^2+10q^3)N_1^2\\
&+ (10+15q+10q^2)N_1^3 - (5+5q)N_1^4 + N_1^5
\end{aligned}$$

Theorem 4.1.

$$N_k = \sum_{i=1}^{k} (-1)^{i-1} P_{i,k}(q) N_1^i$$

where the $P_{i,k}$'s are polynomials with positive integer *coefficients.*

This theorem is proved by Garsia using induction and the fact that the sequence of N_k's satisfy a simple recurrence. For the details, see [GM, Chap. 7]. This result motivates the following combinatorial question:

Question 4.2. What are the objects that the family of polynomials, $\{P_{i,k}\}$, enumerate?

We will answer this questions in due course, in multiples ways, thus providing an alternate proof of Theorem 4.1.

4.1 First answer to Question 4.2

In this section we provide two different combinatorial interpretations for the coefficients of the P_k's.

4.1.1 The Lucas numbers and a (q, t)-analogue

Definition 4.3. Let $S_1^{(n)}$ be the circular shift of set $S \subseteq \{1, 2, \ldots, n\}$ modulo n, i.e. element $x \in S_1^{(n)}$ if and only if $x - 1 \pmod{n} \in S$. We define the $(q, t)-$**Lucas numbers** to be the sequence of polynomials in variables q and t

$$L_n(q,t) = \sum_{S \subseteq \{1,2,\ldots,n\} \;:\; S \cap S_1^{(n)} = \emptyset} q^{\#\text{ even elements in } S} \; t^{\lfloor \frac{n}{2} \rfloor - \#S}. \tag{4.1}$$

Note that this sum is over subsets S with no two numbers circularly consecutive.

These polynomials are a generalization of the sequence of Lucas numbers L_n which have the initial conditions $L_1 = 1$, $L_2 = 3$ (or $L_0 = 2$ and $L_1 = 1$) and satisfy the Fibonacci recurrence $L_n = L_{n-1} + L_{n-2}$. The first few Lucas numbers are

$$1, 3, 4, 7, 11, 18, 29, 47, 76, 123, \ldots$$

As described in numerous sources, e.g. [BY06], L_n is equal to the number of ways to color an $n-$beaded necklace black and white so that no two black beads are consecutive. You can also think of this as choosing a subset of $\{1, 2, \ldots, n\}$ with no consecutive elements, nor the pair $1, n$. (We call this circularly consecutive.) Thus letting q and t both equal one, we get by definition that $L_n(1, 1,) = L_n$.

We will prove the following theorem, which relates our newly defined $(q, t)-$Lucas numbers to the polynomials of interest, namely the N_k's.

Theorem 4.4.

$$1 + q^k - N_k = L_{2k}(q, -N_1) \tag{4.2}$$

for all $k \geq 1$.

To prove this result it suffices to prove that both sides are equal for $k \in \{1, 2\}$, and that both sides satisfy the same three-term recurrence relation. Since

$$\begin{aligned} L_2(q,t) &= 1 + q + t \qquad \text{and} \\ L_4(q,t) &= 1 + q^2 + (2q+2)t + t^2 \end{aligned}$$

we have proven that the initial conditions agree. Note that the sets of (4.1) yielding the terms of these sums are respectively

$$\{1\},\ \{2\},\ \{\ \} \quad \text{and} \quad \{1,3\},\ \{2,4\},\ \{1\},\ \{2\},\ \{3\},\ \{4\},\ \{\ \}.$$

It remains to prove that both sides of (4.2) satisfy the recursion

$$G_{k+1} = (1 + q - N_1)G_k - qG_{k-1}$$

for $k \geq 1$.

Proposition 4.5. *For the* $(q, t)-$*Lucas Numbers* $L_k(q, t)$ *defined as above,*

$$L_{2k+2}(q, t) = (1 + q + t)L_{2k}(q, t) - qL_{2k-2}(q, t). \tag{4.3}$$

Proof. To prove this we actually define an auxiliary set of polynomials, $\{\tilde{L}_{2k}\}$, such that

$$L_{2k}(q, t) = t^k \tilde{L}_{2k}(q, t^{-1}).$$

Thus recurrence (4.3) for the L_{2k}'s translates into

$$\tilde{L}_{2k+2}(q,t) = (1+t+qt)\tilde{L}_{2k}(q,t) - qt^2\tilde{L}_{2k-2}(q,t)$$

for the $\tilde{L}_{2k}$'s. The $\tilde{L}_{2k}$'s happen to have a nice combinatorial interpretation also, namely

$$\tilde{L}_{2k}(q,t) = \sum_{S\subseteq\{1,2,\ldots,2k\}\ :\ S\cap S_1^{(2k)}=\emptyset} q^{\#\text{ even elements in } S}\ t^{\#S}.$$

Recall our slightly different description which considers these as the generating function of 2-colored, labeled necklaces. We will find this terminology slightly easier to work with. We can think of the beads labeled 1 through $2k+2$ to be constructed from a pair of necklaces; one of length $2k$ with beads labeled 1 through $2k$, and one of length 2 with beads labeled $2k+1$ and $2k+2$.

Almost all possible necklaces of length $2k+2$ can be decomposed in such a way since the coloring requirements of the $2k+2$ necklace are more stringent than those of the pairs. However not all necklaces can be decomposed this way, nor can all pairs be pulled apart and reformed as a $(2k+2)$-necklace.

In Figure 4.1 the first necklace is decomposable but the second one is not since black beads 1 and 4 would be adjacent, thus violating the rule. It is clear enough that the number of pairs is $\tilde{L}_2(q,t)\tilde{L}_{2k}(q,t) = (1+t+qt)\tilde{L}_{2k}(q,t)$. To get the third term of the recurrence, i.e. $qt^2\tilde{L}_{2k-2}$, we must define linear analogues, $\tilde{F}_n(q,t)$'s, of the previous generating function. Just as the $\tilde{L}_n(1,1)$'s were Lucas numbers, the $\tilde{F}_n(1,1)$'s will be Fibonacci numbers.

Definition 4.6. The (twisted) $(q,t)-$Fibonacci polynomials, denoted as $\tilde{F}_n(q,t)$, are defined as

$$\tilde{F}_k(q,t) = \sum_{S\subseteq\{1,2,\ldots,k-1\}\ :\ S\cap(S_1^{(k-1)}-\{1\})=\emptyset} q^{\#\text{ even elements in } S}\ t^{\#S}.$$

The summands here are subsets of $\{1,2,\ldots,k-1\}$ such that no two elements are *linearly* consecutive, i.e. we now allow a subset with both the first and last elements. An alternate description of the objects involved are as (linear) chains of $k-1$ beads which are black or white with no two consecutive black beads.

For example, if $k = 2$:

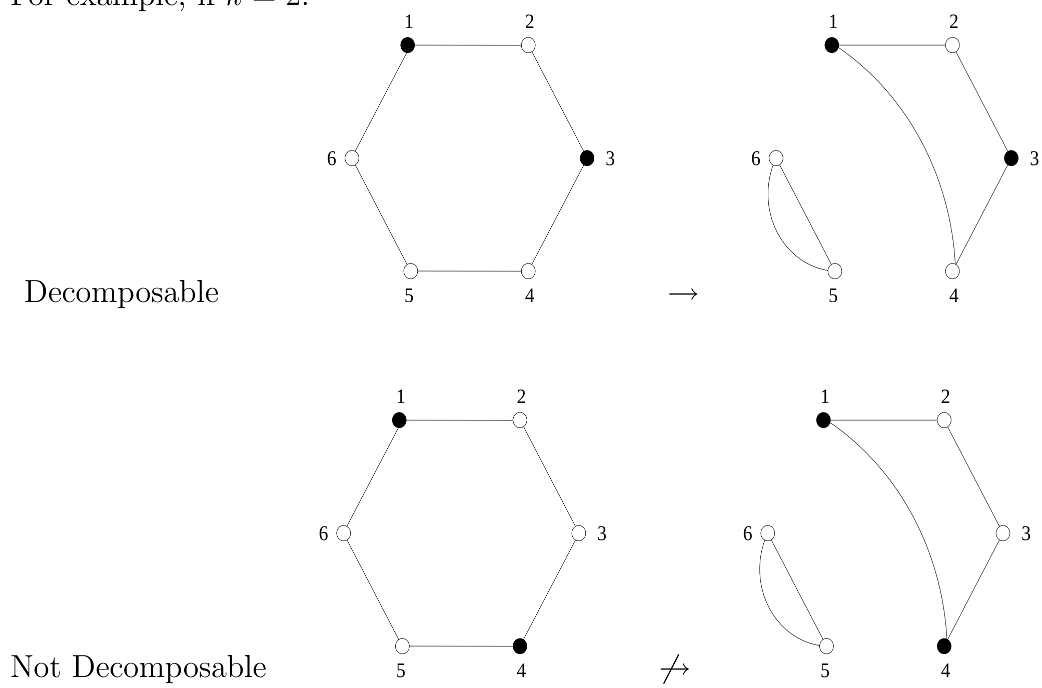


Figure 4.1: Illustrating proof of Proposition 4.5.

With these new polynomials at our disposal, we can calculate the third term of the recurrence, which is the difference between the number of pairs that cannot be recombined and the number of necklaces that cannot be decomposed.

Lemma 4.7. *The number of pairs that cannot be recombined into a longer necklace is* $2qt^2\tilde{F}_{2k-2}(q,t)$.

Proof. We have two cases: either both 1 and $2k+2$ are black, or both $2k$ and $2k+1$ are black. These contribute a factor of qt^2, and imply that beads 2, $2k$, and $2k+1$ are white, or that 1, $2k-1$, and $2k+2$ are white, respectively. In either case, we are left counting chains of length $2k-3$, which have no consecutive black beads. In one case we start at an odd-labeled bead and go to an evenly labeled one, and the other case is the reverse, thus summing over all possibilities yields the same generating function in both cases. $\square$

Lemma 4.8. *The number of* $(2k+2)$*-necklaces that cannot be decomposed into a* 2*-necklace and a* $2k$*-necklace is* $qt^2\tilde{F}_{2k-3}(q,t)$.

Proof. The only ones that cannot be decomposed are those which have beads 1 and $2k$ both black. Since such a necklace would have no consecutive black beads, this implies that beads 2, $2k-1$, $2k+1$, and $2k+2$ are all white. Thus we are reduced to looking at chains of length $2k-4$, starting at an odd, 3, which have no consecutive black beads. □

Lemma 4.9. *The difference of the quantity referred to in Lemma 4.8 from the quantity in Lemma 4.7 is exactly $qt^2\tilde{L}_{2k-2}(q,t)$.*

Proof. It suffices to prove the relation

$$qt^2\tilde{L}_{2k-2}(q,t) = 2qt^2\tilde{F}_{2k-2}(q,t) - qt^2\tilde{F}_{2k-3}(q,t)$$

which is equivalent to

$$qt^2\tilde{L}_{2k-2}(q,t) = qt^2\tilde{F}_{2k-2}(q,t) + q^2t^3\tilde{F}_{2k-4}(q,t) \tag{4.4}$$

since

$$\tilde{F}_{2k-2}(q,t) = qt\tilde{F}_{2k-4}(q,t) + \tilde{F}_{2k-3}(q,t). \tag{4.5}$$

Note that identity (4.5) simply comes from the fact that the $(2k-2)$nd bead can be black or white. Finally we prove (4.4) by dividing by qt^2, and then breaking it into the cases where bead 1 is white or black. If bead 1 is white, we remove that bead and cut the necklace accordingly. If bead 1 is black, then beads 2 and $2k+2$ must be white, and we remove all three of the beads. □

With this lemma proven, the recursion for the $\tilde{L}_{2k}$'s, hence the L_{2k}'s follows immediately. □

Proposition 4.10. *For an elliptic curve C with N_k points over $\mathbb{F}_{q^k}$ we have that*

$$1 + q^{k+1} - N_{k+1} = (1+q-N_1)(1+q^k-N_k) - q(1+q^{k-1}-N_{k-1}).$$

Proof. Recalling that for an elliptic curve C we have the identity

$$N_k = 1 + q^k - \alpha_1^k - \alpha_2^k,$$

we can rewrite the statement of this proposition as

$$\alpha_1^{k+1} + \alpha_2^{k+1} = (\alpha_1 + \alpha_2)(\alpha_1^k + \alpha_2^k) - q(\alpha_1^{k-1} + \alpha_2^{k-1}). \tag{4.6}$$

Noting that $q = \alpha_1\alpha_2$ we obtain this proposition after expanding out algebraically the right-hand-side of (4.6). □

With the proof of Propositions 4.5 and 4.10, we have proven Theorem 4.4.

4.1.2 $(q,t)-$Wheel numbers

Given that we found the Lucas numbers are related to the polynomial formulas $N_k(q, N_1)$, a natural question concerns how alternative interpretations of the Lucas numbers can help us better understand N_k. As noted in [BY06], [Mye71], and [Slo, Seq. A004146], the sequence $\{L_{2n} - 2\}$ counts the number of spanning trees in the wheel graph W_n; a graph which consists of $n+1$ vertices, n of which lie on a circle and one vertex in the center, a hub, which is connected to all the other vertices.

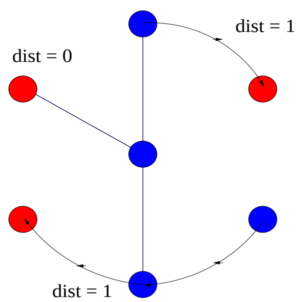
Definition 4.11. An undirected graph $G = (V, E)$ is defined by vertex set V and an edge set E consisting of pairs (v_i, v_j) where v_i and $v_j \in V$. A subgraph of G is defined as $G' = (V', E')$ where V' is a subset of V and E' is a subset of E consisting of edges using only vertices of V. A **spanning tree** of graph G is a connected subgraph G' (there exists a path from any vertex to another using the edges of G') which contains no cycles, i.e. there is exactly one path from one vertex to another.

We note that a spanning tree T of W_n consists of spokes and a collection of disconnected arcs on the rim. Further, since there are no cycles and T is connected, each spoke will intersect exactly one arc. (Since it will turn out to be convenient in the subsequent considerations, we make the – somewhat counter-intuitive – convention that an isolated vertex is considered to be an arc of length 1, and more generally, an arc consisting of k *vertices* is considered as an arc of *length* k.) We imagine the circle being oriented clockwise, and imagine the tail of each arc being the vertex which is the sink for that arc. In the case of an isolated vertex, the lone vertex is the tail of that arc. Since the spoke intersects each arc exactly once, if an arc has length k, meaning that it contains k vertices, there will be k choices of where the spoke and the arc meet. We define the $q-$weight of an arc to be $q^{\text{ number of edges between the spoke and the tail}}$, abbreviating this exponent as $spoke-tail$ distance. We define the $q-$weight of the tree to be the product of the $q-$weights for all arcs on the rim of the tree. This combinatorial interpretation motivates the following definition.

Definition 4.12.

$$\mathcal{W}_n(q,t) = \sum_{T \text{ a spanning tree of } W_n} q^{\text{sum of spoke}-\text{tail distance in } T}\, t^{\#\text{ spokes of } T}.$$

Here the exponent of t counts the number of edges emanating from the central vertex, and the exponent of q is as above.

q^2t^3 

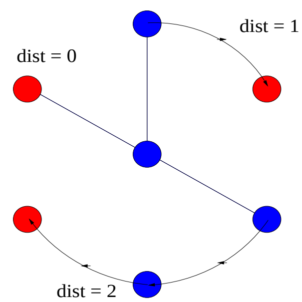
q^3t^3 

Figure 4.2: Illustrating definition of $\mathcal{W}_n(q,t)$.

This definition actually provides exactly the generating function that we desired.

Theorem 4.13.

$$N_k = -\mathcal{W}_k(q, -N_1)$$

for all $k \geq 1$.

Notice that this yields an exact interpretation of the $P_{i,k}$ polynomials as follows:

$$P_{i,k}(q) = \sum_{T \text{ a spanning tree of } W_n \text{ with exactly } i \text{ spokes}} q^{\text{sum of spoke}-\text{tail distance in } T}.$$

We will prove this theorem in two different ways. The first method will utilize Theorem 4.4 and an analogue of the bijection given in [BY06] which relates perfect and imperfect matchings of the circle of length $2k$ and spanning trees of W_k. Our second proof will use the observation that we can categorize the spanning trees based on the sizes of the various connected arcs on the rims. Since this categorization will correspond to partitions, this method will exploit formulas for decomposing power symmetric function p_k into a linear combination of h_λ's, as described in Chapter 2.

4.1.3 First proof of Theorem 4.13: Bijective

There is a simple bijection between subsets (of size at most $n-1$) of $\{1, 2, \ldots, 2n\}$ with no two elements circularly consecutive and spanning trees of the wheel graph W_n. We will use this bijection to give our first proof of Theorem 4.13. The bijection is as follows:

Given a subset S of the set $\{1, 2, \ldots, 2n-1, 2n\}$ with no circularly consecutive elements, we define the corresponding spanning tree T_S of W_n (with the correct q and t weight) in the following way:

1) We will use the convention that the vertices of the graph W_n are labeled so that the vertices on the rim are w_1 through w_n, and the central vertex is w_0.

2) We will exclude the two subsets which consist of all the odds or all the evens from this bijection. Thus we will only be looking at subsets which contain $n-1$ or fewer elements.

3) For $1 \leq i \leq n$, an edge exists from w_0 to w_i if and only if neither $2i-2$ nor $2i-1$ (element 0 is identified with element $2n$) is contained in S.

4) For $1 \leq i \leq n$, an edge exists from w_i to w_{i+1} (w_{n+1} is identified with w_1) if and only if element $2i-1$ or element $2i$ is contained in S.

Proposition 4.14. *Given this construction, T_S is in fact a spanning tree of W_n and further, tree T_S has the same $q-$weights and $t-$weights as set S.*

Proof. Suppose that set S contains k elements. From our above restriction, we have that $0 \leq k \leq n-1$. Since S is a k-subset of a $2n$ element set with no circularly consecutive elements, there will be $(n-k)$ pairs $\{2i-2, 2i-1\}$ with neither element in set S, and k pairs $\{2i-1, 2i\}$ with one element in set S. Consequently, subgraph

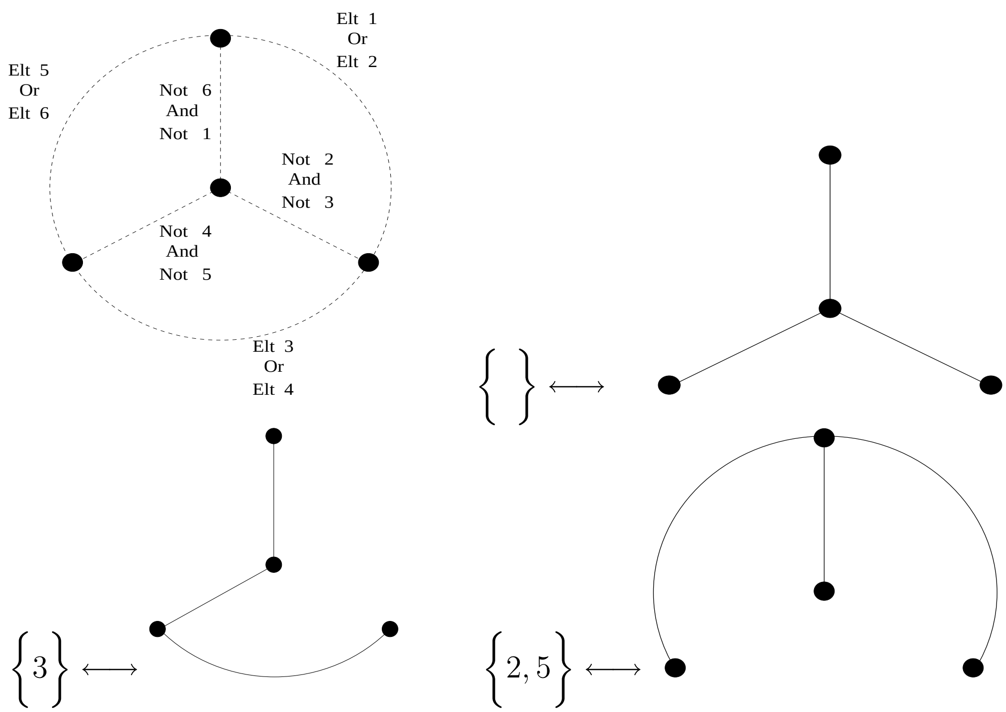


Figure 4.3: Illustrating bijection of Theorem 4.13.

T_S will consist of exactly $(n-k)+k = n$ edges. Since $n = (\#\text{ vertices of } W_n)-1$, to prove T_S is a spanning tree, it suffices to show that each vertex of W_n is included. For every oddly-labeled element of $\{1, 2, \ldots, 2n\}$, i.e. $2i-1$ for $1 \leq i \leq n$, we have the following rubric:

1) If $(2i-1) \in S$ then the subgraph T_S contains the edge from w_i to w_{i+1}.

2) If $(2i-1) \notin S$ and additionally $(2i-2) \notin S$, then T_S contains the spoke from w_0 to w_i.

3) If $(2i-1) \notin S$ and additionally $(2i-2) \in S$, then T_S contains the edge from w_{i-1} to w_i.

Since one of these three cases will happen for all $1 \leq i \leq n$, vertex w_i is incident to an edge in T_S. Also, the central vertex, w_0, has to be included since by our restriction, $0 \leq k \leq n-1$, there are $(n-k) \geq 1$ pairs $\{2i-2, 2i-1\}$ which contain no elements of S.

The number of spokes in T_S is $(n-k)$ which agrees with the $t-$weight of a set S with k elements. Finally, we prove that the q-weight is preserved, by induction on the number of elements in the set S. If set S has no elements, the $q-$weight should be q^0, and spanning tree T_S will consist of n spokes which also has $q-$weight q^0.

Now given a k element subset S $(0 \leq k \leq n-2)$, it is only possible to adjoin an odd number if there is a sequence of three consecutive numbers starting with an even, i.e. $\{2i-2, 2i-1, 2i\}$, which is disjoint from S. Such a sequence of S corresponds to a segment of T_S where a spoke and tail of an arc intersect. (Note this includes the case of vertex w_i being an isolated vertex.)

In this case, subset $S' = S \cup \{2i-1\}$ corresponds to $T_{S'}$, which is equivalent to spanning tree T_S except that one of the spokes w_0 to w_i has been deleted and replaced with an edge from w_i to w_{i+1}. The arc corresponding to the spoke from w_i will now be connected to the next arc, clockwise. Thus the distance between the spoke and the tail of this arc will not have changed, hence the $q-$weight of $T_{S'}$ will be the same as the $q-$weight of T_S.

Alternatively, it is only possible to adjoin an even number to S if there is a sequence $\{2i-1, 2i, 2i+1\}$ which is disjoint from S. Such a sequence of S corresponds to a segment of T_S where a spoke meets the *end* of an arc. (Note this includes the case of vertex w_i being an isolated vertex.)

Here, subset $S'' = S \cup \{2i\}$ corresponds to $T_{S''}$, which is equivalent to spanning tree T_S except that one of the spokes w_0 to w_{i+1} has been deleted and replaced with an edge from w_i to w_{i+1}. The arc corresponding to the spoke from w_{i+1} will now be connected to the *previous* arc, clockwise. Thus the cumulative change to the total distance between spokes and the tails of arcs will be an increase of one, hence the $q-$weight of $T_{S''}$ will be q^1 times the $q-$weight of T_S.

Since any subset S can be built up this way from the empty set, our proof is complete via this induction. □

Since the two sets we excluded, of size k had $(q,t)-$weights q^0t^0 and q^kt^0 respectively, we have proven Theorem 4.13.

4.1.4 Second proof of Theorem 4.13: Via generating function identities

For our second proof of Theorem 4.13, we consider writing the zeta function as an ordinary generating function instead, i.e.

$$Z(C,T) = 1 + \sum_{k\geq 1} H_k T^k. \tag{4.7}$$

In such a form, the H_k's are positive integers which enumerate the number of positive $C(\mathbb{F}_q)$-divisors of degree k, as noted in several places, such as [Mor91].

Proposition 4.15.

$$N_k = \sum_{\lambda\vdash k} (-1)^{l(\lambda)-1} \frac{k}{l(\lambda)} \binom{l(\lambda)}{d_1,\ d_2,\ \ldots\ d_m} \prod_{i=1}^{l(\lambda)} H_{\lambda_i}. \tag{4.8}$$

Proof. Comparing formulas (1.2) and (4.7) for $Z(C,T)$ and taking logarithms, we obtain

$$\frac{N_k}{k} \;=\; \log Z(C,T)\Big|_{T^k} = \log\left(1 + \sum_{n\geq 1} H_n T^n\right)\Big|_{T^k} = \sum_{m\geq 1} \frac{(-1)^{m-1}\left(\sum_{n=1}^{k} H_n T^n\right)^m}{m}\Bigg|_{T^k}.$$

To obtain the coefficient of T^k in

$$\left(H_1 T + H_2 T^2 + \cdots + H_k T^k\right)^m, \tag{4.9}$$

we first select a partition of k with length $\ell(\lambda) = m$. In other words, λ is a vector of positive integers satisfying $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m$. Each occurrence of $\lambda_i = j$ in this partition corresponds to choosing summand $H_j T^j$ in the ith term in product (4.9). Secondly, since the order of these terms does not matter, we include multinomial coefficients. Finally, multiplying through by k yields formula (4.8) for N_k. □

As we saw in Chapter 2, these identities between N_k and H_k are equivalent to those between p_k and h_k and thus the theory of symmetric functions also supplies a proof of Proposition 4.15 specializing to the genus one case.

We now specialize to the case of $g = 1$. Here we can write H_k in terms of N_1 and q. We expand the series

$$Z(E,T) = \frac{1-(1+q-N_1)T+qT^2}{(1-T)(1-qT)} = 1 + \frac{N_1 T}{(1-T)(1-qT)} \tag{4.10}$$

with respect to T, and obtain $H_0 = 1$ and $H_k = N_1(1+q+q^2+\cdots+q^{k-1})$ for $k \geq 1$. Plugging these into formula (4.8), we get polynomial formulas for N_k in terms of q and N_1

$$N_k = \sum_{\lambda \vdash k} (-1)^{l(\lambda)-1} \frac{k}{l(\lambda)} \binom{l(\lambda)}{d_1,\ d_2,\ \ldots\ d_k} \left(\prod_{i=1}^{l(\lambda)} (1+q+q^2+\cdots+q^{\lambda_i - 1}) \right) N_1^{l(\lambda)}.$$

Consequently, Theorem 4.13 is true if and only if we can replace N_1 with $-t$ and then multiply by (-1) and get a true expression for $\mathcal{W}_k$, the (q,t)-weighted number of spanning trees on the wheel graph W_k. We thus provide the following combinatorial argument for the required formula.

Proposition 4.16.

$$\mathcal{W}_k = \sum_{\lambda \vdash k} \frac{k}{l(\lambda)} \binom{l(\lambda)}{d_1,\ d_2,\ \ldots\ d_k} \left(\prod_{i=1}^{l(\lambda)} (1+q+q^2+\cdots+q^{\lambda_i - 1}) \right) t^{l(\lambda)}. \tag{4.11}$$

Proof. We will construct a spanning tree of W_k from the following choices: First we choose a partition $\lambda = 1^{d_1} 2^{d_2} \cdots k^{d_m}$ of k. We let this dictate how many arcs of each length occur, i.e. we have d_1 isolated vertices, d_2 arcs of length 2, etc. Note that this choice also dictates the number of spokes, which is equal to the number of arcs, i.e. the length of the partition.

Second, we pick an arrangement of the $l(\lambda)$ arcs on the circle. After picking one arc to start with, without loss of generality since we are on a circle, we have

$$\frac{1}{l(\lambda)} \binom{l(\lambda)}{d_1,\ d_2,\ \ldots\ d_m}$$

choices for such an arrangement. Third, we pick which vertex w_i of the rim to start with. There are k such choices. Fourth, we pick where the $l(\lambda)$ spokes actually intersect the arcs. There will be $|\mathrm{arc}|$ choices for each arc, and the $q-$weight of this sum will be $(1+q+q^2+\cdots+q^{|\mathrm{arc}|})$ for each arc. Summing up all the possibilities yields (4.11) as desired. $\square$

Thus we have given a second proof of Theorem 4.13.

4.2 More on bivariate Fibonacci polynomials via duality

In this section we explore further properties of various sequences of coefficients arising from the zeta function of a curve, and also more properties regarding bivariate Fibonacci polynomials. Our tools for such investigations will be two different manifestations of duality.

4.2.1 Duality between the symmetric functions h_k and e_k

Recall that in Section 4.1.1, we defined $\tilde{F}_k(q,t)$, i.e. the twisted (q,t)-Fibonacci polynomials. Here we define $F_k(q,t)$, an alternative bivariate analogue of the Fibonacci numbers. The definition of $F_k(q,t)$ is identical to that of $\tilde{F}_k(q,t)$ except for the weighting of parameter t.

Definition 4.17. We define the (q,t)-**Fibonacci polynomials** to be the sequence of polynomials in variables q and t given by

$$F_k(q,t) = \sum_{S \subseteq \{1,2,\ldots,k-1\} \ : \ S \cap (S_1^{(k-1)} - \{1\}) = \emptyset} q^{\#\text{ even elements in } S} \, t^{\lceil \frac{k}{2} \rceil - \#S}.$$

From this definition we obtain the following formulas for the E_k's in the elliptic case.

Theorem 4.18. *If C is a genus one curve, and the E_k's are as above, then for $n \geq 1$, $E_{-n} = 0$, $E_0 = 1$, and*

$$E_n = (-1)^n F_{2n-1}(q, -N_1)$$

where E_k and $F_k(q,t)$ are as defined above.

Before proving Theorem 4.18 we develop two key propositions.

Proposition 4.19. $F_{2n+1}(q,t) = (1+q+t)F_{2n-1}(q,t) - qF_{2n-3}(q,t)$ *for* $n \geq 2$.

Table 4.2: E_k, i.e. $F_{2k-1}(q,t)$'s for small k for the special case of an elliptic curve.

$$
\begin{aligned}
E_1 &= N_1 \\
E_2 &= -(1+q)N_1 + N_1^2 \\
E_3 &= (1+q+q^2)N_1 - (2+2q)N_1^2 + N_1^3 \\
E_4 &= -(1+q+q^2+q^3)N_1 + (3+4q+3q^2)N_1^2 - (3+3q)N_1^3 + N_1^4 \\
E_5 &= (1+q+q^2+q^3+q^4)N_1 - (4+6q+6q^2+4q^3)N_1^2 \\
&+ (6+9q+6q^2)N_1^3 - (4+4q)N_1^4 + N_1^5
\end{aligned}
$$

Proof. This follows the similar logic as the proof of Proposition 4.5 except we can use a more direct method. (One can use the t-weighting of the twisted (q,t)-Fibonacci polynomials instead to see this recursion more clearly, but we will omit this detour.) The polynomial F_{2n+1} is a (q,t)-enumeration of the number of chains of $2n$ beads, with each bead either black or white, and no two consecutive beads both black. Similarly $(1+q+t)F_{2n-1}$ enumerates the concatenation of such a chain of length $2n-2$ with a chain of length 2. One can recover a legal chain of length $2n$ this way except in the case where the $(2n-2)$nd and $(2n-1)$st beads are both black. Such cases are enumerated by qF_{2n-3} and this completes the proof. $\square$

Proposition 4.20. $(-1)^{n+1}E_{n+1} = (1+q-N_1)(-1)^n E_n - q(-1)^{n-1}E_{n-1}$ *for* $n \geq 2$.

Proof. We use the plethystic identity

$$e_k[A+B] = \sum_{i=0}^{k} e_i[A]e_{k-i}[B]$$

for any alphabets A and B. Setting $A = \alpha_1 + \alpha_2$ and $B = 1 + q - \alpha_1 - \alpha_2$, we derive

$$
\begin{aligned}
e_{n+1}[1+q] &= e_{n+1}[1+q-\alpha_1-\alpha_2] + (\alpha_1+\alpha_2)e_n[1+q-\alpha_1-\alpha_2] \\
&+ (\alpha_1\alpha_2)e_{n-1}[1+q-\alpha_1-\alpha_2] \\
&= E_{n+1} + (1+q-N_1)E_n + qE_{n-1}.
\end{aligned}
$$

Since $e_{n+1}[1+q] = 0$ for $n \geq 2$, we obtain the proposition as desired.

This result also follows directly from the generating function for the E_n's which is given by

$$\sum_{n\geq 0}(-1)^n E_n T^n = \frac{(1-T)(1-qT)}{1-(1+q-N_1)T+qT^2}.$$

The denominator of this series, also known as the series' characteristic polynomial, yields the desired linear recurrence for the coefficients of T^{n+1}, whenever $n+1$ exceeds the degree of the numerator.

□

With these two propositions verified, we can also now prove Theorem 4.18.

Proof of Theorem 4.18. It is clear that $E_1 = -F_1(q,-N_1)$, $E_2 = F_3(q,-N_1)$, and $E_3 = -F_5(q,-N_1)$. Propositions 4.19 and 4.20 show that both satisfy the same recurrence relations. Thus we have verified that

$$E_n = (-1)^n F_{2n-1}(q,-N_1).$$

□

Plethysm is a powerful tool and we utilize it below to obtain results of a similar flavor to Proposition 4.20.

Lemma 4.21. *Letting E_k be defined as $e_k[1+q-\alpha_1-\alpha_2]$ where α_1 and α_2 are roots of $T^2-(1+q-N_1)T+q$, we obtain*

$$h_k[\alpha_1+\alpha_2] = (-1)^k E_{k+1}/N_1.$$

Proof. We have for $n \geq 2$ that

$$N_1 E_n = E_{n+1} + (1+q)E_n + qE_{n-1}$$

since $(-1)^{n+1}E_{n+1} = (1+q-N_1)(-1)^n E_n - q(-1)^{n-1}E_{n-1}$ by Proposition 4.20. However by

$$e_k[A-B] = \sum_{i=0}^{k} e_i[A](-1)^{k-i}h_{k-i}[B],$$

we get

$$E_{n+1} = (-1)^{n+1}\Big(h_{n+1}[\alpha_1+\alpha_2] - (1+q)h_n[\alpha_1+\alpha_2] + qh_{n-1}[\alpha_1+\alpha_2]\Big)$$

using $A = 1+q$ and $B = \alpha_1 + \alpha_2$. After verifying initial conditions and comparing with

$$(-1)^{n+1}E_{n+1} = (-1)^{n+1}E_{n+2}/N_1 - (-1)^n(1+q)E_{n+1}/N_1 + (-1)^{n-1}qE_n/N_1$$

we get

$$h_{n+1}[\alpha_1 + \alpha_2] = (-1)^{n+1}E_{n+2}/N_1$$

by induction. □

With this result in mind, we obtain a table of symmetric function e_k and h_k in terms of various alphabets.

Table 4.3: Plethysm of e_k, h_k for elliptic curves.

poly. \ alphabet	$1+q-\alpha_1-\alpha_2$	$1+q$	$\alpha_1+\alpha_2$
e_k	E_k	$e_1 = 1+q,\ e_2 = q$	$e_1 = 1+q-N_1,\ e_2 = q$
h_k	H_k	$1+q+\cdots+q^k$	$(-1)^k E_{k+1}/N_1$

(We had earlier referred to E_k versus $\tilde{E}_k$ and H_k versus $\tilde{H}_k$ for plethysm in the alphabets $1+q-\alpha_1-\alpha_2$ and $\alpha_1+\alpha_2$, respectively.) Notice that the formulas for $e_k[1+q]$ and $h_k[1+q]$ are precisely the $N_1 = 0$ cases of $e_k[\alpha_1+\alpha_2]$ and $h_k[\alpha_1+\alpha_2]$. This should come at no surprise since 1 and q are the two roots of $T^2-(1+q)T+q$.

The plethystic equalities

$$h_k[A+B] = \sum_{i=0}^{k} h_i[A]h_{k-i}[B]$$

and

$$h_k[A-B] = \sum_{i=0}^{k} h_i[A](-1)^{k-i}e_{k-i}[B],$$

as well as the expressions for $e_k[A+B]$ and $e_k[A-B]$ used above, give rise to even more identities for different choices of A and B. We have focused on the ones that we have since they appeared most useful.

The above H_k–E_k (i.e. h_k–e_k) duality generalizes to the case of higher genus curves. However, considering the genus one case further, we take advantage of the simplicity of this particular generating function. Recall, as in (4.10), that by rewriting equation (1.14) we obtain

$$Z(E,T) = 1 + \frac{N_1 T}{(1-qT)(1-T)}$$

when E is an elliptic curve. As an application, we obtain an exponential generating function for the weighted number of spanning trees of the wheel graph,

$$W(q,N_1,T) = \exp\left(\sum_{k\geq 1} \mathcal{W}_k(q,N_1)\frac{T^k}{k}\right).$$

Using $\mathcal{W}_k = -N_k|_{N_1\to -N_1}$, and the fact this is an exponential, we obtain

$$W(q,N_1,T) = \frac{1}{1-\frac{N_1T}{(1-qT)(1-T)}} = \frac{(1-qT)(1-T)}{1-(1+q+N_1)T+qT^2}.$$

Also, rewriting $W(q,t,T)$as an ordinary generating function, we get

$$W(q,t,T) = \sum_{k\geq 0} E_k\Big|_{N_1\to -N_1}(-T)^k = 1 + \sum_{k\geq 1} F_{2k-1}(q,t)T^k.$$

Table 4.4: Plethystic dictionary for elliptic curves and spanning trees.

	Elliptic Curves	Spanning Trees
Generating Function	$\frac{1-(1+q-N_1)T+qT^2}{(1-qT)(1-T)}$	$\frac{(1-qT)(1-T)}{1-(1+q+N_1)T+qT^2}$
$1-(1+q\mp N_1)T+qT^2 =$	$(1-\alpha_1T)(1-\alpha_2T)$	$(1-\beta_1T)(1-\beta_2T)$
N_k (*resp.* $\mathcal{W}_k$)	$p_k[1+q-\alpha_1-\alpha_2]$	$p_k[-1-q+\beta_1+\beta_2]$
$H_k = N_1(1+q+\cdots+q^{k-1})$	$h_k[1+q-\alpha_1-\alpha_2]$	$(-1)^{k-1}e_k[-1-q+\beta_1+\beta_2]$
$(-1)^kE_k = F_{2k-1}(q,\mp N_1)$	$(-1)^ke_k[1+q-\alpha_1-\alpha_2]$	$h_k[-1-q+\beta_1+\beta_2]$

4.2.2 Duality between Lucas and Fibonacci numbers

In addition to the above discussion of how H_k and E_k are dual, this dictionary also highlights a comparison between *elliptic curve–spanning tree* duality and duality between Lucas numbers and Fibonacci numbers. As an application, we obtain a formula for E_k, i.e. $F_{2k-1}(q,t)$, in terms of the polynomial expansion for the $L_{2k}(q,t)$'s. If we recall our definition of $P_{i,k}$'s such that

$N_k = \sum_{i=1}^{k}(-1)^{i+1}P_{i,k}(q)N_1^i$, or equivalently $L_{2k}(q,t) = 1 + q^k + \sum_{i=1}^{k} P_{i,k}(q)t^i$, then we have

Proposition 4.22.

$$E_k = \sum_{i=1}^{k} \frac{(-1)^{k+i} \ \cdot \ i}{k} \ P_{i,k}(q)N_1^i.$$

To verify Proposition 4.22 we need the following combinatorial lemma, which describes a connection between the sets enumerated by Lucas numbers and those sets enumerated by Fibonacci numbers.

Lemma 4.23. *For $1 \leq i \leq k$ and $0 \leq j \leq i$, we have the number, which we denote as $c_{i,j}$, of subsets S_1 of $\{1, 2, \ldots, 2k\}$ with $k-i-j$ odd elements, j even elements, and no two elements circularly consecutive equals*

$$\frac{k}{i} \cdot \# \Big(\text{subsets } S_2 \text{ of } \{1, 2, \ldots, 2k-2\} \text{ with } k-i-j \text{ odd elments, } j \text{ even elements, and no two elements consecutive} \Big).$$

This notation might seem non-intuitive, but we use these indices so that the total number of elements is $k-i$ and the number of even elements is j. Thus the number of subsets S_1 (resp. S_2) directly describes the coefficient of $q^j t^i$ in $L_{2k}(q,t)$ (resp. $F_{2k-1}(q,t)$).

Proof. To prove this result we note that there is a bijection between the number of subsets of the first kind that do not contain $2k-1$ or $2k$ and those of the second kind. Thus it suffices to show that the number of sets S_1 which *do* contain element $2k-1$ or $2k$ is precisely fraction $\frac{k-i}{k}$ of all sets S_1 satisfying the above hypotheses.

Circularly shifting every element of set S_1 by an even amount r, i.e. $\ell \mapsto \ell + r - 1 \pmod{2k} + 1$, does not affect the number of odd elments and even elements. Furthermore, out of the k possible even shifts, $(k-i)$ of the sets, i.e. the cardinality of set S_1, will contain $2k-1$ or $2k$. This follows since for a given element ℓ there is exactly one shift which makes it $2k-1$ (or $2k$) if ℓ is odd (or even), respectively. Since elements cannot be consecutive, there is no shift that sends two different elements to both $2k-1$ and $2k$ simultaneously and thus we get the full $(k-i)$ possible shifts. $\square$

With this lemma proven, we can now show Proposition 4.22.

Proof of Proposition 4.22. We recall that

$$\begin{aligned}
\mathcal{W}_k(q,N_1) = L_{2k} - 1 - q^k &= \sum_{i=1}^{k} P_{i,k}(q) N_1^i = \sum_{i=1}^{k}\sum_{j=0}^{k} c_{i,j} N_1^i q^j \quad \text{and} \\
F_{2k-1}(q,-N_1) &= (-1)^k E_k.
\end{aligned}$$

Furthermore, we just showed via Lemma 4.23 that

$$F_{2k-1}(q,N_1) = \sum_{i=1}^{k}\sum_{j=0}^{k} \frac{i}{k} c_{i,j} N_1^i q^j = \sum_{i=1}^{k} \frac{i}{k} P_{i,k}(q) N_1^i.$$

Using Theorem 4.18 completes the proof. □

Remark 4.24. Alternatively, one can arrive at this result by directly manipulating the generating function. Namely, using the identities as above, we observe that $\frac{1}{Z(E,T)} = \sum_{n\geq 0}(-1)^n E_n T^n$, and so we have

$$\begin{aligned}
\sum_{n\geq 1}(-1)^n E_n T^n &= \frac{1}{Z(E,T)} - 1 = \frac{1}{1+\frac{N_1 T}{(1-qT)(1-T)}} - 1 = \sum_{n\geq 1}(-1)^n \left(\frac{N_1 T}{(1-qT)(1-T)}\right)^n \\
&= -N_1 \frac{\partial}{\partial N_1} \sum_{n\geq 1} \frac{(-1)^{n-1}}{n} \left(\frac{N_1 T}{(1-qT)(1-T)}\right)^n \\
&= -N_1 \frac{\partial}{\partial N_1} \left(\log\left(1 + \frac{N_1 T}{(1-qT)(1-T)}\right)\right) = -N_1 \frac{\partial}{\partial N_1} \log\Big(Z(E,T)\Big),
\end{aligned}$$

which equals $-N_1 \frac{\partial}{\partial N_1}\left(\sum_{k\geq 1} \frac{N_k}{k} T^k\right)$. Rewriting the N_k's using the polynomial formulas of Theorem 4.1, we have

$$\begin{aligned}
\sum_{n\geq 1}(-1)^n E_n T^n &= -N_1 \frac{\partial}{\partial N_1} \left(\sum_{k\geq 1} \frac{1}{k} \sum_{i=1}^{k} (-1)^{i-1} P_{i,k}(q) N_1^i T^k \right) \\
&= \sum_{k\geq 1}\sum_{i=1}^{k} \frac{i}{k} (-1)^i P_{i,k}(q) N_1^i T^k.
\end{aligned}$$

Comparing the coefficients of T^k on both sides completes the proof.

Lemma 4.23 also provides us a way to obtain expressions for $P_{i,k}(q)$, and in particular E_k and N_k, in terms of binomial coefficients.

Proposition 4.25. *For $k \geq 1$ and $1 \leq i \leq k$, we have*

$$P_{i,k}(q) = \sum_{j=0}^{i} \frac{k}{i} \binom{k-1-j}{i-1} \binom{i+j-1}{j} \, q^j.$$

Proof. See [Zel07, Theorem 2.2] or [MP07, Theorem 3] which show by algebraic and combinatorial arguments, respectively, that the number of ways to choose a subset $S \subset \{1, 2, \ldots, 2n\}$ such that S contains q odd elements, r even elements, and no consecutive elements is

$$\binom{n-r}{q}\binom{n-q}{r}.$$

Letting $n = k-1$, $q = k-i-j$ and $r = j$, we obtain

$$\frac{i}{k} P_{i,k}(q) = F_{2k-1}(q, N_1)\Big|_{N_1^i} = \sum_{j=0}^{i} \binom{k-1-j}{i-1}\binom{i+j-1}{j} q^j.$$

∎

Corollary 4.26.

$$N_k(q, N_1) = \sum_{i=1}^{k}\sum_{j=0}^{i} \frac{(-1)^{i+1} \cdot k}{i} \binom{k-1-j}{i-1}\binom{i+j-1}{j} N_1^i \, q^j.$$

and

$$E_k = \sum_{i=1}^{k}\sum_{j=0}^{i} (-1)^{k+i} \binom{k-1-j}{i-1}\binom{i+j-1}{j} N_1^i \, q^j.$$

Remark 4.27. From the proof in Section 4.1.4, we have that

$$\begin{aligned} \mathcal{W}_k(q, N_1) &= \sum_{\lambda \vdash k} \frac{k}{l(\lambda)} \binom{l(\lambda)}{d_1,\ d_2,\ \ldots\ d_r} \left(\prod_{i=1}^{l(\lambda)} (1 + q + q^2 + \cdots + q^{\lambda_i - 1}) \right) N_1^{l(\lambda)} \\ &= \sum_{i=1}^{k} \frac{k}{i} \left(\sum_{\substack{\lambda \vdash k \\ l(\lambda) = i}} \binom{i}{d_1,\ d_2,\ \ldots\ d_r} \prod_{j=1}^{i} (1 + q + q^2 + \cdots + q^{\lambda_j - 1}) \right) N_1^i \end{aligned}$$

which implies also that

$$P_{i,k}(q) = \frac{k}{i} \sum_{\substack{\lambda \vdash k \\ l(\lambda) = i}} \binom{i}{d_1,\ d_2,\ \ldots\ d_r} \prod_{j=1}^{i} (1 + q + q^2 + \cdots + q^{\lambda_j - 1}).$$

Comparing the coefficients of this identity with the coefficients in Proposition 4.25 seems to give a combinatorial identity that seems interesting in its own right.

We have just seen how N_k is equal to $p_k[1+q-\alpha_1-\alpha_2]$ plethystically and how this sequence relates to the sequences $H_k = h_k[1+q-\alpha_1-\alpha_2]$ and $E_k = e_k[1+q-\alpha_1-\alpha_2]$ via symmetric function theory. We close this section with a matrix determinant for $p_k[\alpha_1+\alpha_2] = 1+q^k-N_k$ from [GM, Chapter 7].

Proposition 4.28. $1+q^k-N_k$ *equals*

$$\det\begin{bmatrix} 1+q-N_1 & -1 & 0 & 0 & 0 & 0 \\ -2q & 1+q-N_1 & -1 & 0 & 0 & 0 \\ 0 & -q & 1+q-N_1 & -1 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & \cdots & 1+q-N_1 & -1 \\ 0 & 0 & 0 & \cdots & -q & 1+q-N_1 \end{bmatrix}$$

where this matrix is k-by-k. We denote this matrix as M'_k.

Proof. By the Newton Identities [Sta99], the power symmetric functions p_k can be rewritten in terms of the elementary symmetric functions e_k. In particular, $1+q^k-N_k = 1+q^k-p_k[1+q-\alpha_1-\alpha_2] = p_k[\alpha_1+\alpha_2]$ can be rewritten as

$$\det\begin{bmatrix} e_1[\alpha_1+\alpha_2] & -1 & 0 & 0 & 0 \\ -2e_2[\alpha_1+\alpha_2] & e_1[\alpha_1+\alpha_2] & -1 & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & 0 \\ (-1)^k(k-1)e_{k-1}[\alpha_1+\alpha_2] & (-1)^{k-1}e_{k-2}[\alpha_1+\alpha_2] & \cdots & e_1[\alpha_1+\alpha_2] & -1 \\ (-1)^{k+1}ke_k[\alpha_1+\alpha_2] & (-1)^k e_{k-1}[\alpha_1+\alpha_2] & \cdots & -e_2[\alpha_1+\alpha_2] & e_1[\alpha_1+\alpha_2] \end{bmatrix}.$$

Finally, since $e_1[\alpha_1+\alpha_2] = \alpha_1+\alpha_2 = 1+q-N_1, e_2[\alpha_1+\alpha_2] = \alpha_1\alpha_2 = q$, and $e_k[\alpha_1+\alpha_2] = 0$ for all $k \geq 2$, we have proven the proposition.

□

4.3 Case-Study on $N_2 = (2+2q)N_1 - N_1^2$

In this section, we investigate a method for understanding an elliptic curve E over a finite field $\mathbb{F}_{p^{2k}}$ (p prime) by understanding the elliptic curve restricted to $\mathbb{F}_{p^k}$ as well as a second curve over $\mathbb{F}_{p^k}$ which is known as the (quadratic) twist of $E(\mathbb{F}_{p^k})$. For convenience, we will take q to be p^k and assume $p \geq 5$, i.e. not char

2 or 3. This will allow us to write elliptic curve E as defined by the equation

$$y^2 = x^3 + ax + b$$

where $a, b \in \mathbb{F}_q$. We will let $f(x)$ denote $x^3 + ax + b$, E represent the set of all points with coordinates in the algebraic closure $\overline{\mathbb{F}_q}$, and let $E(\mathbb{F}_{q^n})$ denote the subset $E \cap \mathbb{F}_{q^n}^2$. One of the beauties of elliptic curves is that the sets E and $E(\mathbb{F}_{q^n})$ have additional structure, namely they are abelian groups whose addition we will denote as $\oplus$. By abuse of notation, E and $E(\mathbb{F}_{q^n})$ will signify these groups. We need to define one more operation, and then we will be able to state the main theorem of this section.

Definition 4.29. If $E(\mathbb{F}_q)$ is an elliptic curve with coefficients in $\mathbb{F}_q$ and $\Lambda \in \mathbb{F}_q$, let $E^{t(\Lambda)}(\mathbb{F}_q)$ represent the *quadratic twist* (with respect to Λ) of $E(\mathbb{F}_q)$ defined as follows: if E has equation $y^2 = f(x)$, then $E^{t(\Lambda)}$ has equation

$$y^2 = \Lambda f(x).$$

Proposition 4.30. *$E^{t(\Lambda)}(\mathbb{F}_q)$ is isomorphic to the curve with equation*

$$y'^{\,2} = x'^{\,3} + a\Lambda^{-2}x' + b\Lambda^{-3}.$$

Proof. If $y^2 = \Lambda(x^3 + ax + b)$, then letting $y = \Lambda^2 y'$, $x = \Lambda x'$, we obtain

$$y'^{\,2}\Lambda^4 = x'^{\,3}\Lambda^4 + ax'\Lambda^2 + b\Lambda$$

Dividing through by Λ^4, this becomes

$$y'^{\,2} = x'^{\,3} + a\Lambda^{-2}x' + b\Lambda^{-3}.$$

$\square$

Proposition 4.31. *If we have two elliptic curves over $\mathbb{F}_q$ in the simplified Weierstraß form, i.e.*

$$y^2 = x^3 + Ax + B \tag{4.12}$$
$$y^2 = x^3 + A'x + B' \tag{4.13}$$

then curve (4.12) $\cong$ curve (4.13) if and only if there exists $\omega \in \mathbb{F}_q \setminus \{0\}$ such that $A' = \omega^4 A$ and $B' = \omega^6 B$.

Proof. Two curves are isomorphic if we can change coordinates so that

$$\begin{aligned} x' &= \alpha x + \beta \\ y' &= \gamma x + \delta y + \epsilon \end{aligned}$$

but the only way we can do this so that $y'^{\,2}$ and $x'^{\,3}$ have the same coefficients while y', $x'y'$ and $x'^{\,2}$ have coefficients of zero is if β, γ, ϵ all equal 0, and $\alpha^2 = \delta^3$, which implies there exists $\omega = \frac{\delta}{\alpha}$ such that $\omega^{-2} = \delta$, $\omega^{-3} = \alpha$. Thus there exists $\omega \in \mathbb{F}_q \setminus \{0\}$ such that the transformation $x' = \omega^{-2}x, y' = \omega^{-3}y$ yields an isomorphic curve. After plugging in these into

$$y^2 = x^3 + Ax + B$$

and multiplying through by ω^6, we get the desired equation

$$y^2 = x^3 + \omega^4 Ax + \omega^6 B.$$

$\square$

Proposition 4.32. *If Λ is a square in $\mathbb{F}_q$, then $E^{t(\Lambda)}(\mathbb{F}_q) \cong E(\mathbb{F}_q)$.*

Proof. If $\Lambda = \lambda^2$ for $\lambda \in \mathbb{F}_q$, then we let $y = \lambda y'$ and obtain via this change of coordinates that $y'^{\,2} = f(x)$ whenever (x, y) satisfy $y^2 = \Lambda f(x)$. $\square$

Proposition 4.33. *If Λ is a non-square in $\mathbb{F}_q$, then $E^{t(\Lambda)}(\mathbb{F}_q) \ncong E(\mathbb{F}_q)$, but $E^{t(\Lambda)}(\mathbb{F}_q) \cong E^{t(\Lambda')}(\mathbb{F}_q)$ for any other $\Lambda' \in \mathbb{F}_q$ which is a non-square.*

Proof. The curve $E^{t(\Lambda)}(\mathbb{F}_q)$ is isomorphic to a curve with the equation

$$y'^{\,2} = x'^{\,3} + a\Lambda^{-2}x' + b\Lambda^{-3}.$$

This is the simplified Weierstraß form, and thus $E^{t(\Lambda)}(\mathbb{F}_q)$ is isomorphic to $E(\mathbb{F}_q)$ if only if there exists $\omega \in \mathbb{F}_q \setminus \{0\}$ such that $\Lambda^{-2} = \omega^4, \Lambda^{-3} = \omega^6$, which implies that Λ is a square over $\mathbb{F}_q$. $\Rightarrow\Leftarrow$ $\square$

In light of these results, we will drop the superscript (Λ) from our notation, and let $E^t(\mathbb{F}_q)$ represent $E^{t(\Lambda)}(\mathbb{F}_q)$ where Λ is any non-square of $\mathbb{F}_q$. We now come to the main result of this section.

Theorem 4.34. *If E is a non-singular elliptic curve with coefficients in $\mathbb{F}_q$, and $E^t(\mathbb{F}_q)$ is its quadratic twist over $\mathbb{F}_q$, as defined above, then*

$$|E(\mathbb{F}_{q^2})| = |E(\mathbb{F}_q)| \cdot |E^t(\mathbb{F}_q)|. \tag{4.14}$$

Furthermore, there is an explicit bijection between sets $E(\mathbb{F}_{q^2})$ and $E(\mathbb{F}_q) \times E^t(\mathbb{F}_q)$, as well as a group isomorphism in many cases.

We will prove this theorem in three steps. First we demonstrate the validity of equality (4.14) algebraically. Secondly, we provide an alternative proof of this identity by illustrating an explicit bijection between these two sets. We then discuss the problem of constructing a natural bijection and give a simple criterion for determining when we in fact have a group isomorphism. We begin algebraically.

4.3.1 Algebraic proof

Lemma 4.35. $|E^t(\mathbb{F}_q)| = 2q + 2 - |E(\mathbb{F}_q)|$.

Proof. This result appears several places in the literature, for example [Hus04] We provide a proof of this equality while introducing some new notation that will be used for the proof of Theorem 4.34.

As we saw previously, $f(\alpha)$ for $\alpha \in \mathbb{F}_q$ is either (1) a nonzero square modulo q, (2) a non-square modulo q, or (3) zero. We will let

$$\begin{aligned}
\mathcal{I}_1 &= \#\{\alpha \in \mathbb{F}_q : f(\alpha) = \text{ a nonzero square }\}, \\
\mathcal{I}_{-1} &= \#\{\alpha \in \mathbb{F}_q : f(\alpha) = \text{ a non}-\text{square }\}, \text{ and} \\
\mathcal{I}_0 &= \#\{\alpha \in \mathbb{F}_q : f(\alpha) = 0\}.
\end{aligned}$$

Since we have partitioned $\mathbb{F}_q$, $\mathcal{I}_1 + \mathcal{I}_0 + \mathcal{I}_{-1} = q$. Furthermore,

$$E(\mathbb{F}_q) = 2\mathcal{I}_1 + \mathcal{I}_0 + 1$$

since if $f(\alpha)$ is a nonzero square, $y^2 = f(\alpha)$ has exactly two solutions, $y^2 = 0$ has one solution, and $y^2 = f(\alpha)$, for $f(\alpha)$ a non-square has no solutions. We add one for the point at infinity. Additionally, we obtain

$$E^t(\mathbb{F}_q) = \mathcal{I}_0 + 2\mathcal{I}_{-1} + 1$$

since in this case we are solving $y^2 = \Lambda f(\alpha)$ for Λ a non-square in $\mathbb{F}_q$, and thus the roles of $\mathcal{I}_1$ and $\mathcal{I}_{-1}$ are switched. Consequently,

$$|E(\mathbb{F}_q)| + |E^t(\mathbb{F}_q)| = 2I_{-1} + 2I_0 + 2I_1 + 2 = 2q + 2.$$

See [Sta73] for more exposition on this notation. We now use our formula for $|E(\mathbb{F}_{q^2})|$ in terms of $|E(\mathbb{F}_q)|$ that we earlier obtained via the theory of the zeta function. $\square$

Lemma 4.36. *Using the notation of the above sections,*

$$N_2 = N_1 \cdot (2 + 2q - N_1) = (2q + 2)N_1 - N_1^2.$$

Proof. We can give a quick explicit proof of this fact alone from $E(\mathbb{F}_q)$'s zeta function. To do so, we use the following three relations:

$$\begin{aligned} N_2 &= 1 + q^2 - \alpha_1^2 - \alpha_2^2 \\ N_1 &= 1 + q - \alpha_1 - \alpha_2 \\ \alpha_1\alpha_2 &= q. \end{aligned}$$

Thus $\alpha_1 + \alpha_2 = 1 + q - N_1$, and hence

$$\alpha_1^2 + 2\alpha_1\alpha_2 + \alpha_2^2 = (1 + q - N_1)^2.$$

But on the other hand,

$$\alpha_1^2 + \alpha_2^2 = 1 + q^2 - N_2 \text{ and } \alpha_1\alpha_2 = q,$$

and solving for N_2 in terms of N_1 and q yields the desired result. $\square$

Piecing the last two results together, we obtain $|E(\mathbb{F}_q)| \cdot |E^t(\mathbb{F}_q)| = |E(\mathbb{F}_{q^2})|$.

4.3.2 The explicit bijection

We now wish to prove the existence of an explicit bijection. There will be small differences in the definition of the bijection depending on the value of $\mathcal{I}_0$, noting that $\mathcal{I}_0 \in \{0, 1, 3\}$ since $f(x)$ is a cubic with no multiple roots (E is non-singular). We will highlight those differences as they come up.

Because $\mathbb{F}_q$ is a subfield of $\mathbb{F}_{q^2}$, (in fact there are multiple embeddings), this implies that $E_1 = E(\mathbb{F}_q)$ is a subgroup of $E(\mathbb{F}_{q^2})$. Let E_1' denote the subset of $E(\mathbb{F}_{q^2})$ containing P_∞ as well as points of the form (x, Y) where $x \in \mathbb{F}_q, Y^2 \in \mathbb{F}_q$, but $Y \in \mathbb{F}_{q^2} \setminus \mathbb{F}_q$.

Remark 4.37. We can actually explicitly construct E_1' by fixing Λ to be a specific non-square of $\mathbb{F}_q$ and considering points of $E(\mathbb{F}_{q^2})$ of the form $(x, \lambda^{-1}y)$ such that $x, y \in \mathbb{F}_q$ and $\lambda \in \mathbb{F}_{q^2} \setminus \mathbb{F}_q$ satisfies $\lambda^2 = \Lambda$. If we choose Λ to be a different non-square (e.g. $\Lambda' = c^2\Lambda$ and $\lambda' = c\lambda$) then $(x, \lambda^{-1}y)$ would still have the form $(x, \lambda'^{-1}y')$ by letting $y' = cy \in \mathbb{F}_q$. Thus E_1' does not actually depend on the choice of Λ.

Lemma 4.38. *E_1' is actually a subgroup, as opposed to simply a subset.*

Proof. If $P_1 = (x_1, \lambda^{-1}y_1)$ and $P_2 = (x_2, \lambda^{-1}y_2)$, (with $x_1 \neq x_2$) then

$$P_1 \oplus P_2 = \Bigg(\frac{(y_2 - y_1)^2}{(x_2 - x_1)^2} \lambda^{-2} - (x_1 + x_2) \quad , \quad \frac{(x_2y_2 - x_1y_1 + 2x_1y_2 - 2x_2y_1)}{(x_2 - x_1)}\lambda^{-1} - \frac{(y_2 - y_1)^3}{(x_2 - x_1)^3}\lambda^{-3} \Bigg)$$

and

$$2P_1 = \left(\frac{(3x_1^2 + a)^2\lambda^2}{4y_1^2} - 2x_1, \quad -\frac{(3x_1^2 + a)^3\lambda^3}{8y_1^3} + \frac{3x_1(3x_1^2 + a)\lambda}{2y_1} - y_1\lambda^{-1} \right).$$

Since $\lambda^2 = \Lambda \in \mathbb{F}_q$, implies that $P_1 \oplus P_2$ and $2P_1$ both have desired form $(x_3, \lambda^{-1}y_3)$ with $x_3, y_3 \in \mathbb{F}_q$. Lastly, if we add $(x, \lambda^{-1}y_1)$ to $(x, \lambda^{-1}y_2)$ for $y_1 \neq y_2$, we get P_∞. $\square$

Lemma 4.39. *The group E_1' is isomorphic to $E^t(\mathbb{F}_q)$.*

Proof. By Proposition 4.30, $E^t(\mathbb{F}_q)$ is isomorphic to an equation of the form

$$y'^{\,2} = x'^{\,3} + \Lambda^{-2}ax' + \Lambda^{-3}b,$$

where $\Lambda \in \mathbb{F}_q$ is a non-square, via the transformations

$$y' = \Lambda^{-2}y \text{ and } x' = \Lambda^{-3}x.$$

Also $E^t(\mathbb{F}_{q^2})$ is isomorphic to $E(\mathbb{F}_{q^2})$ since Λ is a square in $\mathbb{F}_{q^2}$. Thus we have $E^t(\mathbb{F}_{q^2}) \cong E(\mathbb{F}_{q^2})$ which respectively have subgroups $E^t(\mathbb{F}_q)$ and E_1'. Furthermore if we let Ψ be the explicit isomorphism $(x, y) \mapsto (\lambda^{-2}x, \lambda^{-3}y)$ from $E^t(\mathbb{F}_{q^2})$ to $E(\mathbb{F}_{q^2})$, then

$$\Psi(E^t(\mathbb{F}_q)) \subset E_1'$$

since $\lambda^2 \in \mathbb{F}_q$ but $\lambda \notin \mathbb{F}_q$ and we get the opposite inclusion as Ψ^{-1} maps E_1' onto $E^t(\mathbb{F}_q)$. Thus Ψ is an isomorphism between $E^t(\mathbb{F}_q)$ and E_1'. $\square$

We note that E_1 and E_1' are both subgroups of $E(\mathbb{F}_{q^2})$, and thus we can define another subgroup of $E(\mathbb{F}_{q^2})$, namely $E_1 \cdot E_1'$, which is the group of elements of the form $P \oplus Q$ such that $P \in E_1$, $Q \in E_1'$. We have a surjective homomorphism

$$\phi : E(\mathbb{F}_q) \times E^t(\mathbb{F}_q) \to E_1 \cdot E_1' \leq E(\mathbb{F}_{q^2})$$

defined by

$$(P, Q) \mapsto P \oplus \Psi(Q).$$

It is a homomorphism since Ψ is an isomorphism and $P \mapsto P$ is the identity isomorphism, and it is surjective since by construction, $E_1 \cdot E_1'$ is the set of all elements of the form $P \oplus \Psi(Q)$.

Proposition 4.40. *If $\mathcal{I}_0 = 0$, then we have the equality of groups $E_1 \cdot E_1' = E(\mathbb{F}_{q^2})$, hence map ϕ is an isomorphism, and therefore a bijection, between*

$$E(\mathbb{F}_q) \times E^t(\mathbb{F}_q) \text{ and } E(\mathbb{F}_{q^2}).$$

Proof. Since $\mathcal{I}_0 = 0$, there are no points of the form $(x, 0)$ in either E_1 or E_1'. Thus all finite points of E_1 are different from the finite points of E_1', and vice-versa. Hence,

$$E_1 \cap E_1' = \{P_\infty\},$$

where P_∞ is the identity element of $E(\mathbb{F}_{q^2})$. Consequently, the Cartesian product $E_1 \times E_1'$ is isomorphic to $E_1 \cdot E_1'$. By the isomorphism $E_1 \cong E(\mathbb{F}_q)$ and $E_1' \cong E^t(\mathbb{F}_q)$, we obtain $E(\mathbb{F}_q) \times E^t(\mathbb{F}_q) \cong E_1 \cdot E_1'$.

Since $|E_1 \times E_1'| = |E(\mathbb{F}_q)| \cdot |E^t(\mathbb{F}_q)| = |E(\mathbb{F}_{q^2})|$, and $E_1 \cdot E_1' \leq E(\mathbb{F}_{q^2})$, the isomorphism $E_1 \times E_1' \cong E_1 \cdot E_1'$ implies that $|E_1 \cdot E_1'| = |E(\mathbb{F}_{q^2})|$, and consequently

$$E_1 \cdot E_1' = E(\mathbb{F}_{q^2}).$$

Since $|E(\mathbb{F}_q) \times E^t(\mathbb{F}_q)| = |E(\mathbb{F}_{q^2})|$ from earlier results, the surjective homomorphism ϕ between $E(\mathbb{F}_q) \times E^t(\mathbb{F}_q)$ and $E(\mathbb{F}_{q^2})$ must be an isomorphism. $\square$

In the case of $\mathcal{I}_0 = 1$, the cubic $f(x)$ factors as $(x - x_0)g(x)$ where g is an irreducible quadratic over $\mathbb{F}_q$, but over $\mathbb{F}_{q^2}$ the quadratic g splits and there exist $x_1, x_2 \in \mathbb{F}_{q^2} \setminus \mathbb{F}_q$ such that $(x_1, 0)$ and $(x_2, 0) \in E(\mathbb{F}_{q^2}) \setminus E(\mathbb{F}_q)$. Also $(x_0, 0)$ is an element of $E(\mathbb{F}_q)$, and all three of these have order 2 since the inverse of (x, y) is defined as $(x, -y)$ over $E(\mathbb{F}_q)$ or $E(\mathbb{F}_{q^2})$.

Proposition 4.41. *If $\mathcal{I}_0 = 1$ then ϕ is a 2-to-1 map. This is equivalent to proving $E_1 \cdot E_1'$ has index 2 in $E(\mathbb{F}_{q^2})$, or that ϕ has kernel $\{(P_\infty, P_\infty), ((x_0, 0), (x_0, 0))\}$. Furthermore, we can use surjective homomorphism ϕ to construct a map $\overline{\phi}$ from $E(\mathbb{F}_q) \times E(\mathbb{F}_{q^2})$ into all of $E(\mathbb{F}_{q^2})$ which is a bijection.*

Proof. We first show that if $R = P \oplus Q \in E_1 \cdot E_1' \leq E(\mathbb{F}_{q^2})$, then there exist unique $P' \neq P$ and $Q' \neq Q$ such that $R = P' \oplus Q'$. We let $P' = (x_0, 0) \oplus P$ and $Q' = (x_0, 0) \oplus Q$. It is clear that $P' \neq P$ and $Q' \neq Q$ are both satisfied since E_1 and E_1' are groups with identity P_∞. Furthermore $E_1 \cap E_1' = \{P_\infty, (x_0, 0)\}$ since $E_1 \ni (x_0, 0) = (x_0, 0 \cdot \lambda) \in E_1'$, but $(x, \lambda y) \notin E_1$ for all nonzero $y \in \mathbb{F}_q$. (Note that this gives an alternate proof that the point $(x_0, 0)$ has order two since $E_1 \cap E_1'$ is a closed subgroup.)

The group $E_1 \cdot E_1'$ is abelian so we can rewrite $P' \oplus Q'$ as

$$(x_0, 0) \oplus P \oplus (x_0, 0) \oplus Q = (x_0, 0) \oplus (x_0, 0) \oplus P \oplus Q = P \oplus Q.$$

If P'' and Q'' also satisfied $R = P'' \oplus Q''$ then $P \ominus P''$ would equal $Q'' \ominus Q$. However, one of these is an element of E_1 and one is an element of E_1', which implies $P \ominus P'' = Q'' \ominus Q \in \{P_\infty, (x_0, 0)\}$. Hence $P'' = P$ or P', and similarly $Q'' = Q$ or Q'.

Picking $\alpha \in E(\mathbb{F}_{q^2}) \setminus E_1 \cdot E_1'$, we next find that $E(\mathbb{F}_{q^2})$ decomposes as $E_1 \cdot E_1' \sqcup \alpha \oplus E_1 \cdot E_1'$. Note that this is a disjoint union since if there exists $P, P' \in E_1$ and

$Q, Q' \in E_1'$ such that $R = P \oplus Q = \alpha \oplus P' \oplus Q'$, then $\alpha = (P \ominus P') \oplus (Q \ominus Q') \in E_1 \cdot E_1'$, a contradiction. Furthermore, this union actually contains all of $E(\mathbb{F}_{q^2})$ since $|E_1 \cdot E_1'| = |E(\mathbb{F}_{q^2})|/2$.

Thus we can construct a bijection $\overline{\phi}$ between $E(\mathbb{F}_q) \times E^t(\mathbb{F}_q)$ and $E(\mathbb{F}_{q^2})$ by the following: for every coset of $E(\mathbb{F}_q) \times E^t(\mathbb{F}_q) \Big/ \Big\{(P_\infty, P_\infty), ((x_0, 0), (x_0, 0))\Big\}$, we pick one of the elements $\in \{(P, Q), (P \oplus (x_0, 0), Q \oplus (x_0, 0))\}$ and distinguish it from the other one. Let Γ be the set of distinguished elements. Then we define $\overline{\phi}$ piece-meal:

$$\begin{aligned} \Gamma &\rightarrow E_1 \cdot E_1' \\ \Big((x_0, 0), (x_0, 0)\Big) \oplus \Gamma &\rightarrow \alpha \oplus E_1 \cdot E_1' \quad \text{via the maps} \\ \beta &\mapsto \phi(\beta) \in E_1 \cdot E_1' \\ \Big((x_0, 0), (x_0, 0)\Big) \oplus \beta &\mapsto \alpha \oplus \phi(\beta) \in \alpha \oplus E_1 \cdot E_1' \end{aligned}$$

for $\beta \in \Gamma$. $\square$

Proposition 4.42. *If $\mathcal{I}_0 = 3$ then ϕ is a 4-to-1 map. This is equivalent to proving $E_1 \cdot E_1'$ has index 4 in $E(\mathbb{F}_{q^2})$, or that ϕ has kernel*

$$\{(P_\infty, P_\infty), ((x_0, 0), (x_0, 0)), ((x_1, 0), (x_1, 0)), ((x_2, 0), (x_2, 0))\}.$$

Furthermore, we can use surjective homomorphism ϕ to construct a map $\overline{\phi}$ from $E(\mathbb{F}_q) \times E(\mathbb{F}_{q^2})$ into all of $E(\mathbb{F}_{q^2})$ which is a bijection.

Proof. For this case, we will prove the result by computing the kernel of ϕ. We find that $\phi((P, Q)) = P_\infty$ if and only if $P \oplus \Psi(Q) = P_\infty$, where $P \in E_1, \Psi(Q) \in E_1'$. Since E_1 and E_1' are closed under inverses, both P and $\Psi(Q)$ must also be in $E_1 \cap E_1'$. Thus $P, \Psi(Q) \in \{P_\infty, (x_0, 0), (x_1, 0), (x_2, 0)\}$. However, P and $\Psi(Q)$ must be inverses and each of these choices are the identity or an involution, and thus we have the kernel as desired.

Picking $\alpha \in E(\mathbb{F}_{q^2}) \setminus E_1 \cdot E_1'$, $\beta \in E(\mathbb{F}_{q^2}) \setminus (E_1 \cdot E_1' \cup \alpha \oplus E_1 \cdot E_1')$, and $\gamma \in E(\mathbb{F}_{q^2}) \setminus (E_1 \cdot E_1' \cup \alpha \oplus E_1 \cdot E_1' \cup \beta \oplus E_1 \cdot E_1')$, we get that $E(\mathbb{F}_{q^2})$ decomposes as

$$E_1 \cdot E_1' \sqcup \alpha \oplus E_1 \cdot E_1' \sqcup \beta \oplus E_1 \cdot E_1' \sqcup \gamma \oplus E_1 \cdot E_1'.$$

Note that it is clear that we can successively pick α, β, and γ since $E_1 \cdot E_1'$ has index 4 in $E(\mathbb{F}_{q^2})$ This four-tuple is a disjoint union since if an element were in the intersection of any two of them, we would have an element of the form α, (respectively $\beta, \gamma, \beta \ominus \alpha, \gamma \ominus \alpha$, or $\gamma \ominus \beta$) would be in E_1, (respectively $E_1, E_1, \alpha E_1, \alpha E_1$, or βE_1), which would be a contradiction. Thus it is a union which spans $E(\mathbb{F}_{q^2})$ by comparing the sizes of $E_1 \cdot E_1'$ and $E(\mathbb{F}_{q^2})$.

Thus we can construct a bijection $\overline{\phi}$ between $E(\mathbb{F}_q) \times E^t(\mathbb{F}_q)$ and $E(\mathbb{F}_{q^2})$ analogous to the above construction: for every coset C_i of

$$E(\mathbb{F}_q) \times E^t(\mathbb{F}_q) \Big/ \Big\{ (P_\infty, P_\infty), ((x_0,0),(x_0,0)), ((x_1,0),(x_1,0)), ((x_2,0),(x_2,0)) \Big\},$$

we pick one of the elements of C_i and distinguish it from the other three. Let Γ be the set of distinguished elements. Then we define $\overline{\phi}$ piece-meal:

$$\begin{aligned}
\Gamma &\to E_1 \cdot E_1' \\
\Big((x_0,0),(x_0,0)\Big) \oplus \Gamma &\to (x_1,0) \oplus E_1 \cdot E_1' \\
\Big((x_1,0),(x_1,0)\Big) \oplus \Gamma &\to (x_1,0) \oplus E_1 \cdot E_1' \\
\Big((x_2,0),(x_2,0)\Big) \oplus \Gamma &\to (x_1,0) \oplus E_1 \cdot E_1' \quad \text{via the maps} \\
\omega &\mapsto \phi(\omega) \in E_1 \cdot E_1' \\
\Big((x_0,0),(x_0,0)\Big) \oplus \omega &\mapsto \alpha \oplus \phi(\omega) \in \alpha \oplus E_1 \cdot E_1' \\
\Big((x_1,0),(x_1,0)\Big) \oplus \omega &\mapsto \beta \oplus \phi(\omega) \in \beta \oplus E_1 \cdot E_1' \\
\Big((x_2,0),(x_2,0)\Big) \oplus \omega &\mapsto \gamma \oplus \phi(\omega) \in \gamma \oplus E_1 \cdot E_1'
\end{aligned}$$

for $\omega \in \Gamma$. □

Thus putting the last three propositions together, corresponding to the three cases $\mathcal{I}_0 = 0$, 1, or 3, we have proven Theorem 4.34, illustrating an explicit bijection yielding equality (4.14).

However, except for the case when $\mathcal{I}_0 = 0$, the bijection constructed was not necessarily an isomorphism, and was not natural (since it depends on the choice of coset representatives to place in distinguished set Γ). Consequently, in the next section we address this issue, providing a simple criterion for when an isomorphism between $E(\mathbb{F}_{q^2})$ and $E(\mathbb{F}_q) \times E^t(\mathbb{F}_q)$ exists, how to construct it in these cases, and what goes wrong in the other cases.

4.3.3 Determining when there is an isomorphism

Theorem 4.43. *If $\mathcal{I}_0 = 0$ or 1, then not only do we have a bijection but we have that*

$$|E(\mathbb{F}_q)|_2 = |E^t(\mathbb{F}_q)|_2 \iff E(\mathbb{F}_q) \times E^t(\mathbb{F}_q) \cong E(\mathbb{F}_{q^2}).$$

Here the notation $|G|_p$ signifies the exponent of p in cardinality $|G|$ (if group G contains $p^k m$ elements, with p and m relatively prime, then $|G|_p = k$). If $\mathcal{I}_0 = 3$, then $E(\mathbb{F}_q) \times E^t(\mathbb{F}_q)$ is never isomorphic to $E(\mathbb{F}_{q^2})$, though we always have an explicit bijection between them.

We prove this theorem by dividing it into cases. We begin, my noticing that in the case $\mathcal{I}_0 = 0$, that neither $E(\mathbb{F}_q)$ nor $E^t(\mathbb{F}_q)$ contain any points of the form $(x, 0)$, i.e. no elements of order two. Thus $|E(\mathbb{F}_q)|_2 = 0 = |E^t(\mathbb{F}_q)|_2$ in this case, and the hypotheses of Theorem 4.43 are satisfied for every elliptic curve E with $\mathcal{I}_0 = 0$. Furthermore, as seen in the proof of Proposition 4.40, we indeed have an isomorphism in this case. Turning our attention to the $\mathcal{I}_0 = 1$ case, the groups $E(\mathbb{F}_q)$ and $E^t(\mathbb{F}_q)$ both have a single element of order two, and thus have cyclic decompositions as

$$E(\mathbb{F}_q) \cong \mathbb{Z}_{2^k} \times G \text{ and } E^t(\mathbb{F}_q) \cong \mathbb{Z}_{2^{k'}} \times G'$$

where $|G|$ and $|G'|$ are both odd.Using the notation as above, we have subgroups of $E(\mathbb{F}_{q^2})$, E_1 and E_1', such that $E(\mathbb{F}_q) \cong E_1$, $E^t(\mathbb{F}_q) \cong E_1'$. We use these decompositions of E_1 and E_1' to describe the possible group structures for $E_1 \cdot E_1'$ and $E(\mathbb{F}_{q^2})$ explicitly.

Proposition 4.44. *If $\mathcal{I}_0 = 1$ and $E(\mathbb{F}_q) \cong \mathbb{Z}_{2^k} \times G$ and $E^t(\mathbb{F}_q) \cong \mathbb{Z}_{2^{k'}} \times G'$ where*

$|G|$ and $|G'|$ are both odd, then

$$E_1 \cdot E_1' \cong \left(\mathbb{Z}_{2^k} \cdot \mathbb{Z}_{2^{k'}}\right) \times G \times G' \quad (4.15)$$

$$\cong \mathbb{Z}_{2^{k-1}} \times \mathbb{Z}_{2^{k'}} \times G \times G'. \quad (4.16)$$

Furthermore,

$$E(\mathbb{F}_{q^2}) \cong \mathbb{Z}_{2^k} \times \mathbb{Z}_{2^{k'}} \times G \times G' \text{ or} \quad (4.17)$$

$$E(\mathbb{F}_{q^2}) \cong \mathbb{Z}_{2^{k-1}} \times \mathbb{Z}_{2^{k'+1}} \times G \times G'. \quad (4.18)$$

Proof. Since $E_1 \cap E_1' = \{P_\infty, (x_0, 0)\}$ contains elements of order one and two, we have that the subgroups G and G' of odd order satisfy $G \cap G' = \{P_\infty\}$, hence $G \cdot G' \cong G \times G'$. So after distributing the $\cdot$ over the $\times$, we obtain (4.15).

Let α signify a generator of $\mathbb{Z}_{2^k}$, and let β be a generator of $\mathbb{Z}_{2^{k'}}$. We then define element $\gamma \in E_1 \cdot E_1'$ to be $\alpha \oplus (2^{k'-k})\beta$. Notice that if $0 < d < 2^{k-1}$ then $d\alpha \in E_1, \notin E_1'$, and $(d \cdot 2^{k'-k})\beta \notin E_1, \in E_1'$. Thus $d\gamma = d\alpha \oplus (d \cdot 2^{k'-k})\beta$ is not the identity element of $E(\mathbb{F}_{q^2})$ in this case. However, if $d = 2^{k-1}$, then $(2^{k-1})\alpha$ is an element in E_1 of order two, hence $(x_0, 0)$, and $2^{k-1}(2^{k-k'})\beta = (2^{k'-1})\beta$ is an element in E_1' of order two, hence $(x_0, 0)$. Thus $d\gamma = (x_0, 0) \oplus (x_0, 0) = P_\infty$, and we conclude γ has order 2^{k-1}.

Let $\langle\alpha\rangle$ denote the cyclic subgroup of E_1 generated by α, $\langle\beta\rangle$ denote the cyclic subgroup of E_1' generated by β, and $\langle\gamma\rangle$ denote the cyclic subgroup of $E_1 \cdot E_1'$ generated by γ. We now need to show that

$$\langle\alpha\rangle \cdot \langle\beta\rangle = \langle\gamma\rangle \cdot \langle\beta\rangle \cong \langle\gamma\rangle \times \langle\beta\rangle.$$

We shall use multiplicative notation for our group to do so, i.e. we now write α^d to denote $d\alpha$, etc. We get the first equality since if we choose i between 0 and $2^{k-1} - 1$, and $j' = j - i(2^{k'-k}) \mod 2^{k'}$ between 0 and $2^{k'} - 1$, then $\gamma^i \oplus \beta^{j'} = \alpha^i \oplus \beta^{i(2^{k'-k})+j'} = \alpha^i \oplus \beta^j$. Furthermore, $\beta^{2^{k'-1}} = (x_0, 0) = \alpha^{2^{k-1}}$, thus restricting i so that $0 \le i \le 2^{k-1} - 1$ still includes all elements of $\langle\alpha\rangle \cdot \langle\beta\rangle$.

We get the second equality since $\gamma^d = \alpha^d \oplus \beta^{d(2^{k'-k})} \neq \beta^e$ for any value of d, e other than $\gamma^0 = P_\infty = \beta^0$ since more generally $\alpha^d \oplus \beta^{d'} = \beta^e$ implies $\alpha^d = \beta^{e'}$ and $\langle\alpha\rangle \cap \langle\beta\rangle = \{(x_0, 0), P_\infty\}$. However, since the order of γ is 2^{k-1}, we presume $d <$

2^{k-1} in which case P_∞ is the only point in the intersection, i.e. $\langle\gamma\rangle \cap \langle\beta\rangle = \{P_\infty\}$. Thus we have proven (4.16).

Now, since $E_1 \cdot E_1'$ has index two in $E(\mathbb{F}_{q^2})$, after doubling, we find that

$$\begin{aligned} E(\mathbb{F}_{q^2}) &\cong \mathbb{Z}_{2^k} \times \mathbb{Z}_{2^{k'}} \times G \times G' \text{ or} \\ E(\mathbb{F}_{q^2}) &\cong \mathbb{Z}_{2^{k-1}} \times \mathbb{Z}_{2^{k'+1}} \times G \times G' \text{ or} \\ E(\mathbb{F}_{q^2}) &\cong \mathbb{Z}_{2^{k-1}} \times \mathbb{Z}_{2^{k'-1}} \times \mathbb{Z}_2 \times G \times G'. \end{aligned}$$

However the third case is not actually possible since such a decomposition would imply that $E(\mathbb{F}_{q^2})$ would have more than three elements of order two, contradicting Corollary 3.21. Note that we do not encounter such a problem in (4.17) or (4.18) since even though these expressions are written as the decomposition of four or more cyclic subgroups, since $|G|$ and $|G'|$ are odd, G and G' can absorb $\mathbb{Z}_{2^k}$ and $\mathbb{Z}_{2^{k'}}$ into them respectively. ☐

We recall that in Section 4.3.2, in the case $\mathcal{I}_0 = 1$, we defined bijection $\overline{\phi}$ as

$$\begin{aligned} \Gamma &\rightarrow E_1 \cdot E_1' \\ \Big((x_0,0),(x_0,0)\Big) \oplus \Gamma &\rightarrow \alpha \oplus E_1 \cdot E_1' \quad \text{via the maps} \\ \beta &\mapsto \phi(\beta) \in E_1 \cdot E_1' \\ \Big((x_0,0),(x_0,0)\Big) \oplus \beta &\mapsto \alpha \oplus \phi(\beta) \in \alpha \oplus E_1 \cdot E_1' \end{aligned}$$

for $\beta \in \Gamma$, where α is an element of $E(\mathbb{F}_{q^2}) \setminus E_1 \cdot E_1'$ and Γ is a set of distinguished representatives of the cosets of $E(\mathbb{F}_q) \times E^t(\mathbb{F}_q) \Big/ \Big\{(P_\infty, P_\infty), ((x_0,0),(x_0,0))\Big\}$. In fact, we can say more.

Proposition 4.45. *If $\mathcal{I}_0 = 1$ and $|E(\mathbb{F}_q)| \equiv 2 \mod 4$ then we can pick Γ and α accordingly so that $\overline{\phi}$ is not only a bijection but an isomorphism of groups.*

Proof. Since $2q+2 \equiv 0 \mod 4$ for q odd we obtain $|E^t(\mathbb{F}_q)| \equiv 2 \mod 4$ if and only if $|E(\mathbb{F}_q)| \equiv 2 \mod 4$. Note that we know that $|E(\mathbb{F}_q)|$ (and $|E^t(\mathbb{F}_q)|$) are even when $\mathcal{I}_0 = 1$ since $|E(\mathbb{F}_q)| = 2\mathcal{I}_1 + \mathcal{I}_0 + 1$ and $|E^t(\mathbb{F}_q)| = 2\mathcal{I}_{-1} + \mathcal{I}_0 + 1$.

Thus $|E(\mathbb{F}_q)| = 2k$ for k odd, and $|E^t(\mathbb{F}_q)| = 2k'$ for k' odd. Hence as groups, $E(\mathbb{F}_q) \cong \mathbb{Z}_2 \times G$ and $E^t(\mathbb{F}_q) \cong \mathbb{Z}_2 \times G'$ with $|G|$ and $|G'|$ odd. Furthermore, since the only element of order two in either $E(\mathbb{F}_q)$ or $E^t(\mathbb{F}_q)$ is $(x_0, 0)$, we can write these explicitly as

$$\begin{aligned} E(\mathbb{F}_q) &= \Big\{P_\infty, (x_0, 0)\Big\} \cdot G \\ E^t(\mathbb{F}_q) &= \Big\{P_\infty, (x_0, 0)\Big\} \cdot G'. \end{aligned}$$

Hence $E(\mathbb{F}_q) \times E^t(\mathbb{F}_q)$ equals

$$\Big\{(P_\infty, P_\infty), (P_\infty, (x_0, 0)), ((x_0, 0), P_\infty), ((x_0, 0), (x_0, 0))\Big\} \cdot \Big(G \times G'\Big).$$

Consequently $E(\mathbb{F}_q) \times E^t(\mathbb{F}_q) \Big/ \Big\{(P_\infty, P_\infty), ((x_0, 0), (x_0, 0))\Big\}$ is isomorphic to

$$\Big\{(P_\infty, P_\infty), (P_\infty, (x_0, 0))\Big\} \cdot \Big(G \times G'\Big),$$

and we can choose the distinguished set Γ to be

$$\Big\{(P_\infty, P_\infty), (P_\infty, (x_0, 0))\Big\} \cdot \Big(G \times G'\Big)$$

for G and G' subgroups of $E(\mathbb{F}_q)$ and $E^t(\mathbb{F}_q)$ as defined above. Thus in this case Γ is not only a set but a group, thus $\phi : E(\mathbb{F}_q) \times E^t(\mathbb{F}_q) \to E_1 \cdot E_1'$ restricts to an isomorphism $\phi|_\Gamma$ from Γ to $E_1 \cdot E_1'$.

We can extend $\phi|_\Gamma$ to an isomorphism $\overline{\phi}$ from $E(\mathbb{F}_q) \times E^t(\mathbb{F}_q) \to E(\mathbb{F}_{q^2})$ by setting

$$\overline{\phi}\Big(((x_0, 0), (x_0, 0))\Big) = (x_1, 0) \notin E_1 \cdot E_1',$$

i.e. let $\alpha = (x_1, 0)$.

Note firstly that Γ and $E_1 \cdot E_1'$ are isomorphic, and so the number of elements of order two in each of them are the same. Since $G \times G'$ has odd order, Γ has only one element of order two, and consequently, $(x_0, 0)$ must be the only element of order two in $E_1 \cdot E_1'$. Hence $(x_1, 0), (x_2, 0) \notin E_1 \cdot E_1'$. Secondly, we have the decompositions $E(\mathbb{F}_q) \times E^t(\mathbb{F}_q) = \Gamma \sqcup ((x_0, 0), (x_0, 0)) \oplus \Gamma$ and $E(\mathbb{F}_{q^2}) = E_1 \cdot E_1' \sqcup (x_1, 0) \oplus E_1 \cdot E_1'$, and that map $\overline{\phi}$ is a bijection from earlier arguments. Thus to prove $\overline{\phi}$ is an isomorphism, it suffices to prove that $\overline{\phi}$ is a homomorphism, and since Γ is a group, $\overline{\phi}$ is a homomorphism if and only if

$$\overline{\phi}\Big(((x_0, 0), (x_0, 0)) \oplus \beta\Big) = (x_1, 0) \oplus \overline{\phi}(\beta) = (x_1, 0) \oplus \phi(\beta)$$

and

$$\overline{\phi}\Big(((x_0,0),(x_0,0)) \oplus \beta \oplus ((x_0,0),(x_0,0))\Big) = \phi(\beta).$$

Map $\overline{\phi}$ satisfies both of these since $((x_0,0),(x_0,0))$ and $(x_1,0)$ both have order two in their respective groups. $\square$

Alternatively, we could have mapped $((x_0,0),(x_0,0)) \mapsto (x_2,0)$ since

$$(x_1,0) \notin E_1 \cdot E_1' \iff (x_2,0) \notin E_1 \cdot E_1'$$

by $(x_0,0) \oplus (x_1,0) = (x_2,0)$ and the fact each of these three elements have order two.

Proposition 4.46. *If $\mathcal{I}_0 = 1$, $|E(\mathbb{F}_q)| \equiv 0 \mod 4$, and $|E(\mathbb{F}_q)|_2 = |E^t(\mathbb{F}_q)|_2$, then $E(\mathbb{F}_q) \times E^t(\mathbb{F}_q) \cong E(\mathbb{F}_{q^2})$ via the isomorphism φ which maps $E(\mathbb{F}_q) \times \{P_\infty\}$ to $E_1 \leq E(\mathbb{F}_{q^2})$, and sends $\beta \in E^t(\mathbb{F}_q)$ to $\gamma \in E_1 \cdot E_1'$, where β, γ are generators as described in the proof of Proposition 4.44.*

This case takes more work then the $|E(\mathbb{F}_q)| \equiv 2 \mod 4$ case. Namely, we begin with the following auxiliary results. For any group G and $n \in \mathbb{N}$, let $G[n]$ denote the subgroup of G consisting of elements with order dividing n, i.e. the n-torsion elements.

Lemma 4.47. *Let $|E(\mathbb{F}_q)|_2 = k$ and $|E^t(\mathbb{F}_q)|_2 = k'$, and assume without loss of generality that $k \leq k'$. Then $E(\overline{\mathbb{F}_q})[2^k] \subset E(\mathbb{F}_{q^2})$ if and only if the group decomposition of $E(\mathbb{F}_{q^2})$ is as in case (4.17).*

Proof. If we have (4.17), then $E(\mathbb{F}_{q^2})[2^k] \cong \mathbb{Z}_{2^k} \times \mathbb{Z}_{2^k}$, which contains all $(2^k)^2$ elements of $E(\overline{\mathbb{F}_q})[2^k]$. Thus $E(\mathbb{F}_{q^2})[2^k]$ is not only a subset of $E(\overline{\mathbb{F}_q})[2^k]$, but is actually equal to it. Thus

$$E(\mathbb{F}_{q^2}) \supset E(\mathbb{F}_{q^2})[2^k] = E(\overline{\mathbb{F}_q})[2^k].$$

On the other hand, if we do not have (4.17), then by above arguments, we must have (4.18), which implies that

$$E(\mathbb{F}_{q^2})[2^k] = \Big(\mathbb{Z}_{2^{k-1}} \times \mathbb{Z}_{2^{k'+1}}\Big)[2^k] = \mathbb{Z}_{2^{k-1}} \times \mathbb{Z}_{2^k}$$

since $k \leq k'$. Thus $\left|E(\mathbb{F}_{q^2})[2^k]\right| = 2^{k-1} \cdot 2^k$. However, $\left|E(\overline{\mathbb{F}_q})[2^k]\right| = (2^k)^2$, and so $E(\overline{\mathbb{F}_q})[2^k] \not\subset E(\mathbb{F}_{q^2})[2^k]$, hence $E(\overline{\mathbb{F}_q})[2^k] \not\subset E(\mathbb{F}_{q^2})$. $\square$

Lemma 4.48. *If $\mathcal{I}_0 = 1$ and k, k' signify $|E(\mathbb{F}_q)|_2$, $|E^t(\mathbb{F}_q)|_2$ respectively, then $k = k'$ if and only if (4.17).*

Proof. We assume that $k = k'$ and that (4.18) holds. Subgroup $E_1 \cdot E_1'$ has index two in $E(\mathbb{F}_{q^2})$ and is isomorphic to $\mathbb{Z}_{2^{k-1}} \times \mathbb{Z}_{2^k} \cong \langle\gamma\rangle \cdot \langle\beta\rangle \cong \langle\gamma\rangle \times \langle\beta\rangle$. However $E(\mathbb{F}_{q^2})$ is isomorphic to $E^t(\mathbb{F}_{q^2})$, this is a quadratic twist over $\mathbb{F}_q$ which is always a square in $\mathbb{F}_{q^2}$ regardless of whether or not it is a square in $\mathbb{F}_q$, and so we have $E_1 \cdot E_1' \cong \langle\gamma\rangle \cdot \langle\alpha\rangle \cong \langle\gamma\rangle \times \langle\alpha\rangle$ as well, switching the roles of $\langle\beta\rangle$ and $\langle\alpha\rangle$. In the case (4.18), β (resp. α), which has order 2^k, must have a square root in $E(\mathbb{F}_{q^2}) \setminus E_1 \cdot E_1'$, since $E(\mathbb{F}_{q^2}) \cong \mathbb{Z}_{2^{k-1}} \times \mathbb{Z}_{2^{k+1}}$.

This implies that there exists $\delta, \epsilon \in E(\mathbb{F}_{q^2}) \setminus E_1 \cdot E_1'$ such that $\delta^2 = \beta$ and $\epsilon^2 = \alpha$. Consequently, $\delta\epsilon$ is the square-root of $\alpha\beta$, which is γ when $k = k'$. Since γ has order 2^{k-1}, the element $\delta\epsilon$ has order 2^k. Matching orders, equation (4.18) implies that $E(\mathbb{F}_{q^2}) \cong \langle\gamma\rangle \cdot \langle\delta\rangle = \langle\gamma\rangle \cdot \langle\epsilon\rangle$, and we can write δ (resp. ϵ), which are elements of $E(\mathbb{F}_{q^2})$, in the form $\gamma^i\beta^j$, for j odd (resp. $\gamma^{i'}\alpha^{j'}$ for j' odd).

However, we have now reached a contradiction since

$$\delta^2\epsilon^2 = \gamma = \gamma^{2i+2i'}\beta^{2j}\alpha^{2j'} = \gamma^{2i+2i'+2j}\alpha^{2(j'-j)}$$

assuming without loss of generality that $j \leq j'$. However, $\langle\gamma\rangle \cap \langle\alpha\rangle = \{P_\infty\}$, hence $j = j'$ and

$$\gamma = \gamma^{2i+2i'+2j}.$$

But this is impossible since γ has even order and so γ^1 cannot be equal to γ^{2m} for any m.

Going the other direction, (4.17) implies that $E(\mathbb{F}_{q^2}) \cong \mathbb{Z}_{2^k} \times \mathbb{Z}_{2^{k'}}$. The order of γ is 2^{k-1} and $E_1 \cdot E_1' \cong \langle\gamma\rangle \times \langle\beta\rangle$, so there exists $\delta \in E(\mathbb{F}_{q^2}) \setminus E_1 \cdot E_1'$ such that $\delta^2 = \gamma = \alpha\beta^{2^{k'-k}}$. Now assume $k < k'$, which implies the exponent of β is even, and there exists element $\epsilon \in E(\mathbb{F}_{q^2}) \setminus E_1 \cdot E_1'$ satisfying $\epsilon^2 = \alpha$ (namely we let $\epsilon = \delta/\beta^{2^{k'-k-1}}$). Element $\epsilon \notin E_1 \cdot E_1'$ since $\beta \in E_1'$ and $E_1 \cdot E_1'$ is a subgroup of $E(\mathbb{F}_{q^2})$.

Thus $\delta\epsilon \in E_1 \cdot E_1' \cong \langle\gamma\rangle \times \langle\beta\rangle$, and δ of order 2^k, ϵ of order 2^{k+1}, so $\delta\epsilon$ has order 2^{k+1}. Hence $\delta\epsilon = \beta^i\gamma^j$ for $i \neq 0$. Also from definition of δ and ϵ, we get $\delta^2\epsilon^2 = \alpha^2\beta^{2^{k'-k}}$ hence we get the alternate representation

$$\delta\epsilon = \alpha\beta^{2^{k'-k-1}} = \gamma\beta^{2^{k'-k-1}-1},$$

which has an odd exponent of β and hence we get a contradiction analogous to the last case since elements in $\langle\gamma\rangle \times \langle\beta\rangle$ have unique representations. □

Proof of Proposition 4.46. We summarize these various results as follows.

Claim 4.49. Given that $\mathcal{I}_0 = 0$ or 1 and $E(\mathbb{F}_q) \cong \mathbb{Z}_{2^k} \times G$, $E^t(\mathbb{F}_q) \cong \mathbb{Z}_{2^{k'}} \times G'$, the following are equivalent:

- $k = k'$
- $E(\mathbb{F}_{q^2}) \cong \mathbb{Z}_{2^k} \times \mathbb{Z}_{2^{k'}} \times G \times G'$.
- $E(\mathbb{F}_{q^2}) \cong E(\mathbb{F}_q) \times E^(\mathbb{F}_q)$
- $E(\overline{\mathbb{F}_q})[2^k] \subset E(\mathbb{F}_{q^2})$

Claim 4.50. Given the same hypotheses, the following are equivalent:

- $k < k'$
- $E(\mathbb{F}_{q^2}) \cong \mathbb{Z}_{2^{k-1}} \times \mathbb{Z}_{2^{k'+1}} \times G \times G'$.
- $E(\mathbb{F}_{q^2}) \ncong E(\mathbb{F}_q) \times E^(\mathbb{F}_q)$
- $E(\overline{\mathbb{F}_q})[2^k] \not\subset E(\mathbb{F}_{q^2})$

In the literature [MOV93], an elliptic curve E satisfying $E(\overline{\mathbb{F}_q})[2^k] \subset E(\mathbb{F}_{q^2})$ is known as a curve with a certain embedding degree. Consequently Claims 4.49 and 4.50 therefore clearly delineate equivalent conditions and the ramifications on the group structure. □

To make this clearer, we note that if $\mathcal{I}_0 = 1$, $|E(\mathbb{F}_q)| \equiv 0 \mod 4$, and $|E(\mathbb{F}_q)|_2 \neq |E^t(\mathbb{F}_q)|_2$, then $E(\mathbb{F}_q) \times E^t(\mathbb{F}_q) \ncong E(\mathbb{F}_{q^2})$. Nonetheless, we obtain a bijection between them, and furthermore we know that

$$\begin{aligned} E(\mathbb{F}_q) &\cong \mathbb{Z}_{2^k} \times G \\ E(\mathbb{F}_q)^t &\cong \mathbb{Z}_{2^{k'}} \times G' \end{aligned}$$

for some $k, k' \geq 2$, such that $k \neq k'$ and $|G|, |G'|$ odd based on the hypotheses. Then

$$E(\mathbb{F}_{q^2}) \cong \mathbb{Z}_{2^{k-1}} \mathbb{Z}_{2^{k'+1}} \times G \times G'.$$

This follows since we proved previously that a bijection existed between them. However, in the case where $k \neq k'$, we have (4.18) by the above arguments and claims.

In the case where $\mathcal{I}_0 = 3$, the cubic $f(x)$ factors as $(x - x_0)(x - x_1)(x - x_2)$ over $\mathbb{F}_q$ and

$$E_1 \cap E_1' = \{P_\infty, (x_0, 0), (x_1, 0), (x_2, 0)\}.$$

Note that as a group $E_1 \cap E_1' \cong \mathbb{Z}_2 \times \mathbb{Z}_2$.

Proposition 4.51. *The groups $E(\mathbb{F}_q) \times E^t(\mathbb{F}_q)$ and $E(\mathbb{F}_{q^2})$ are never isomorphic when $\mathcal{I}_0 = 3$, but we do always obtain the bijection as previously seen.*

Proof. When $\mathcal{I}_0 = 3$, both $E(\mathbb{F}_q)$ and $E^t(\mathbb{F}_q)$ have three elements of order two. In fact $E(\mathbb{F}_q) \cap E^t(\mathbb{F}_q) = \left\{ P_\infty, (x_0, 0), (x_1, 0), (x_2, 0) \right\}$ where $(x_0, 0), (x_1, 0)$, and $(x_2, 0)$ are the three elements of order two. Thus

$$E(\mathbb{F}_q) \cong \mathbb{Z}_{2^a} \times \mathbb{Z}_{2^b} \times G \text{ and}$$

$$E^t(\mathbb{F}_q) \cong \mathbb{Z}_{2^c} \times \mathbb{Z}_{2^d} \times G'$$

for a, b, c, and $d \geq 1$. This means that $E(\mathbb{F}_q) \times E^t(\mathbb{F}_q)$ cannot be decomposed into less than four cyclic subgroups, but that contradicts Corollary 3.21. $\square$

Conjecture 4.52. *Just as in the $\mathcal{I}_0 = 1$ case, we can explicitly describe how to choose the representatives for the bijection. Namely, we can actually choose α, β, and γ to be elements of order 4 such that their squares are respectively $(x_0, 0), (x_1, 0)$, and $(x_2, 0)$ so that each of these square roots will live in disjoint cosets of $E_1 \cdot E_1'$.*

With these special cases complete, the proof of Theorem 4.43 is complete.

Conjecture 4.53. *In the case $\mathcal{I}_0 = 3$ the author conjectures that we still can describe the group decomposition explicitly, namely if we write*

$$\begin{aligned} E_1 &\cong \mathbb{Z}_a \times \mathbb{Z}_b \text{ and} \\ E_1' &\cong \mathbb{Z}_c \times \mathbb{Z}_d \end{aligned}$$

with $a \leq b$ and $c \leq d$, then

$$E(\mathbb{F}_{q^2}) \cong \mathbb{Z}_{ad} \times \mathbb{Z}_{bc}.$$

4.4 Geometric interpretations of fractions N_k/N_1

We now generalize the techniques of the previous section. The expressions for N_k, in terms of q and N_1, are always divisible by N_1 nd in the case $k = 2$ we saw $N_2 = N_1(2q + 2 - N_1)$ and $2q + 2 - N_1 = |E^t(\mathbb{F}_q)|$, the number of points (over $\mathbb{F}_q$) on the twist of elliptic curve E. This motivate the following query.

Question 4.54. Is there a geometric way to understand $\frac{N_k}{N_1}$ in general?

Theorem 4.55. *The quantity N_k/N_1 has a geometric interpretation as the number of points occurring in a prime divisor D such that $d \cdot D$ is linear equivalent to $k \cdot P_\infty$ for some $d|k$. Alternatively, we can think of this as the number of points $P \in E(\overline{\mathbb{F}_q})$ which satisfy the identity*

$$P + \pi(P) + \pi^2(P) + \cdots + \pi^{k-1}(P) \equiv kP_\infty.$$

However, before discussing how to prove this theorem via exact sequences and elliptic cyclotomic polynomials, as we will later on and in Section 5.3.2, we spend this section giving intuition and providing examples for small values of k.

We start by re-examining the $k = 2$ case. In this instance, the result states that N_2/N_1 should be the number of points $P \in E(\overline{\mathbb{F}_q})$ such that $P + \pi(P)$ is linearly equivalent to $2P_\infty$.

In the case where $P \in E(\mathbb{F}_q)$, we have $\pi(P) = P$ and this relation is equivalent to $2P \equiv 2P_\infty$, which is true if and only if $P = P_\infty$ or $(x_0, 0)$ for some $x_0 \in \mathbb{F}_q$. In other words, $2P \equiv 2P_\infty$ if and only if P is a point of order 1 or 2 in the group of the elliptic curve.

For a point $P \in E(\mathbb{F}_q \setminus \mathbb{F}_{q^2})$, P is not contained in any $1-$ or $2-$Frobenius cycle, and thus it would be impossible for such a point to satisfy $P + \pi(P) \equiv 2P_\infty$. Thus the only other possible points we have to consider are those contained in $E(\mathbb{F}_{q^2} \setminus \mathbb{F}_q)$ satisfying $P + \pi(P) \equiv 2P_\infty$. However, since $P \in E(\mathbb{F}_{q^2} \setminus \mathbb{F}_q)$ implies that $\pi(P) = -P$, i.e. $\pi\Big((x, y)\Big) = (x, -y)$, the only way this is true is if P lies on a vertical line $x = a$ for some $a \in \mathbb{F}_q$. This implies that P has an $x-$coordinate in $\mathbb{F}_q$ but a $y-$coordinate in $\mathbb{F}_{q^2} \setminus \mathbb{F}_q$.

Taking the union which includes the point at infinity, points of the form $(x_0, 0)$ and points of the form (a, β), we have exactly described the elements of $E^t(\mathbb{F}_q)$. Hence the theorem exactly agrees with the case we have previously discussed. Looking at

$$N_3/N_1 = 3(1 + q + q^2) - 3(1 + q)N_1 + N_1^2$$

we note that the terms on the right are three different ways of constructing a line in $\mathbb{P}^2(\overline{\mathbb{F}_q})$ whose defining equation has coefficients in $\mathbb{F}_q$.

$$\begin{aligned} 1 + q + q^2 &= \text{The number of projective lines of form } aX + bY + cZ = 0 \text{ with } a, b, c \in \mathbb{F}_q \\ (1 + q)N_1 &= \text{The number of ways to pick an } \mathbb{F}_q\text{−point, and slope, which determines a line} \\ N_1^2 &= \text{The number of ways to pick two points over } \mathbb{F}_q\text{, which will determine a line.} \end{aligned}$$

There are five kinds of lines we can have (analogous to the three kinds of vertical lines $x = a$ we had in the case $k = 2$, which were delineated by $\mathcal{I}_{-1}, \mathcal{I}_0$, and $\mathcal{I}_1$). Let $\mathcal{J}_{111}$ denote the number of lines (with defining equation having coefficients in $\mathbb{F}_q$) which go through three distinct points in $E(\mathbb{F}_q)$. Let $\mathcal{J}_{21}$ denote the number of lines which go through two distinct points in $E(\mathbb{F}_q)$, and is tangent with multiplicity two at one of them. Let $\mathcal{J}_3$ denote the number of lines which go through one point in $E(\mathbb{F}_q)$, and is an inflection point with multiplicity three. Let $\mathcal{J}^{21}$ denote the number of lines which go through one point in $E(\mathbb{F}_q)$ and two distinct points in $E(\mathbb{F}_{q^2} \setminus \mathbb{F}_q)$. Finally, let $\mathcal{J}^3$ denote the number of lines which go through three distinct points in $E(\mathbb{F}_{q^3} \setminus \mathbb{F}_q)$.

By comparing our three constructions of lines, we obtain

$$\begin{aligned} 1 + q + q^2 &= \mathcal{J}_{111} + \mathcal{J}_{21} + \mathcal{J}_3 + \mathcal{J}^{21} + \mathcal{J}^3 \\ (1+q)N_1 &= 3\mathcal{J}_{111} + 2\mathcal{J}_{21} + \mathcal{J}_3 + \mathcal{J}^{21} \\ N_1^2 &= 6\mathcal{J}_{111} + 3\mathcal{J}_{21} + \mathcal{J}_3 \end{aligned}$$

Consequently,

$$3(1+q+q^2) - 3(1+q)N_1 + N_1^2 = \mathcal{J}_3 + 3\mathcal{J}^3$$

and by noting the definitions of $\mathcal{J}_3$ and $\mathcal{J}^3$, we have now proven the theorem in the case of $k = 3$.

It appears the proof should work in general via this inclusion-exclusion- construction of rational functions technique. For example, in the case of $k = 4$, we should be computing the number of quadratics $aXZ + bX^2 + cYZ + dZ^2 = 0$ that can be constructed in various ways. To figure out which constructions we need to compare, we break-up the expression for N_4/N_1 according to partition, i.e.

$$N_4/N_1 = 4(1+q+q^2+q^3) - 4(1+q+q^2)N_1 - 2(1+q)^2N_1 + 4(1+q)N_1^2 - N_1^3$$

It is clear that there are eleven types of quadratics, depending on the number of points (with multiplicities) over the various subfields. Further $(1 + q + q^2 + q^3)$ and N_1^3 clearly count quadratics (3 points determine a quadratic), but not as clear why the other terms count the number of ways to construct a certain family of quadratics. Nonetheles, based on algebraic (as opposed to geometric) enumeration of these quantities based on their role as counting the number of positive divisors, we obtain

$$\begin{aligned}
(1+q+q^2+q^3) &= A_1+A_2+A_3+A_4+A_5+A_6+A_7+A_8+A_9+A_{10}+A_{11} \\
(1+q+q^2)N_1 &= A_1+2A_2+2A_3+3A_4+4A_5+2A_6+A_7+A_{10} \\
(1+q)^2N_1 &= A_1+2A_2+3A_3+4A_4+6A_5+2A_6+2A_7+2A_8+A_9 \\
(1+q)N_1^2 &= A_1+3A_2+4A_3+7A_4+12A_5+2A_6+A_7 \\
N_1^3 &= A_1+4A_2+6A_3+12A_4+24A_5.
\end{aligned}$$

Thus using the previous expression for N_4/N_1 as a weighted signed sum of these terms, we obtain

$$N_4/N_1 = A_1 + 2A_9 + 4A_{11}.$$

Here we enumerate the eleven types of quadratics in the following order:

A_1 through A_5 counts the number with all points in $E(\mathbb{F}_q)$ but varying multiplicities (all possible partitions of 4 in usual order 4, 31, 22, 211, 1111).

A_6 counts the number with one 2−cycle and two distinct points in $E(\mathbb{F}_q)$,

A_7 counts the number with one 2−cycle and one point in $E(\mathbb{F}_q)$ with multiplicity two,

A_8 counts the number with two distinct 2−cycles,

A_9 counts the number with one 2−cycle with multiplicity two,

A_{10} counts the number with one 3−cycle and one point in $E(\mathbb{F}_q)$, and

A_{11} counts the number with one 4−cycle.

Again, the definitions of A_1, A_9, and A_{11} immediately imply the result for $k = 4$. For $k = 5$, there are 17 kinds of curves with equation

$$aZ^2 + bXZ + cYZ + dX^2 + eXY = 0.$$

There are seven partitions of five, and the matrix of expansion coefficients in this case is

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 2 & 3 & 3 & 4 & 5 \\ 1 & 2 & 3 & 4 & 5 & 7 & 10 \\ 1 & 3 & 4 & 7 & 8 & 13 & 20 \\ 1 & 3 & 5 & 8 & 11 & 18 & 30 \\ 1 & 4 & 7 & 13 & 18 & 33 & 60 \\ 1 & 5 & 10 & 20 & 30 & 60 & 120 \\ 1 & 1 & 2 & 1 & 2 & 1 & 0 \\ 1 & 2 & 3 & 2 & 4 & 2 & 0 \\ 1 & 3 & 4 & 6 & 6 & 6 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 2 & 0 & 2 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 2 & 1 & 2 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}^T$$

After applying the signed coefficients c_λ's, we obtain $N_k/N_1 = A_1 + 5A_{17}$ which gives the right geometric interpretation. Note that precise definitions of A_1 through A_{17} omitted for this case but like the $k = 4$ case, A_1 counts the number where one point of $E(\mathbb{F}_q)$ has multiplicity 5, and A_{17} counts the number with one 5$-$cycle. To prove this result in general, we mention the following few approaches.

1) Based on the algebraic definition of H_k as the number of positive divisors, i.e. multi-cycles with k points, we can break up the sum $N_k = \sum_\lambda c_\lambda H_{\lambda_1} \cdots H_{\lambda_r}$ into more elementary structures so after summing the positive and negative terms together, we are left with an expression which is nonnegative and only includes a small subset of these elementary structures as terms. Since each H_k is divisible by N_1 there is no loss by dividing the entire expression by N_1 as long as the elementary structures are chosen in a way that they are all divisible by N_1.

2) We generalize the various cases (corresponding to elementary structures) as geometric configurations of points. Then we should be counting the number of curves with defining equation (on $Z = 1$ patch) given by $a_1 + a_2x + a_3y + a_4x^2 + a_5xy + a_6x^3 + a_7x^2y + \cdots + a_kM_k$ where monomial

$$M_k = \begin{cases} 1 \text{ if } k = 1 \\ x^{\frac{k}{2}} \text{ if } 2|k \\ x^{\frac{k-3}{2}} y \text{ if } 2 \not| k \text{ and } k \geq 3 \end{cases}.$$

Each of the terms in the expansion of N_k/N_1 according to partitions signifies a way of designating a subset of such curves, with some curves being designated multiple times with different data. Then an inclusion-exclusion argument or algebraic formula for such multiplicities should be able to prove that N_k/N_1 equals a nonnegative sum of a small subset of the terms with the right form.

We obtain general expressions $N_k = N_1 \cdot |V_k|$ where V_k equals the variety of points satisfying $P + \pi(P) + \cdots + \pi^{k-1}(P) \equiv kP_\infty$. This is called the trace-zero variety in the literature, e.g. [Fre01]. We provide the following explicit proof of this identity.

Proposition 4.56. *We have*

$$N_k/N_1 \quad = \quad \left| Ker\ (1 + \pi + \pi^2 + \cdots + \pi^{k-1}) \right|.$$

Proof. One can prove this result simply by observing

$$(1 - \pi^k) = (1 - \pi)(1 + \pi + \pi^2 + \cdots + \pi^{k-1})$$

and since these maps are group homomorphisms, we obtain

$$\begin{aligned} \left| \text{Ker } (1 - \pi^k) \right| &= \left| \text{Ker}(1 - \pi) \right| \cdot \left| Ker\ (1 + \pi + \pi^2 + \cdots + \pi^{k-1}) \right|, \qquad i.e \\ N_k &= N_1 \cdot \left| Ker\ (1 + \pi + \pi^2 + \cdots + \pi^{k-1}) \right|. \end{aligned}$$

□

In the literature, this is also commonly cited by appealing to Weil descent or Weil restriction. Because of the importance of this particular variety, we provide a second elementary proof of this equality.

Alternate proof of Corollary 4.56. Since $\pi(P_\infty) = P_\infty = \pi^{-1}(P_\infty)$, we have that any element P in the kernel of $Tr_k = 1 + \pi + \cdots + \pi^{k-1}$ must also satisfy

$$(1 + \pi + \cdots + \pi^{k-1})\pi(P) = (\pi + \pi^2 + \cdots + \pi^k)(P) = P_\infty.$$

Putting these two together, we get that such a P will satisfy $(1-\pi^k)(P) = P_\infty$. In particular, $P \in E(\mathbb{F}_{q^k})$, and we conclude Ker $Tr_k \subseteq E(\mathbb{F}_{q^k})$. On the other hand, if R is in the image of $1+\pi+\cdots+\pi^{k-1}$ acting on $Q \in E(\mathbb{F}_{q^k})$, then

$$(1+\pi+\cdots+\pi^{k-1})\pi(Q) = (\pi+\pi^2+\cdots+\pi^k)(Q) = (1+\pi+\cdots+\pi^{k-1})(Q),$$

hence $(1-\pi)R = P_\infty$, i.e. $R \in E(\mathbb{F}_q)$, and so Im $Tr_k \subseteq E(\mathbb{F}_q)$.

We wish to prove the following sequence

$$0 \longrightarrow Ker\ (1+\pi+\cdots+\pi^{k-1}) \longrightarrow E(\mathbb{F}_{q^k})\ \xrightarrow{1+\pi+\cdots+\pi^{k-1}} E(\mathbb{F}_q) \longrightarrow 0$$

is exact; which would imply

$$\frac{|E(\mathbb{F}_{q^k})|}{|E(\mathbb{F}_q)|} = \left| Ker\ (1+\pi+\pi^2+\cdots+\pi^{k-1}) \right|.$$

The only part we have left to prove is the fact that $Tr_k : E(\mathbb{F}_{q^k}) \to E(\mathbb{F}_q)$ is surjective. This can be verified by Hilbert's Theorem 90. [DF91].

Theorem 4.57 (Additive Version of Hilbert's Theorem 90)**.** *Let L/K be a finite cyclic Galois extension (of degree k) with $Gal(L/K) = \langle\sigma\rangle$. An element $y \in L$ satisfies*

$$\sum_{\tau \in Gal(L/K)} \tau(y) = \sum_{i=0}^{k-1} \sigma^i(y) = \phi_k(y) = 0$$

if and only if there exists $x \in L$ such that $y = x - \sigma(x)$.

By this Theorem, we rephrase the problem of finding the image of Tr_k as finding the kernel of operator $1-\pi$, which is $E(\mathbb{F}_q)$. However, we can also prove surjectivity by elementary means, as done in [GM, Ch. 1] for $\mathbb{F}_{q^k} \xrightarrow{Tr_k} \mathbb{F}_q$. We thus use this proof by considering how $\pi : E(\mathbb{F}_{q^k}) \to E(\mathbb{F}_{q^k})$ acts on each of two coordinates. By abuse of notation we now use π to denote the map from $\mathbb{F}_{q^k} \mapsto \mathbb{F}_{q^k}$ which sends α to α^q. Similarly Tr_k will be $1+\pi+\pi^2+\cdots+\pi^{k-1}$. The trace map is linear over $\mathbb{F}_q$, satisfying $Tr_k(c_1\alpha+c_2\beta) = c_1Tr_k(\alpha)+c_2Tr_k(\beta)$ for all $c_1, c_2 \in \mathbb{F}_q$ and $\alpha, \beta \in \mathbb{F}_{q^k}$. Also we have that for $\alpha \in \mathbb{F}_{q^k}$ the property

$$Tr_k(\alpha x) = 0 \text{ for all } x \in \mathbb{F}_{q^k} \text{ if and only if } \alpha = 0$$

since the equation $Tr_k(x) = 0$ is of degree q^{k-1} and thus cannot have more than q^{k-1} solutions in $\mathbb{F}_{q^k}$. Since $\mathbb{F}_{q^k}$ has q^k elements, we can certainly find $\alpha \in \mathbb{F}_{q^k}$ such that $Tr_k(\alpha) \neq 0$. Thus we let $Tr_k(\alpha) = c_1$ for $c_1 \in \mathbb{F}_q \setminus \{0\}$, and by using linearity of the trace map, we have $Tr_k(c_2\alpha/c_1) = c_2$ for all $c_2 \in \mathbb{F}_q$. Thus $Tr_k = 1 + \pi + \pi^2 + \cdots + \pi^{k-1}$ is surjective from $\mathbb{F}_{q^k}$ onto $\mathbb{F}_q$. $\square$

While the author has not worked out the details, this numeric identity should also give rise an explicit bijection for higher k via coset decomposition, as in the $k = 2$ case. Unfortunately, as seen even in that case, hope for a natural bijection is doubtful since the most natural type of bijection, a group isomorphism, cannot be constructed in general.

4.5 Acknowledgement

Much of the material in Chapter 4 has been submitted for publication in the paper "Combinatorial Aspects of Elliptic Curves" by Gregg Musiker. The dissertation author is the primary investigator and author of this paper.

5 Determinantal formulas for N_k

In subsection 4.1.1, we introduced the (q, t)-Lucas Numbers, which corresponded to $1 + q^k - N_k$ yet still helped produce a generating function for $-N_k$ directly in subsection 4.1.2. Similarly, we now illustrate a determinantal formula for N_k in terms of q and N_1 which at first glance looks analogous to the matrix of Proposition 4.28. The upshot to the revised determinantal formula is that the eigenvalues of matrix M_k, which are defined below, are factors of N_k, a statement that is not true for the matrix of Proposition 4.28.

Theorem 5.1. *Let* $M_1 = [-N_1]$, $M_2 = \begin{bmatrix} 1+q-N_1 & -1-q \\ -1-q & 1+q-N_1 \end{bmatrix}$, *and for* $k \geq 3$, *let* M_k *be the* k*-by-*k *"three-line" circulant matrix*

$$\begin{bmatrix} 1+q-N_1 & -1 & 0 & \dots & 0 & -q \\ -q & 1+q-N_1 & -1 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & -q & 1+q-N_1 & -1 & 0 \\ 0 & \dots & 0 & -q & 1+q-N_1 & -1 \\ -1 & 0 & \dots & 0 & -q & 1+q-N_1 \end{bmatrix}.$$

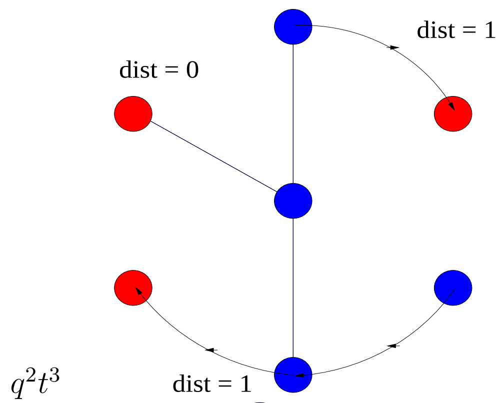
Then the sequence of integers $N_k = \#C(\mathbb{F}_{q^k})$ *satisfies the relation*

$$N_k = -\det M_k \quad \text{for all } k \geq 1.$$

We provide three proofs of this theorem, one which relies on graph theory, one which utilizes the three term recurrence from Section 4.1.1, and one which introduces a new sequence of polynomials which are interesting in their own right.

5.1 First proof of Theorem 5.1: Via graph theory

In subsection 4.1.3, we proved that N_k can be written as $-\mathcal{W}_k(q, -N_1)$ where $\mathcal{W}_k$ is a (q,t)-analogue of the number of spanning trees of W_k, where each tree is given a certain (q,t)-weighting. An alternative definition of $\mathcal{W}_k(q,t)$ uses a deformation of the wheel graph such that each edge incident to the central hub is replaced with t bi-directed edges, and every two adjacent vertices along the rim are connected via q edges going clockwise and 1 edge going counter-clockwise.



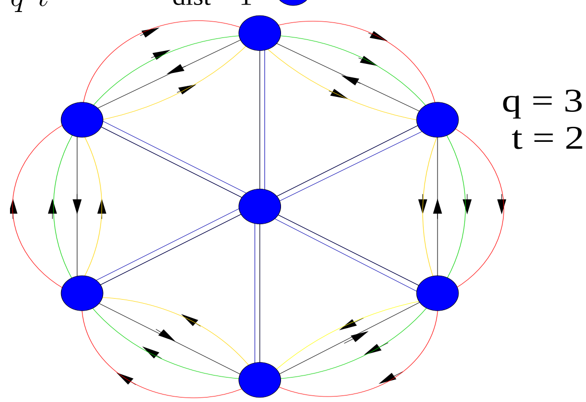


Figure 5.1: A second definition of $\mathcal{W}_k(q,t)$.

With this definition of the (q,t)-W_k, we no longer have to weight the spanning trees to obtain $\mathcal{W}_k(q,t)$; instead the (q,t)-weighting is implicit in the definition of the (q,t)-wheel graph. More precisely we obtain

Lemma 5.2. *$\mathcal{W}_k(q,t)$ as defined in Section 4.13 is equal to the (without weighting) number of directed rooted spanning trees of (q,t)-W_k which are rooted at the central hub.*

Having dispensed with the weightings, we can appeal to the directed multi-graph version of the Matrix-Tree Theorem to count (in the ordinary sense) the number of spanning trees of (q,t)-W_k with root v_0. Before describing this theorem, we provide some necessary terminology that will also be used again in Chapter 6. A directed multi-graph, as the name and picture implies, is a directed version of the simple graphs we earlier defined which also allow multiple edges between a given pair of vertices. We call the number of outgoing edges of a given vertex, the **outdegree**, and denote this quantity as $d(v_i)$. Additionally, we will let $d(v_i, v_j)$ denote the number of directed edges from v_i to v_j. The **Laplacian** matrix L of a graph is defined by entries $L_{ii} = d(v_i)$ and $L_{i,j} = -d(v_i, v_j)$. Finally we define a rooted spanning tree, with root v_0, to be an oriented spanning tree such that all edges flow away from v_0.

Theorem 5.3 (Matrix-Tree Theorem)**.** *The number of rooted spanning trees, with root v_0, of graph G is given as the determinant of the matrix L_0 where L_0 is the reduced Laplacian matrix, i.e. matrix L with the column and row corresponding to root v_0 removed.*

Proof. See [Sta99, Ch. 5]. $\square$

In the case of the (q,t)-wheel graph W_k, we obtain Laplacian matrix

$$L = \begin{bmatrix} 1+q+t & -1 & 0 & \dots & 0 & -q & -t \\ -q & 1+q+t & -1 & 0 & \dots & 0 & -t \\ \dots & \dots & \dots & \dots & \dots & \dots & -t \\ 0 & \dots & -q & 1+q+t & -1 & 0 & -t \\ 0 & \dots & 0 & -q & 1+q+t & -1 & -t \\ -1 & 0 & \dots & 0 & -q & 1+q+t & -t \\ -t & -t & -t & \dots & -t & -t & kt \end{bmatrix}$$

where the last row and column correspond to the hub vertex, which happens to be the root. By the Matrix-Tree theorem, the number of directed rooted spanning trees is $\det L_0$ where L_0 is matrix L with the last row and last column deleted. We have the identities

$$N_k = -\mathcal{W}_k(q, -N_1) \tag{5.1}$$

$$M_k = L_0\Big|_{t=-N_1} \quad \text{and thus} \tag{5.2}$$

$$\mathcal{W}_k(q, t) = \det L_0 \quad \text{implies} \tag{5.3}$$

$$-\mathcal{W}_k(q, -N_1) = -\det L_0\Big|_{t=-N_1} \quad \text{so we get} \tag{5.4}$$

$$N_k = -\det M_k. \tag{5.5}$$

Thus we have proven Theorem 5.1.

5.1.1 The Smith normal form of matrices M_k

Before discussing the other proofs of Theorem 5.1, and related topics, we stop to discuss a combinatorially interesting feature of these matrices. As we have written the M_k's, they are sparse circulant matrices with very simple entries. However, the Smith normal forms of these matrices are also quite nice. Recall that the Smith normal form of an integral matrix is unchanged by

1. Multiplication of a row or a column by -1.

2. Addition of an integer multiple of a row or column to another.

3. Swapping of two rows or two columns.

In particular, the determinant of the matrix is unchanged by these operations. To be precise a matrix has a Smith normal form when its entries are defined over a principal ideal domain R such as $\mathbb{Z}$ or $F[x]$ where F is a field. In general, operation (1) would be expressed as "multiplication of a row or a column by a unit in R," however when $R = \mathbb{Z}$ the only units are ± 1. The matrices we consider have entries which are integral polynomials in the constants q and N_1 (or t). Thus to obtain the Smith normal form, we must fix q and N_1 (resp. t) to be specific integers before proceeding. Nonetheless, even with this caveat, we will be able to provide a combinatorial description of the Smith normal form of our matrices.

Theorem 5.4. *The Smith normal form of M_k is equivalent to*

$$\begin{bmatrix} 1 & 0 & \dots & 0 & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 & 0 & 0 \\ 0 & 0 & \dots & 0 & qE_{k-1}/N_1 - 1 & -qE_k/N_1 \\ 0 & 0 & \dots & 0 & E_k/N_1 & -E_{k+1}/N_1 - 1 \end{bmatrix}$$

where the E_k's are the signed bivariate Fibonacci polynomials from subsection 4.2. Note that the lower-right 2-by-2 block will reduce to $\begin{bmatrix} m_1 & 0 \\ 0 & m_2 \end{bmatrix}$ such that $m_1|m_2$ as integers once q and N_1 are evaluated as specific numbers.

Before proving this theorem, we provide the following Lemma that will be a key step in our proof. This Lemma describes a matrix identity which is an immediate corollary to Proposition 4.20.

Lemma 5.5.

$$\begin{bmatrix} 0 & -q \\ 1 & 1+q-N_1 \end{bmatrix}^n = \begin{bmatrix} q \cdot (-1)^{n-1} E_{n-1}/N_1 & q \cdot (-1)^n E_n/N_1 \\ (-1)^{n-1} E_n/N_1 & (-1)^n E_{n+1}/N_1 \end{bmatrix}$$

for all $n \geq 2$.

Proof. We prove this by induction on n. The initial conditions

$$\begin{bmatrix} 0 & -q \\ 1 & 1+q-N_1 \end{bmatrix}^2 = \begin{bmatrix} -q & -q(1+q-N_1) \\ 1+q-N_1 & (1+q-N_1)^2 - q \end{bmatrix} = \begin{bmatrix} -q \cdot E_1/N_1 & q \cdot E_2/N_1 \\ -E_2/N_1 & E_3/N_1 \end{bmatrix}$$

$$\begin{bmatrix} 0 & -q \\ 1 & 1+q-N_1 \end{bmatrix}^3 = \begin{bmatrix} -q(1+q-N_1) & q^2 - q(1+q-N_1)^2 \\ -q + (1+q-N_1)^2 & -E_4/N_1 \end{bmatrix} = \begin{bmatrix} q \cdot E_2/N_1 & -q \cdot E_3/N_1 \\ E_2/N_1 & -E_4/N_1 \end{bmatrix}$$

are clear. Furthermore,

$$\begin{aligned} \begin{bmatrix} 0 & -q \\ 1 & 1+q-N_1 \end{bmatrix} & \times \begin{bmatrix} q \cdot (-1)^{n-1} E_{n-1}/N_1 & q \cdot (-1)^n E_n/N_1 \\ (-1)^{n-1} E_n/N_1 & (-1)^n E_{n+1}/N_1 \end{bmatrix} \\ & = \begin{bmatrix} q \cdot (-1)^n E_n/N_1 & q \cdot (-1)^{n+1} E_{n+1}/N_1 \\ a_2 & b_2 \end{bmatrix} \end{aligned}$$

where

$$\begin{aligned} a_2 &= q \cdot (-1)^{n-1} E_{n-1}/N_1 - (1+q-N_1) \cdot (-1)^n E_n/N_1 \quad \text{and} \\ b_2 &= q \cdot (-1)^n E_n/N_1 - (1+q-N_1) \cdot (-1)^{n+1} E_{n+1}/N_1. \end{aligned}$$

Thus the inductive step, i.e. $a_2 = (-1)^n E_{n+1}/N_1$ and $b_2 = (-1)^{n+1} E_{n+2}/N_1$, follows from the recursion of Proposition 4.20. □

Proof of Theorem 5.4. To begin we note after permuting rows cyclically and multiplying through all rows by (-1) that we get

$$M_k \equiv \begin{bmatrix} 1 & 0 & \dots & 0 & q & -1-q+N \\ -1-q-N & 1 & 0 & \dots & 0 & q \\ q & -1-q-N & 1 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & 0 & q & -1-q+N & 1 & 0 \\ 0 & \dots & 0 & q & -1-q+N & 1 \end{bmatrix}.$$

Since this matrix is lower-triangular with ones on the diagonal, besides the upper-right corner of three elements, we can add a multiple of the first row to the second and third rows, respectively, and obtain a new matrix with vector

$$V = [1, 0, 0, \dots, 0]^T$$

as the first column. Since we can add multiples of columns to one another as well, we also obtain a matrix with vector V^T as the first row.

This new matrix will again be lower triangular with ones along the diagonal, except for nonzero entries in four spots in the last two columns of rows two and three. By the symmetry and sparseness of this matrix, we can continue this process, which will always shift the nonzero block of four in the last two columns down one row. This process will terminate with a block diagonal matrix consisting of $(k-2)$ 1-by-1 blocks of element 1 followed by a single 2-by-2 block which will be more complicated. To explicitly identity these elements, we consider the following recursive argument.

Let $\begin{bmatrix} a_1'' & b_1'' \\ a_2' & b_2' \\ a_3 & b_3 \\ a_4 & b_4 \\ a_5 & b_5 \\ \vdots & \vdots \\ a_k & b_k \end{bmatrix}$ signify the last two columns of matrix M_k. Following the above construction, we obtain $\begin{bmatrix} 0 & 0 \\ a_2'' & b_2'' \\ a_3' & b_3' \\ a_4 & b_4 \\ a_5 & b_5 \\ \vdots & \vdots \\ a_k & b_k \end{bmatrix}$ after one iteration, and $\begin{bmatrix} 0 & 0 \\ 0 & 0 \\ a_3'' & b_3'' \\ a_4' & b_4' \\ a_5 & b_5 \\ \vdots & \vdots \\ a_k & b_k \end{bmatrix}$ after the next, where

$$\begin{aligned} a_i'' &= (1+q-N_1)a_{i-1}'' + a_i' \\ b_i'' &= (1+q-N_1)b_{i-1}'' + b_i' \\ a_{i+1}' &= -qa_{i-1}'' + a_{i+1} \\ b_{i+1}' &= -qb_{i-1}'' + b_{i+1} \end{aligned}$$

for $2 \leq i \leq k-1$. Consequently,

$$\begin{bmatrix} a_m'' \\ a_{m+1}'' \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -q & 1+q-N_1 \end{bmatrix} \begin{bmatrix} a_{m-1}'' \\ a_m'' \end{bmatrix} + \begin{bmatrix} 0 \\ a_{m+1} \end{bmatrix} \quad \text{and} \tag{5.6}$$

$$\begin{bmatrix} b_m'' \\ b_{m+1}'' \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -q & 1+q-N_1 \end{bmatrix} \begin{bmatrix} b_{m-1}'' \\ b_m'' \end{bmatrix} + \begin{bmatrix} 0 \\ b_{m+1} \end{bmatrix}. \tag{5.7}$$

Since we have $a_1'' = q,\ b_1'' = -1-q+N_1,\ b_2' = q,\ a_{k-1} = 1,\ a_k = -1-q+N_1,$

$b_k = 1$, and the rest of the a_i and b_i equal 0, we obtain

$$\begin{aligned}
\begin{bmatrix} a''_{k-2} \\ a''_{k-1} \end{bmatrix} &= \begin{bmatrix} 0 & 1 \\ -q & 1+q-N_1 \end{bmatrix}^{k-3} \begin{bmatrix} a''_1 \\ a''_2 \end{bmatrix} + \begin{bmatrix} 0 \\ a_{k-1} \end{bmatrix} \\
&= \begin{bmatrix} 0 & 1 \\ -q & 1+q-N_1 \end{bmatrix}^{k-3} \begin{bmatrix} q \\ q(1+q-N_1) \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix}.
\end{aligned}$$

Analogously,

$$\begin{aligned}
\begin{bmatrix} b''_{k-2} \\ b''_{k-1} \end{bmatrix} &= \begin{bmatrix} 0 & 1 \\ -q & 1+q-N_1 \end{bmatrix}^{k-3} \begin{bmatrix} b''_1 \\ b''_2 \end{bmatrix} + \begin{bmatrix} 0 \\ b_{k-1} \end{bmatrix} \\
&= \begin{bmatrix} 0 & 1 \\ -q & 1+q-N_1 \end{bmatrix}^{k-3} \begin{bmatrix} -1-q+N_1 \\ q-(1+q-N_1)^2 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \end{bmatrix}.
\end{aligned}$$

Putting this together we get

$$\begin{bmatrix} a''_{k-2} & b''_{k-2} \\ a''_{k-1} & b''_{k-1} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -q & 1+q-N_1 \end{bmatrix}^{k-3} \begin{bmatrix} q & -1-q+N_1 \\ -1-q+N_1 & q-(1+q-N_1)^2 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$

which simplifies to

$$\begin{bmatrix} a''_{k-2} & b''_{k-2} \\ a''_{k-1} & b''_{k-1} \end{bmatrix} = (-1) \begin{bmatrix} 0 & 1 \\ -q & 1+q-N_1 \end{bmatrix}^{k-1} + \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}.$$

Finally we get

$$\begin{aligned}
\begin{bmatrix} a''_{k-1} & b''_{k-1} \\ a''_{k} & b''_{k} \end{bmatrix} &= (-1) \begin{bmatrix} 0 & 1 \\ -q & 1+q-N_1 \end{bmatrix}^{k} + \begin{bmatrix} 0 & 1 \\ -q & 1+q-N_1 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} + \begin{bmatrix} 0 & a_k \\ 0 & b_k \end{bmatrix} \\
&= (-1) \begin{bmatrix} 0 & 1 \\ -q & 1+q-N_1 \end{bmatrix}^{k} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.
\end{aligned}$$

At this point we recall Lemma 5.5 which states

$$\begin{bmatrix} 0 & -q \\ 1 & 1+q-N_1 \end{bmatrix}^{k} = \begin{bmatrix} q \cdot (-1)^{k-1} E_{k-1}/N_1 & q \cdot (-1)^k E_k/N_1 \\ (-1)^{k-1} E_k/N_1 & (-1)^k E_{k+1}/N_1 \end{bmatrix}$$

for all $k \geq 2$. To finish the proof we multiply the last two rows by a power of (-1) and take the transpose, neither of which effects the Smith normal form. $\square$

Besides showing another connection between the Fibonacci numbers and the N_k's, this theorem will be used again in Chapter 6.

5.2 Second proof of Theorem 5.1: Using orthogonal polynomials

Recall from the zeta function of an elliptic curve, $Z(E,T)$, we derived a three term recurrence relation for the sequence $\{G_k = 1 + q^k - N_k\}$:

$$G_{k+1} = (1 + q - N_1)G_k - qG_{k-1}. \tag{5.8}$$

Such a relation is indicative of an interpretation of the $1 + q^k - N_k$'s as a sequence of orthogonal polynomials. In particular, any sequence of orthogonal polynomials, $\{P_k(x)\}$, satisfies

$$P_{k+1}(x) = (a_k x + b_k)P_k(x) + c_k P_{k-1}(x) \tag{5.9}$$

where a_k, b_k and c_k are constants that depend on $k \in \mathbb{N}$. Additionally, it is usual to initialize $P_{-k}(x) = 0$, $P_0(x) = 1$, and $P_1(x) = a_0 x + b_0$.

Since we can think of the bivariate $N_k(q, N_1)$ as univariate polynomials in variable N_1 with constants from field $\mathbb{Q}(q)$, it follows that recurrence (5.8) is such an example, with

$$\begin{array}{rcll} a_k & = & -1 & \text{for } k \geq 0 \\ b_k & = & 1+q & \text{for } k \geq 0, \\ c_1 & = & -2q & \quad\text{and} \\ c_k & = & -q & \text{for } k \geq 2 \end{array}$$

in the case. (Note that we must take c_1 to be $2q$ because we originally defined $L_0(q,t)$ as 2.) One of the properties of a sequence of orthogonal polynomials is an interpretation as the determinants of a family of tridiagonal k-by$k$ matrices. In particular, we obtain a second proof of Proposition 4.28.

Proof. Given a sequence of orthogonal polynomials satisfying $P_0(x) = 1$, $P_1(x) =$

a_0x+b_0 and recurrence (5.9), we have the formula [IPS00]

$$P_k(x) = \det \begin{bmatrix} a_0x+b_0 & c_1 & 0 & 0 & 0 & 0 \\ -1 & a_1x+b_1 & c_2 & 0 & 0 & 0 \\ 0 & -1 & a_2x+b_2 & c_3 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & \cdots & a_{k-2}x+b_{k-2} & c_k \\ 0 & 0 & 0 & \cdots & -1 & a_{k-1}x+b_{k-2} \end{bmatrix}.$$

Plugging in the a_i, b_i, and c_i's as above yields the formula. □

Recall that we obtained these same formulas, i.e. determinants of matrices M'_k in Section 4.2. We can prove Theorem 5.1 by an algebraic manipulation of matrix M_k followed by use of Proposition 4.28. Namely, by using the multilinearity of the determinant, and expansions about the first row followed by the first column, we obtain

$$\det(M_k) = \det(A_k) + \det(B_k) + \det(C_k) + \det(D_k)$$

where A_k, B_k, C_k, and D_k are the following k-by-k matrices:

$$A_k = \begin{bmatrix} 1+q-N_1 & -1 & 0 & 0 & 0 & 0 \\ -q & 1+q-N_1 & -1 & 0 & 0 & 0 \\ 0 & -q & 1+q-N_1 & -1 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & \cdots & 1+q-N_1 & -1 \\ 0 & 0 & 0 & \cdots & -q & 1+q-N_1 \end{bmatrix}.$$

$$B_k = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & -q \\ -q & 1+q-N_1 & -1 & 0 & 0 & 0 \\ 0 & -q & 1+q-N_1 & -1 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & \cdots & 1+q-N_1 & -1 \\ 0 & 0 & 0 & \cdots & -q & 1+q-N_1 \end{bmatrix}.$$

$$C_k = \begin{bmatrix} 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 1+q-N_1 & -1 & 0 & 0 & 0 \\ 0 & -q & 1+q-N_1 & -1 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & \cdots & 1+q-N_1 & -1 \\ -1 & 0 & 0 & \cdots & -q & 1+q-N_1 \end{bmatrix}.$$

$$D_k = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & -q \\ 0 & 1+q-N_1 & -1 & 0 & 0 & 0 \\ 0 & -q & 1+q-N_1 & -1 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & \cdots & 1+q-N_1 & -1 \\ -1 & 0 & 0 & \cdots & -q & 1+q-N_1 \end{bmatrix}.$$

Cyclic permutation of the rows of B_k and the columns of C_k yield upper-triangular matrices with -1's (resp. $-q$)'s on the diagonal. Given that the sign of such a cyclic permutation is $(-1)^{k-1}$, we obtain $\det(B_k)+\det(C_k) = -q-1$. Additionally, by expanding $\det(D_k)$ about the first row followed by the first column, we obtain $\det(D_k) = -q\det(A_{k-2})$. In conclusion

$$1 + q^k + \det(M_k) = \det(A_k) - q\det(A_{k-2}).$$

By analogous methods we obtain

$$\det M_k' = \det(A_k) - q\det(A_{k-2})$$

and thus the desired formula $\det M_k = -N_k$.

5.2.1 Explicit connection to orthogonal polynomials

We now push the analysis of the last section further, writing the $\{1+q^k-N_k\}$'s explicitly in terms of a sequence of classical orthogonal polynomials. We let $T_k(x)$ denote the kth Chebyshev (Tchebyshev) polynomials of the first kind, which are defined as $\cos(k\theta)$ written out in terms of x such that $\theta = \arccos x$. Equivalently, we can define $T_k(x)$ as the expansion of $\alpha^k + \beta^k$ in terms of powers of $\cos\theta$ where

$$\begin{aligned} \alpha &= \cos\theta + i\sin\theta \\ \beta &= \cos\theta - i\sin\theta. \end{aligned}$$

Theorem 5.6. *Considering the $(1+q^k-N_k)$'s as univariate polynomials in N_1 over the field $\mathbb{Q}(q)$, we obtain*

$$1 + q^k - N_k = 2q^{k/2} T_k\Big((1+q-N_1)/2q^{1/2}\Big).$$

Proof. We note that Chebyshev polynomials satisfy initial conditions $T_0(x) = 1$, and $T_1(x) = x$ and the three-term recurrence

$$T_{k+1}(x) = 2xT_k(x) - T_{k-1}(x)$$

for $k \geq 1$ since

$$\begin{aligned} T_{k+1}(x) &= \alpha^{k+1} + \beta^{k+1} \\ &= (\alpha+\beta)(\alpha^k+\beta^k) - \alpha\beta(\alpha^{k-1}+\beta^{k-1}) \\ &= 2\cos\theta\, T_k(x) - T_{k-1}(x) \\ &= 2xT_k(x) - T_{k-1}(x). \end{aligned}$$

Let $x = \frac{1+q-N_1}{2\sqrt{q}}$. Clearly Theorem 5.6 holds for $k = 1$, and additionally the $\frac{1+q^k-N_k}{2q^{k/2}}$'s satisfy the same recurrence as the $T_k(x)$'s. Namely

$$\begin{aligned} \frac{1+q^{k+1}-N_{k+1}}{2q^{(k+1)/2}} &= \frac{(1+q-N_1)(1+q^k-N_k) - q(1+q^{k-1}-N_{k-1})}{2q^{(k+1)/2}} \\ &= 2\left(\frac{1+q-N_1}{2q^{1/2}}\right)\left(\frac{1+q^k-N_k}{2q^{k/2}}\right) - \left(\frac{1+q^{k-1}-N_{k-1}}{2q^{(k-1)/2}}\right). \end{aligned}$$

$\square$

Another way to foresee the appearance of Chebyshev polynomials is by noting that in the case that we plug in $q = 0$ or $q = 1$, we obtain a family of univariate polynomials $\tilde{N}_k$ with the property $\tilde{N}_{mk} = \tilde{N}_m(\tilde{N}_k) = \tilde{N}_k(\tilde{N}_m)$. It is a fundamental theorem of Chebyshev polynomials that families of univariate polynomials with such a property are very restrictive. In particular, from [BT51] as described on page 33 of [BE95]: If $\{\tilde{N}_k\}$ is a sequence of integral univariate polynomials of degree k with the property

$$\tilde{N}_{mn} = \tilde{N}_m(\tilde{N}_n) = \tilde{N}_n(\tilde{N}_m)$$

for all positive integers m and n, then $\tilde{N}_k$ must either be a linear transformation of

1. x^k or

2. $T_k(x)$, the Chebyshev polynomial of the first kind,

where a linear transformation of a polynomial $f(x)$ is of the form

$$A \cdot f\Big((x-B)/A\Big) + B \text{ or equivalently } \Big(f(\overline{A}x + \overline{B}) - \overline{B}\Big)\Big/\overline{A}.$$

In particular we get formulas for $\mathcal{W}_k(0, N_1)$ and $\mathcal{W}_k(1, N_1)$ (resp. $N_k(0, N_1)$ and $N_k(1, N_1)$) which are indeed linear transformations of x^k and $T_k(x)$ respectively.

Proposition 5.7.

$$\begin{aligned} N_k(0, N_1) &= -(1 - N_1)^k + 1, && (5.10) \\ N_k(1, N_1) &= -2T_k(-N_1/2 + 1) + 2. && (5.11) \end{aligned}$$

Proof. The coefficient of N_1^m in $\mathcal{W}_k(0, N_1)$ is the number of directed spanning trees of W_k with m spokes and arcs always directed counter-clockwise. In particular it is only the placement of the spokes that matter at this point since the placement of the arcs is now forced. Thus the coefficient of N_1^m in $\mathcal{W}_k(0, N_1)$ is $\binom{k}{m}$ for all $1 \leq m \leq k$. Thus the generating function $\mathcal{W}_k(0, N_1)$ satisfies

$$\mathcal{W}_k(0, N_1) = (1 + N_1)^k - 1$$

since the constant term of $\mathcal{W}_k(0, N_1)$ is zero. Using the relation $N_k(q, N_1) = -\mathcal{W}_k(q, -N_1)$ completes the proof in the $q = 0$ case. We also note that $-(1-x)^k + 1$ is a linear transformation of x^k via $A = -1$ and $B = 1$. The case for $q = 1$ is a corollary of Theorem 5.6. ∎

For higher values of q, we lose some of the symmetry and thus cannot apply the Fundamental Theorem of Chebyshev polynomials. However, it seems fruitful to consider the theory of Chebyshev polynomials when considering alternate polynomial expressions or expansions of $\mathcal{W}_k(q, N_1)$. For example, putting together Proposition 5.7 with a result of [ZYG05], namely Theorem 12, we get the following result.

Theorem 5.8. *For $n = N_1 \geq k \geq 3$, let $T(K_n - C_k)$ signify the number of spanning trees in the graph $K_n - C_k$ formed by taking the complete graph on n vertices and removing the k edges of a k-cycle. Then we have as a formal expression*

$$T(K_n - C_k) = (-1)^{k-1} n^{n-k-2} N_k(1, n).$$

Proof. In [ZYG05], the authors develop a formula in terms of Chebyshev polynomials for the number of spanning trees of various graphs. In particular, they find that

$$T(K_n - C_k) = n^{n-k-2}\left[\left(\sqrt{\frac{n}{4}} + \sqrt{\frac{n-4}{4}}\right)^k - \left(-\sqrt{\frac{n}{4}} + \sqrt{\frac{n-4}{4}}\right)^k\right]^2$$

which after several steps of algebra is found to be equal to

$$n^{n-k-2}(-1)^k(2T_k(-n/2+1) - 2).$$

More specifically, we use relation

$$T_k(x) = \frac{1}{2}\left[\left(x + \sqrt{x^2-1}\right)^k + \left(x - \sqrt{x^2-1}\right)^k\right]$$

from Equation (19) of [BP86]. Plugging in $x = -n/2 + 1$, we get

$$\begin{aligned}(-1)^k\Big(2T_k(-n/2+1) - 2\Big) &= \left(n/2 - 1 - \sqrt{\frac{n(n-4)}{4}}\right)^k \\ &+ \left(n/2 - 1 + \sqrt{\frac{n(n-4)}{4}}\right)^k + 2(-1)^{k-1}.\end{aligned}$$

On the other hand, expanding

$$\begin{aligned}\left[\left(\sqrt{\frac{n}{4}} + \sqrt{\frac{n-4}{4}}\right)^k - \left(-\sqrt{\frac{n}{4}} + \sqrt{\frac{n-4}{4}}\right)^k\right]^2 &= \left(\sqrt{\frac{n}{4}} + \sqrt{\frac{n-4}{4}}\right)^{2k} + \left(-\sqrt{\frac{n}{4}} + \sqrt{\frac{n-4}{4}}\right)^{2k} \\ &- 2\left(\sqrt{\frac{n}{4}} + \sqrt{\frac{n-4}{4}}\right)^k\left(-\sqrt{\frac{n}{4}} + \sqrt{\frac{n-4}{4}}\right)^k,\end{aligned}$$

we obtain

$$-2\left(\sqrt{\frac{n}{4}}+\sqrt{\frac{n-4}{4}}\right)^k\left(-\sqrt{\frac{n}{4}}+\sqrt{\frac{n-4}{4}}\right)^k = -2\left(\frac{n-4}{4}-\frac{n}{4}\right)^k = 2(-1)^{k-1}$$

and

$$\left(\sqrt{\frac{n}{4}}+\sqrt{\frac{n-4}{4}}\right)^{2k} = \left(\frac{n}{4}+2\sqrt{\frac{n(n-4)}{16}}+\frac{n-4}{4}\right)^k = \left(n/2-1+\sqrt{\frac{n(n-4)}{4}}\right)^k.$$

Analogously

$$\left(-\sqrt{\frac{n}{4}}+\sqrt{\frac{n-4}{4}}\right)^{2k} = \left(\frac{n}{4}-2\sqrt{\frac{n(n-4)}{16}}+\frac{n-4}{4}\right)^k = \left(n/2-1-\sqrt{\frac{n(n-4)}{4}}\right)^k.$$

We thus have $T(K_n - C_k) = n^{n-k-2}(-1)^k\Big(2T_k(-N_1/2+1)-2\Big)$ which equals $n^{n-k-2}(-1)^k\Big(-N_k(1,n)\Big)$ by Proposition 5.7. $\square$

5.3 Third proof of Theorem 5.1: Using the zeta function

Alternatively, we note that we can factor

$$N_k = 1 + q^k - \alpha_1^k - \alpha_2^k$$

using the fact that $q = \alpha_1\alpha_2$. Consequently,

$$N_k = (1-\alpha_1^k)(1-\alpha_2^k)$$

and we can factor each of these two terms using cyclotomic polynomials. We recall that $(1-x^k)$ factors as

$$1 - x^k = \prod_{d|k} Cyc_d(x)$$

where $Cyc_d(x)$ is a monic irreducible polynomial with integer coefficients. We can similarly factor N_k as

$$N_k = \prod_{d|k} Cyc_d(\alpha_1)Cyc_d(\alpha_2).$$

These factors are therefore bivariate analogues of the cyclotomic polynomials, and we will refer to them henceforth as **elliptic cyclotomic polynomials**, denoted as $ECyc_d$.

Definition 5.9. We define the elliptic cyclotomic polynomials to be a sequence of polynomials in variables q and N_1 such that for $d \geq 1$,

$$ECyc_d = Cyc_d(\alpha_1)Cyc_d(\alpha_2),$$

where α_1 and α_2 are the two roots of

$$T^2 - (1 + q - N_1)T + q.$$

We verify that they can be expressed in terms of q and N_1 by the following proposition.

Proposition 5.10. *Writing down $ECyc_d$ in terms of q and N_1 yields irreducible bivariate polynomials with integer coefficients.*

Proof. Firstly we have

$$\alpha_1^j + \alpha_2^j = (1 + q^j - N_j) \in \mathbb{Z}$$

for all $j \geq 1$ and expanding a polynomial in α_1 multiplied by the same polynomial in α_2 yields terms of the form $\alpha_1^i\alpha_2^i(\alpha_1^j + \alpha_2^j)$. Secondly the quantity N_j is an integral polynomial in terms of q and N_1 by Theorem 4.1 and $\alpha_1^i\alpha_2^i = q^i$. Putting these relations together, and the fact that Cyc_d is an integral polynomial itself, we obtain the desired expressions for $ECyc_d$.

Now let us assume that $ECyc_d$ is factored as $F(q, N_1)G(q, N_1)$. The polynomial $Cyc_d(x)$ factors over the complex numbers as

$$Cyc_d(x) = \prod_{\substack{j=1 \\ \gcd(j,d)=1}}^{d} (1 - \omega^j x)$$

where ω is a dth root of unity. Thus $F(q, N_1) = \prod_{i \in S}(1-\omega^i\alpha_1)\prod_{j \in T}(1-\omega^j\alpha_2)$ for some nonempty subsets S, T of elements relatively prime to d. The only way F can be integral is if F equals its complex conjugate $\overline{F}$. However, α_1 and α_2 are complex conjugates by the Riemann hypothesis for elliptic curves [Has34, Sil92] (Hasse's Theorem), and thus $F = \overline{F}$ implies that the sets S and T are equal. Since $Cyc_d(x)$ is known to be irreducible, the only possibility is $S = T = \{j : \gcd(j, d) = 1\}$, and thus $F(q, N_1) = ECyc_d$, $G(q, N_1) = 1$. □

Remark 5.11. Alternatively, the integrality of the $ECyc_d$'s follows from the Fundamental Theorem of Symmetric Functions that states that a symmetric polynomial with integer coefficients can be rewritten as an integral polynomial in $e_1, e_2, \ldots$. In this case, $Cyc_d(\alpha_1)Cyc_d(\alpha_2)$ is a symmetric polynomial in two variables so $e_1 = \alpha_1 + \alpha_2 = 1 + q - N_1$, $e_2 = \alpha_1\alpha_2 = q$, and $e_k = 0$ for all $k \geq 3$. Thus we obtain an expression for $ECyc_d$ as a polynomial in q and N_1 with integer coefficients.

We can factor N_k, i.e. the $ECyc_d$'s even further, if we no longer require our expressions to be integral.

$$\begin{aligned} N_k &= \prod_{j=1}^{k}(1 - \alpha_1\omega_k^j)(1 - \alpha_2\omega_k^j) \\ &= \prod_{j=1}^{k}(1 - (\alpha_1 + \alpha_2)\omega_k^j + (\alpha_1\alpha_2)\omega_k^{2j}) \\ &= (-1)\prod_{j=1}^{k}(-\omega_k^{k-j})(1 - (1 + q - N_1)\omega_k^j + (q)\omega_k^{2j}) \\ &= -\prod_{j=1}^{k}\left((1 + q - N_1) - q\omega_k^j - \omega_k^{k-j}\right). \end{aligned}$$

Furthermore, the eigenvalues of a circulant matrix are well-known, and involve roots of unity analogous to the expression precisely given by the second equation above. (For example Loehr, Warrington, and Wilf [LWW04] provide an analysis of a more general family of three-line-circulant matrices from a combinatorial perspective. Using their notation, our result can be stated as

$$N_k = \Phi_{k,2}(1 + q - N_1, -q)$$

where $\Phi_{p,q}(x, y) = \prod_{j=1}^{p}(1 - x\omega^j - y\omega^{qj})$ and ω is a primitive pth root of unity. It is unclear how our combinatorial interpretation of N_k, in terms of spanning trees, relates to theirs, which involves permutation enumeration.) In particular, we prove Theorem 5.1 since $\det M_k$ equals the product of M_k's eigenvalues, which are precisely given as the k factors of $-N_k$ in second equation above.

5.3.1 Combinatorics of elliptic cyclotomic polynomials

In this subsection we further explore properties of elliptic cyclotomic polynomials, noting that they are more than auxiliary expressions that appear in the derivation of a proof. To start with, by Möbius inversion, we can use the identity

$$N_k = \prod_{d|k} ECyc_d(q, N_1) \tag{5.12}$$

to define elliptic cyclotomic polynomials directly as

$$ECyc_k(q, N_1) = \prod_{d|k} {N_d}^{\mu(k/d)} \tag{5.13}$$

in addition to the alternative definition

$$ECyc_k(q, N_1) = \prod_{\substack{j=1 \\ \gcd(j,d)=1}}^{k} \Big((1 + q - N_1) - q\omega_k^j - \omega_k^{k-j}\Big). \tag{5.14}$$

In particular, $ECyc_1 = N_1$ and $ECyc_p = N_p/N_1$ if p is prime. We note several commonalities among these polynomials, as described in the following propositions. These properties are further rationale for our choice of name for this family of polynomials.

Proposition 5.12. *We have*

$$ECyc_d|_{N_1=0} = C(d)Cyc_d(q) \tag{5.15}$$

$$ECyc_d|_{N_1=2q+2} = C'(d)Cyc_d(-q) \tag{5.16}$$

where $C(d)$ and $C'(d)$ are the functions from $\mathbb{Z}_{>0}$ to $\mathbb{Z}_{\geq 0}$ such that

$$C(d) = \begin{cases} 0 \text{ if } d = 1 \\ p \text{ if } d = p^k \text{ for } p \text{ prime} \\ 1 \text{ otherwise} \end{cases}$$

Table 5.1: Elliptic cyclotomic polynomials $ECyc_k(q, N_1)$ for small k.

$$\begin{aligned}
ECyc_4 &= N_1^2 - (2+2q)N_1 + 2(1+q^2)\\
ECyc_6 &= N_1^2 - (1+q)N_1 + (1-q+q^2)\\
ECyc_8 &= N_1^4 - (4+4q)N_1^3 + (6+8q+6q^2)N_1^2 - (4+4q+4q^2+4q^3)N_1\\
&+ 2(1+q^4)\\
ECyc_9 &= N_1^6 - (6+6q)N_1^5 + (15+24q+15q^2)N_1^4 - (21+36q+36q^2+21q^3)N_1^3\\
&+ (18+27q+27q^2+27q^3+18q^4)N_1^2 - (9+9q+9q^2+9q^3+9q^4+9q^5)N_1\\
&+ 3(1+q^3+q^6)\\
ECyc_{10} &= N_1^4 - (3+3q)N_1^3 + (4+3q+4q^2)N_1^2 - (2+q+q^2+2q^3)N_1\\
&+ (1-q+q^2-q^3+q^4)\\
ECyc_{12} &= N_1^4 - (4+4q)N_1^3 + (5+8q+5q^2)N_1^2 - (2+2q+2q^2+2q^3)N_1\\
&+ (1-q^2+q^4)
\end{aligned}$$

and

$$C'(d) = \begin{cases} -2 \text{ if } d = 1 \\ 0 \text{ if } d = 2 \\ p \text{ if } d = 2p^k \text{ for } p \text{ prime (including 2)} \\ 1 \text{ otherwise} \end{cases}.$$

Proof. In the case that $N_1 = 0$, the characteristic quadratic equation factors as

$$1 - (1 + q - N_1)T + qT^2 = (1 - T)(1 - qT).$$

Consequently, $\alpha_1 = 1$ and $\alpha_2 = q$ in this special case. (Note this is strictly formal since $N_1 = 0$ is impossible, and thus it is not contradictory that the Riemann Hypothesis fails.) Nonetheless, we still have $ECyc_d = Cyc_d(\alpha_1)Cyc_d(\alpha_2)$, and consequently,

$$ECyc_d|_{N_1=0} = Cyc_d(1)Cyc_d(q).$$

Finally the value of $Cyc_d(1)$ equals the function defined as $C(d)$ above [Slo, Seq. A020500].

For the reader's convenience we also provide a simple proof of this equality. It is clear that $Cyc_1(q) = 1-q$ and $Cyc_p(q) = 1+q+q^2+\cdots+q^{p-1}$ so by induction on $k \geq 1$, assume that $Cyc_{p^k}(1) = p$.

$$\frac{1-q^{p^k}}{1-q} = 1+q+q^2+\cdots+q^{p^k-1} = \prod_{j=1}^{k} Cyc_{p^j}(q).$$

Plugging in $q = 1$, and by induction we get $p^k = p^{k-1} \cdot Cyc_{p^k}(1)$, thus we have $Cyc_{p^k}(1) = p$. We now proceed to show $Cyc_d(1) = 1$ if $d = p_1^{k_1} p_2^{k_2} \cdots p_r^{k_r}$ for any $r \geq 2$. For this we use k such that $d|k$. We assume $k = p_1^{k'_1} p_2^{k'_2} \cdots p_r^{k'_r}$.

$$\begin{aligned}
\frac{1-q^k}{1-q} &= 1+q+q^2+\cdots+q^{k-1} \\
&= \left(\prod_{j_1=1}^{k'_1} Cyc_{p_1^{j_1}}(q)\right)\left(\prod_{j_2=1}^{k'_2} Cyc_{p_2^{j_2}}(q)\right)\cdots\left(\prod_{j_r=1}^{k'_r} Cyc_{p_r^{j_r}}(q)\right) \\
&\times \left(\prod_{d \text{ is another divisor of } k} Cyc_d(q)\right).
\end{aligned}$$

The expression $\left.\frac{1-q^k}{1-q}\right|_{q=1}$ equals k, and the first r products on the right-hand-side equal $p_1^{k'_1},\ p_2^{k'_2},\ \ldots,\ p_r^{k'_r}$ respectively. Thus the last set of factors, i.e. the cyclotomic polynomials of d with two or more prime factors, must all equal the value 1.

We prove (5.16) analogously. When $N_1 = 2q+2$ (again this is strictly formal), the characteristic equation factors as

$$1-(1+q-N_1)T + qT^2 = (1+T)(1+qT)$$

implying $\alpha_1 = -1$ and $\alpha_2 = -q$. Additionally, $C'(d) = Cyc_d(-1)$ was observed by Ola Veshta on Jun 01 2001, as cited on [Slo, Seq. A020513]. $\square$

Proposition 5.13. *For $d \geq 2$,*

$$\deg_{N_1} ECyc_d = \deg_q ECyc_d = \phi(d),$$

where the Euler ϕ function which counts the number of integers between 1 and $d-1$ which are relatively prime to d.

Proof. As noted in Remark 5.11, we can write $ECyc_d$ as an integral polynomial in $e_1 = \alpha_1 + \alpha_2 = 1 + q - N_1$ and $e_2 = \alpha_1\alpha_2 = q$. The highest degree of N_1 in $ECyc_d$ is therefore equal to the highest degree of $e_1 = \alpha_1 + \alpha_2$, which is the same as the largest m such that $\alpha_1^m\alpha_2^0$ (resp. $\alpha_1^0\alpha_2^m$) is a term in $Cyc_d(\alpha_1)Cyc_d(\alpha_2)$. Thus $\deg_{N_1} ECyc_d(q, N_1) = \deg_{\alpha_1} Cyc_d(\alpha_1) = \phi(d)$. Analogously, the degree of q comes from the highest power of $(\alpha_1\alpha_2)^m$ in $Cyc_d(\alpha_1)Cyc_d(\alpha_2)$. Thus we have shown

$$\deg_q ECyc_d \leq \phi(d).$$

Equality follows from the first half of Proposition 5.12 when $d \geq 2$ since the constant term with respect to N_1, which equals $C(d)Cyc_d(q)$, has degree $\phi(d)$. □

Finally, if one examines the expressions for $ECyc_d(q, N_1)$, one will note that they appear alternating in sign just as the polynomials for N_k, except for the constant term which equals $C(d)Cyc_d(q)$ by Proposition 5.12. More precisely, the author finds the following empirical evidence for such a claim:

Proposition 5.14. *For d between 2 and 104, we obtain*

$$ECyc_d(q, N_1) = Cyc_d(1) \cdot Cyc_d(q) + \sum_{i=1}^{\phi(d)} (-1)^i Q_{i,d}(q) N_1^i$$

where $Q_{i,d}$ is a univariate polynomial with positive *integer coefficients.*

However, the conjecture fails for $d = 105$. In particular,

$$\begin{aligned} ECyc_{105}(q, N_1) &= Cyc_{105}(1) \cdot Cyc_{105}(q) + \sum_{i=1}^{\phi(d)} (-1)^i Q_{i,d}(q) N_1^i \\ &+ \Big(2q^{40} + 18q^{39} + 33q^{38} + 33q^{37} + 33q^{36} + 21q^{35} + 10q^{34} \\ &+ 10q^{13} + 21q^{12} + 33q^{11} + 33q^{10} + 33q^9 + 18q^8 + 2q^7\Big) N_1 \end{aligned}$$

where the $Q_{i,d}$'s are univariate polynomials with positive integer coefficients. (Note that there are 46 coefficients of N_1 in the expansion of $ECyc_{105}(q, N_1)$, only 14 of which have the unexpected sign.)

The number $105 = 3 \cdot 5 \cdot 7$ is significant and interesting from a number theoretic point of view. This number is also the first d such that ordinary cyclotomic polynomial Cyc_d has a coefficient other than $-1, 0$, or 1.

$$\begin{aligned} Cyc_{105} &= 1 + x + x^2 - x^5 - x^6 - 2x^7 - x^8 - x^9 + x^{12} + x^{13} + x^{14} \\ &+ x^{15} + x^{16} + x^{17} - x^{20} - x^{22} - x^{24} - x^{26} - x^{28} + x^{31} + x^{32} \\ &+ x^{33} + x^{34} + x^{35} + x^{36} - x^{39} - x^{40} - 2x^{41} - x^{42} - x^{43} \\ &+ x^{46} + x^{47} + x^{48}. \end{aligned}$$

Despite this counter-example, we still can prove that the coefficients of the $ECyc_d$'s alternate in sign for an infinite number of d's. Specifically, we note that $ECyc_{2^m}$ resemble the coefficients of $N_{2^{m-1}}$, and moreover the pattern we find is

Proposition 5.15.

$$ECyc_{2^m} = 2Cyc_{2^{m-1}}(q) - N_{2^{m-1}}. \tag{5.17}$$

In particular, for i between 1 and $\phi(2^m) = 2^{m-1}$, we get

$$Q_{i,2^m} = P_{i,2^{m-1}} \tag{5.18}$$

where the $P_{i,k}$ are the coefficients of N_k.

Note that in our proof we will use the fact that $ECyc_d$ can be written as

$$Cyc_d(1) \cdot Cyc_d(q) + \sum_{i=1}^{\phi(d)} (-1)^i Q_{i,d}(q) N_1^i$$

where the $Q_{i,d}$'s are univariate polynomials with *possibly* negative coefficients. Therefore, our proof of Proposition 5.15 will actually extend Proposition 5.14 to the case where d is a power of 2 since we previously showed that the $P_{i,d}$'s alternate.

Proof. We note that $Cyc_{2^{m-1}} = 1+q^{2^{m-1}}$ and that (5.18) follows from (5.17). Also, $ECyc_{2^m} = N_{2^m}/N_{2^{m-1}}$ and thus it suffices to prove

$$N_{2^m} = (2 + 2q^{2^{m-1}})N_{2^{m-1}} - N_{2^{m-1}}^2.$$

However, this is a special case of

$$N_2(q, N_1) = (2 + 2q)N_1(q, N_1) - N_1(q, N_1)^2$$

where we plug in $q^{2^{m-1}}$ in the place of q. $\square$

Unfortunately, formulas for $Q_{i,d}$'s in terms of $P_{i,k}$'s when d is not a power of 2 are not as simple. On the other hand, the last part of this proof highlights a principle that has the potential to open up a new direction. Namely, $N_k(q, N_1)$ is defined as the number of points on $E(\mathbb{F}_{q^k})$ where q itself can also be a power of p. Consequently,

$$N_{m\cdot k}(q,\ N_1) = \#E(\mathbb{F}_{q^{m\cdot k}}) = N_m\biggl(q^k,\ N_k\biggr). \tag{5.19}$$

While this relation is immediate given our definition of $N_k = \#E(\mathbb{F}_{q^k})$, when we translate this relation in terms of spanning trees, the relation

$$\mathcal{W}_{mk}(q,t) = \mathcal{W}_m\biggl(q^k, \mathcal{W}_k(q,t)\biggr) \tag{5.20}$$

seems much more novel. Furthermore, in this case, this relation involves only positive integer coefficients and thus motivates exploration for a bijective proof. As noted in Section 5.2.1, such a compositional formula is indicative of the appearance of a linear transformation of x^k or $T_k(x)$, which is also clear from the three-term recurrence satisfied by the $1 + q^k - N_k$'s.

5.3.2 Geometric interpretation of elliptic cyclotomic polynomials

Despite the fact that the above expressions of elliptic cyclotomic polynomials do not have positive coefficients nor coefficients with alternating signs, we can nonetheless describe a set of geometric objects which the elliptic cyclotomic polynomials enumerate.

Theorem 5.16. *We have*

$$ECyc_d = \left| Ker\biggl(Cyc_d(\pi)\biggr) :\ E(\overline{\mathbb{F}_q}) \to E(\overline{\mathbb{F}_q}) \right|$$

where π denotes the Frobenius map, and $Cyc_d(\pi)$ is an element of $End(E) = End(E(\overline{\mathbb{F}_q}))$.

Proof. One of the key properties of the Frobenius map is the fact that $E(\mathbb{F}_{q^k}) = Ker(1 - \pi^k)$, where $1 - \pi^k$ is an element of $End(E)$. See [Sil92] for example. The map $(1-\pi^k)$ factors into cyclotomic polynomials in $End(E)$ since the endomorphism ring contains both integers and powers of π.

Since the maps $Cyc_d(\pi)$ are each group homomorphisms, it follows that the cardinality of $\left|\text{Ker }\left(Cyc_{d_1}Cyc_{d_2}(\pi)\right)\right|$ equals $\left|\text{Ker } Cyc_{d_1}(\pi)\right| \cdot \left|\text{Ker } Cyc_{d_2}(\pi)\right|$. Thus

$$\prod_{d|k} ECyc_d = N_k = \left|\text{Ker }(1-\pi^k)\right| = \left|\text{Ker} \prod_{d|k} Cyc_d(\pi)\right| = \prod_{d|k}\left|\text{Ker } Cyc_d(\pi)\right|,$$

and since the last equation is true for all $k \geq 1$, we must have the relations

$$ECyc_d = \left|\text{Ker } Cyc_d(\pi)\right|. \tag{5.21}$$

for all $d \geq 1$. $\square$

5.4 Acknowledgement

Much of the material in Chapter 5 has been submitted for publication in the paper "Combinatorial Aspects of Elliptic Curves" by Gregg Musiker. The dissertation author is the primary investigator and author of this paper.

6 Connections between elliptic curves and chip-firing

In Chapter 4 we explored elliptic curves from a combinatorial viewpoint, finding that $N_k = \#E(\mathbb{F}_{q^k})$, the number of points over $\mathbb{F}_{q^k}$, could be written as an integral polynomial only depending on q and N_1. This motivated the main topic of that chapter, which was the search for a combinatorial interpretation of these coefficients, one such interpretation involving spanning trees of wheel graphs.

In this chapter, we continue this journey. As discussed in Chapter 3, an elliptic curve E has an abelian group structure, and in this chapter we describe a family of abelian groups whose orders are given by the sequence $\{\mathcal{W}_k(q, N_1)\}$, i.e. groups that are equinumerous with the weighted number of spanning trees of the wheel graph.

6.1 Introduction to chip-firing games

We now step away from elliptic curves momentarily and discuss some fundamental results from the theory of chip-firing games on graphs. The main source for these details is [Big99], though there is an extensive literature on the subject, for example [Mer05, Wag00]. At first glance, this topic might appear totally unrelated to elliptic curves, but we will shortly flesh out the connection. Given a directed (loop-less) graph G, we define a configuration C to be a vector of nonnegative integers, with a coordinate for each vertex of the graph, letting C_i denote the integer corresponding to vertex v_i. One can think of this assignment as a collection of chips placed on each of the vertices. We say that a given vertex v_i can *fire* if the number of chips it holds, C_i, is greater than or equal to its out-degree. If so, firing leads to a new configuration where a chip travels along each outgoing edge incident to v_i. Thus we obtain a configuration C' where $C'_j = C_j + d(v_i, v_j)$ and $C'_i = C_i - d(v_i)$. Here $d(v_i, v_j)$ equals the number of directed edges from v_i to v_j, and $d(v_i)$ is the out-degree of v_i, which of course equals $\sum_{j \neq i} d(v_i, v_j)$.

Many interesting problems arise from this definition. For example, it can be shown [LP01] that the set of configurations reachable from an initial choice of a vector forms a distributive lattice. Thus one can ask combinatorial questions such as examining the structure of this lattice as a poset. Other computations such as the minimal number or expected number of firings necessary to reach configuration C' from C are also common in dynamical systems. In this field, critical configurations are often referred to as the abelian sandpile model [Mer05].

In this classical model, we consider the $\mathbb{Z}$-by-$\mathbb{Z}$ lattice, and presume we are given an initial configuration where each lattice point (site) has a collection of grains of sand on top of it. We further suppose that once a site contains ≥ 4 grains of sand, it topples, sending one grain of sand to each of its neighbors. In this way, by adding sand to this system at a given point, one can cause an *avalanche*. Namely that particular pile of sand will topple onto its neighbors, which in turn might now have too much sand and there will be a smoothing out process of this nature until an equilibrium is achieved. This is known as the *abelian* sandpile model because if two grains are added at two different sites, the resulting equilibrium is independent of the order in which the grains are added. This same notion can be applied in more generality for any graph where we place chips on the vertices, as we will shortly discuss.

For the purposes of relating this topic to an elliptic curve, we will not need the theory of chip-firing games in generality, but consider a variant of the standard chip-firing game, known as the *dollar game*, due to Biggs [Big99]. This game is also a special case of a game with boundary studied by Chung and Ellis [CE02]. In the dollar game, we have the same set-up as before with three changes.

1. We designate one vertex v_0 to be the bank, and allow C_0 to be negative. All the other C_i's still must be nonnegative.

2. To limit extraneous configurations, we presume that the sum $\sum_{i=0}^{\#V-1} C_i = 0$. (Thus in particular, C_0 will be non-positive.)

3. The bank, i.e. vertex v_0, is only allowed to fire if no other vertex can fire. Note that since we now allow C_0 to be negative, v_0 is allowed to fire even when it is smaller than its outdegree.

With this set-up in mind, we define a configuration to be **stable** if v_0 is the only vertex that can fire. We define a configuration C to be **recurrent** if there is a firing sequence which leads back to C. Note that this will necessarily require the use of v_0 firing. We call a configuration **critical** if it is both stable and recurrent.

Proposition 6.1. *For any initial configuration satisfying rules* (1) *and* (2) *above, there exists a* unique *critical configuration that can be reached by a firing sequence, subject to rule* (3)*.*

Proof. See [Big99]. □

We define the **critical group of graph** G, with respect to vertex v_0 to be the set of critical configurations, with addition given by $C_1 \oplus C_2 = \overline{C_1 + C_2}$. Here $+$ signifies the usual pointwise vector addition and $\overline{C_3}$ represents the unique critical configuration reachable from C_3. When v_0 is understood, we will abbreviate this group as the critical group of graph G, and denote it as $\mathcal{C}(G)$.

Theorem 6.2 (Biggs 1999, [Big99])**.** *$\mathcal{C}(G)$ is in fact an abelian (associative) group.*

Proof. If we consider the initial configuration $C_3 = C_1 + C_2$, then by Proposition 6.1, there is a unique critical configuration reachable from C_3. Additionally, we can compute $(C_0 \oplus C_1) \oplus C_2$ or $C_0 \oplus (C_1 \oplus C_2)$ by adding together $C_0 + C_1 + C_2$ pointwise, and then reducing once at the end, rather than reducing twice. Thus associativity and commutativity follow. □

6.2 Connection to elliptic curves

In this section, we describe an alternative definition for the critical group which expresses it in a form more closely resembling the definition of the Picard group or Jacobian of an algebraic variety. Recall that divisors on elliptic curve E over $\mathbb{F}_q$ are formal integral linear combinations of points on $E(\overline{\mathbb{F}_p})$ which are invariant under Frobenius endomorphism π which fixes finite field $\mathbb{F}_q$ ($q = p^k$). We consider relations of the form $D = \sum_i n_i P_i \sim 0$ whenever D is the divisor of a rational function. For an elliptic curve, this simply includes relations generated by those of the form $P + Q + R - 3P_\infty \sim 0$. Furthermore, for elliptic curves, the Abel-Jacobi map provides an isomorphism between the set of equivalence classes $[P - P_\infty]$ and the set of points $P \in E(\mathbb{F}_q)$ [Lan82]. We thus encode all of these relations as a matrix, L_0, and then the Picard group or Jacobian of the elliptic curve is given as $\mathbb{Z}^{\#E(\mathbb{F}_q)}/\text{Im } L_0$.

Returning to the theory of chip-firing games, the literature for this subject occasionally uses the terms Picard group or Jacobian for the critical group as well, e.g. [Lor00]. Let $\mathbb{Z}^{\#V}$ be the set of divisors on the set of vertices V. That is, we consider formal integral (possibly negative) linear combinations of v_1 through $v_{\#V}$. Alternatively we can think of these as the set of homomorphisms from V to $\mathbb{Z}$ or integral vectors of length $\#V$. Let L represent the Laplacian matrix for directed graph G, as defined in Section 5.1., that is $L_{ii} = d(v_i)$ and $L_{i,j} = -d(v_i, v_j)$. The Laplacian will be a singular matrix with a nontrivial nullspace. However, if we take the minor which omits the row and column corresponding to v_0, then we get a nonsingular matrix L_0. The critical group of the graph (V, E) is isomorphic to $\mathbb{Z}^{\#V-1}/\text{Im } L_0$.

While it is more economical to define the group structure in terms of this cokernel, the advantage of the definition via chip-firing is that distinguishing the critical configurations allows us to canonically select coset representatives thereby writing down the explicit elements for this group presentation. Nonetheless, the definition as $\mathbb{Z}^{\#V-1}/\text{Im } L_0$ allows us to use the Matrix-Tree Theorem, as described in Section 5.1, to identify $|\mathcal{C}(G)|$ as the number of spanning trees in G.

In particular, we now have a family of groups, i.e. the critical groups of the (q,t)-wheel graphs, whose orders equal $\mathcal{W}_k(q,t) = -N_k(q,-t)$, We thus turn our attention to the critical group of the (q,t)-wheel graph for $q \geq 0$ and $t \geq 1$, and compare and contrast these groups with the group on elliptic curve $E(\mathbb{F}_{q^k})$ for

$k \geq 1$ and various E's.

Remark 6.3. While it now suffices to work in terms of these groups of critical configurations, for completeness we provide here a natural bijection between spanning trees of the (q, t)-wheel graphs and critical configurations. Such a natural bijection does not exist in general, although Biggs and Winkler have an algorithmic bijection, as appears in [BW] and also reproduced in [EI02]. Nonetheless, in this case, one could define the desired group structure directly on (colored) spanning trees.

Proposition 6.4. *There exists an explicit* bijection *between critical configurations and spanning trees (at least in the case of the directed (q, t)-wheel multi-graph). This map induces an isomorphism of groups.*

Specifically pick one of the vertices on the rim to be v_1, and label v_2 through v_k clockwise. Label the central hub as v_0. For i between 1 and k, if $1 \leq C_i \leq q$, then fill in the arc between v_{i-1} and v_i, labeling it with the number C_i. (In the case of $i = 1$ we use the arc between v_k and v_1 instead.) If $1 + q \leq C_i \leq q + t$ then fill in the spoke between v_0 and v_i and label it with number C_i. After filling in the edges as indicated we will get a subgraph of a spanning tree. To complete this subgraph to a tree, fill in additional arcs using the following rule: one may fill in an arc from v_{i-1} to v_i, and label it with a q, if and only if $C_i \in \{1 + q, \ldots, q + t\}$. In other words, if $C_i = 0$ then this will contribute no arc nor a spoke.

Proof. We defer the proof of this theorem until Section 6.3 where we precisely describe which critical configurations actually arise. It will then be clear that the list of configurations that show up as the image of a spanning tree, and the list of possible critical configurations, are equivalent. Since the described map is injective by construction, we have the desired bijection. □

6.2.1 Group structure

We now return to the main topic at hand, namely elliptic curves. An elliptic curve over a finite field has a well-known group structure. In fact, it is the product of at most two cyclic groups. One way to prove this is by showing that for $gcd(N,p) = 1$, the $[N]$-torsion subgroup of $E(\overline{\mathbb{F}}_p)$ (also denoted as $E[N]$) is isomorphic to $\mathbb{Z}/N\mathbb{Z} \times \mathbb{Z}/N\mathbb{Z}$ and that $E[p^r]$ is either 0 or $\mathbb{Z}/p^r\mathbb{Z}$.

Since we know that the critical group of graphs are also abelian groups, this motivates the question: what is the group decomposition of the $\mathcal{C}(G)$'s? The case of a simple wheel graph W_k was explicitly found by Biggs to be

$$\mathbb{Z}/L_k\mathbb{Z} \times \mathbb{Z}/L_k\mathbb{Z} \quad \text{or} \quad \mathbb{Z}/F_{k-1}\mathbb{Z} \times \mathbb{Z}/5F_{k-1}\mathbb{Z}$$

depending on whether k is odd or even, respectively [Big99]. Here L_k is the kth Lucas number and F_k is the kth Fibonacci number.

Determining such structures of critical groups has been the subject of several papers recently, e.g. [JNR03, Max06], and a common tool is the Smith normal form of the Laplacian. Fortunately, we already know the Smith normal form for the case we care about, namely for the (q,t)-wheel graphs.

Theorem 6.5. *$\mathcal{C}(W_k(q,N_1))$ is isomorphic to at most two cyclic groups, a property that this sequence of critical groups shares with the family of elliptic curve groups over finite fields.*

Proof. By Theorem 5.4, the Smith Normal form of the reduced Laplacian L_0 for the graphs $W_k(q,t)$ consists of a diagonal of ones followed by at most two integers greater than one. Since the Smith normal form of M gives the cyclic decomposition of the group defined by coker $M = \mathbb{Z}^k/\text{Im } M$, we conclude these critical groups can be decomposed into at most two cyclic groups. $\square$

In addition to a presentation for $\mathcal{C}(W_k(q,N_1))$, we also get a more explicit presentation of $E(\mathbb{F}_{q^k})$ in certain cases.

Theorem 6.6. *If $E(\mathbb{F}_q) \cong \mathbb{Z}/N_1\mathbb{Z}$, as opposed to the product of two cyclic groups, and $End(E) \cong \mathbb{Z}[\pi]$, then*

$$E(\mathbb{F}_q^k) \cong \mathbb{Z}^k \Big/ M_k\mathbb{Z}^k$$

for all $k \geq 1$. That is, $E(\mathbb{F}_{q^k})$ is the cokernel of the image of M_k. Furthermore, there exists a point $P \in E(\mathbb{F}_{q^k})$ with property $\pi^m(P) \neq P$ for all $1 < m < k$ such that we can take $\mathbb{Z}^k$ as being generated by $\{P, \pi(P), \ldots, \pi^{k-1}(P)\}$ under this presentation.

Proof. A theorem of Lenstra [Len96] says that an *ordinary* elliptic curve over $\mathbb{F}_q$ has a group structure in terms of its endomorphism ring, namely,

$$E(\mathbb{F}_{q^k}) \cong End(E) \Big/ (\pi^k - 1).$$

Wittman [Wit01] gives an explicit description of the possibilities for $End(E)$, given q and $E(\mathbb{F}_q)$. It is well known, e.g. [Sil92], that the endomorphism ring in the ordinary case is an order in an imaginary quadratic field. This means that

$$End(E) \cong \mathcal{O}_g = \mathbb{Z} \oplus g\delta\mathbb{Z}$$

for some $g \in \mathbb{Z}_{\geq 0}$ and $\delta = \sqrt{D}$ or $\frac{1+\sqrt{D}}{2}$ according to d's residue modulo 4. Wittman shows that for a curve E with conductor f, the possible g's that occur satisfy $g|f$ as well as

$$n_1 = \gcd(a-1, g/f).$$

The conductor f and constant a are computed by rewriting the Frobenius map as $\pi = a + f\delta$, and n_1 is the unique positive integer such that $E(\mathbb{F}_q) \cong \mathbb{Z}/n_1\mathbb{Z} \times \mathbb{Z}/n_2\mathbb{Z} \quad (n_1|n_2)$.

We focus here on the case when $g = f$ and $End(E) \cong \mathbb{Z}[\pi]$. In particular, n_1 must be equal to one in this case, and so the condition that $End(E) = \mathbb{Z}[\pi]$ is actually a sufficient hypothesis. Since $E(\mathbb{F}_{q^k}) \cong \mathbb{Z}[\pi]/(1-\pi^k)$ in this case, we get

$$E(\mathbb{F}_{q^k}) \cong \mathbb{Z}[x]/(x^2 - (1+q-N_1)x + q, \quad x^k - 1)$$

with x transcendent over $\mathbb{Q}$. Thus

$$\begin{aligned} E(\mathbb{F}_{q^k}) \cong \mathbb{Z}\{1, x, x^2, \ldots, x^{k-1}\} \Big/ \\ \Big(& x^2 - (1+q-N_1)x + q,\ x^3 - (1+q-N_1)x^2 + qx,\ \ldots, \\ & x^{k-1} - (1+q-N_1)x^{k-2} + qx^{k-3},\ 1 - (1+q-N_1)x^{k-1} + qx^{k-2}, \\ & x - (1+q-N_1) + qx^{k-1} \Big) \end{aligned}$$

and using matrix M_k, as defined above, we obtain the desired presentation for $E(\mathbb{F}_{q^k})$ in this case. $\square$

Question 6.7. What can we say in the case of another endomorphism ring, or the case when $E(\mathbb{F}_q)$ is not cyclic?

6.2.2 Analogues of elliptic cyclotomic polynomials

We found for elliptic curves that $ECyc_d(q, N_1)$ counted the number of points in the kernel of the isogeny $Cyc_d(\pi)$ where π is the Frobenius isogeny. Since

$$N_k = \prod_{d|k} ECyc_d(q, N_1)$$

and $\mathcal{W}_k(q,t) = -N_k\Big|_{N_1 \to -t}$, it also makes sense to consider the decomposition

$$\mathcal{W}_k(q,t) = \prod_{d|k} WCyc_d(q,t)$$

where $WCyc_d(q,t) = -ECyc_d|_{N_1 \to -t}$.

Table 6.1: The polynomials $WCyc_d(q,t)$ for small d.

$$\begin{aligned}
WCyc_1 &= t \\
WCyc_2 &= t + 2(1+q) \\
WCyc_3 &= t^2 + (3+3q)t + 3(1+q+q^2) \\
WCyc_4 &= t^2 + (2+2q)t + 2(1+q^2) \\
WCyc_5 &= t^4 + (5+5q)t^3 + (10+15q+10q^2)t^2 + (10+15q+15q^2+10q^3)t \\
&+ 5(1+q+q^2+q^3+q^4) \\
WCyc_6 &= t^2 + (1+q)t + (1-q+q^2) \\
WCyc_8 &= t^4 + (4+4q)t^3 + (6+8q+6q^2)t^2 + (4+4q+4q^2+4q^3)t + 2(1+q^4) \\
WCyc_9 &= t^6 + (6+6q)t^5 + (15+24q+15q^2)t^4 + (21+36q+36q^2+21q^3)t^3 \\
&+ (18+27q+27q^2+27q^3+18q^4)t^2 \\
&+ (9+9q+9q^2+9q^3+9q^4+9q^5)t + 3(1+q^3+q^6) \\
WCyc_{10} &= t^4 + (3+3q)t^3 + (4+3q+4q^2)t^2 + (2+q+q^2+2q^3)t \\
&+ (1-q+q^2-q^3+q^4) \\
WCyc_{12} &= t^4 + (4+4q)t^3 + (5+8q+5q^2)t^2 + (2+2q+2q^2+2q^3)t + (1-q^2+q^4)
\end{aligned}$$

We ask the same question as before, namely does there exist a combinatorial or geometric interpretation of these polynomials.

Remark 6.8. The coefficients of the $WCyc_d$'s are always integers, but not necessarily positive, as seen in the constant coefficient, as well as in the counter-example $WCyc_{105}$. Nonetheless, plugging in specific integers $q \geq 0$ and $t \geq 1$ do in fact result in positive expressions, which factor $\mathcal{W}_k(q,t)$. It is these values that we are interested in understanding.

Indeed, we consider the following properties of the $\mathcal{C}(W_k(q,t))$'s that allow us to derive a result analogous to the elliptic cyclotomic case.

Proposition 6.9. *The identity map induces an injective group homomorphism between $\mathcal{C}(W_{k_1}(q,t))$ and $\mathcal{C}(W_{k_2}(q,t))$ whenever $k_1|k_2$. More precisely, we let $\mathcal{C}(W_{k_1}(q,t))$ embed into $\mathcal{C}(W_{k_2}(q,t))$ by letting $w \in \mathcal{C}(W_{k_1}(q,t))$ map to the word $www\ldots w \in \mathcal{C}(W_{k_2}(q,t))$ using $\frac{k_2}{k_1}$ copies of w.*

Define ρ to be the rotation map on $\mathcal{C}(W_k(q,t))$. If we consider elements of the critical group to be configuration vectors, then we mean circular rotation of the elements to the right. On the other hand, ρ acts by rotating the rim vertices of W_k clockwise if we view elements of $\mathcal{C}(W_k(q,t))$ as spanning trees.

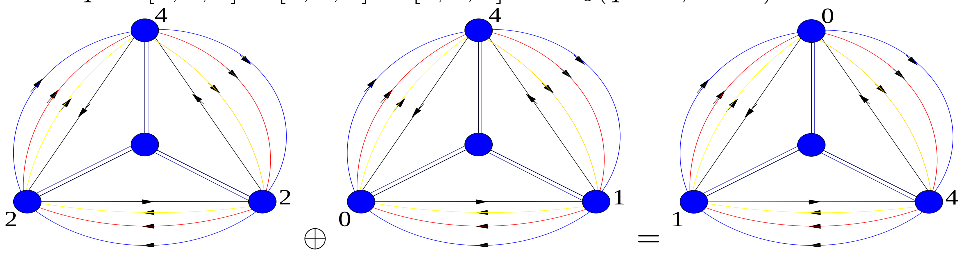
Proposition 6.10. *The kernel of $(1-\rho^{k_1})$ acting on $\mathcal{C}(W_{k_2}(q,t))$ is subgroup $\mathcal{C}(W_{k_1}(q,t))$ whenever $k_1|k_2$.*

Proof. We prove both of these propositions simultaneously, by noting that chip firing is a local process. Namely, if k_1 divides k_2 and we add two configurations of $W_{k_1}(q,t)$ together pointwise to get configuration C, then lift C to a length k_2 configuration C' of $W_{k_2}(q,t)$ by periodically extending length k_1 vector C. Then the claim is that if C reduces to unique critical configuration $\overline{C}$, then C' also reduces to $\overline{C}$'s periodic extension. To see this, observe that every time vertex $v \in W_{k_1}(q,t)$ fires in the reduction algorithm, then we could simultaneously fire the set of vertices of $W_{k_2}(q,t)$ in the image of v after lifting. In other words, if $v_i \in W_{k_1}(q,t)$ fires, we fire $\{v'_i, v'_{i+k_2/k_1}, v'_{i+2k_2/k_1}, \ldots\} \in W_{k_2}(q,t)$ thus obtaining the lift of the configuration reached after v fires. $\square$

We therefore can define a direct limit

$$\mathcal{C}(\overline{W}(q,t)) \cong \bigcup_{k=1}^{\infty} \mathcal{C}(W_k(q,t))$$

Example: $[2,4,2] \oplus [0,4,1] \equiv [1,0,4]$ in $\mathcal{W}_3(q=3, t=2)$ versus



$[2,4,2,2,4,2] \oplus [0,4,1,0,4,1] \equiv [1,0,4,1,0,4]$ in $\mathcal{W}_6(q=3, t=2)$

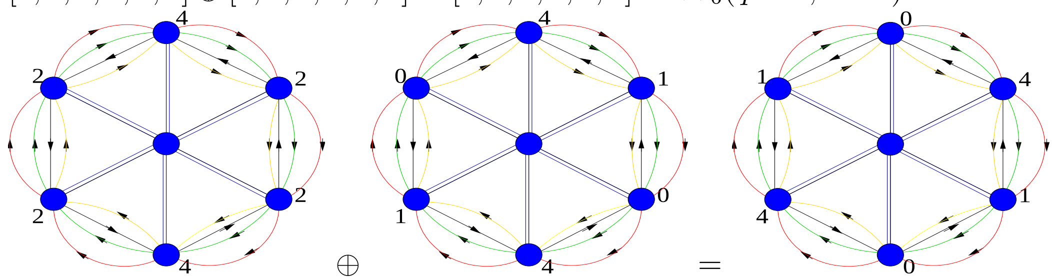


Figure 6.1: Illustrating Propositions 6.9 and 6.10.

where ρ provides the transition maps.

Another view of $\mathcal{C}(\overline{W}(q,t))$ is as the set of bi-infinite words which are (1) periodic, and (2) have fundamental subword equal to a configuration vector in $\mathcal{C}(W_k(q,t))$ for some $k \geq 1$. In this interpretation, map ρ acts on $\mathcal{C}(\overline{W}(q,t))$ also. In this case, ρ is the shift map, and in particular we obtain

$$\mathcal{C}(W_k(q,t)) \cong Ker(1-\rho^k) : \mathcal{C}(\overline{W}(q,t)) \to \mathcal{C}(\overline{W}(q,t)).$$

We now can describe our variant of Theorem 5.16.

Theorem 6.11.

$$WCyc_d = \left| Ker\Big(Cyc_d(\rho)\Big) : \ \mathcal{C}(\overline{W}(q,t)) \to \mathcal{C}(\overline{W}(q,t)) \right|$$

where ρ denotes the shift map, and $\mathcal{C}(\overline{W}(q,t))$ is the direct limit of the sequence $\{\mathcal{C}(W_k(q,t))\}_{k=1}^{\infty}$.

Proof. The proof is analogous to the elliptic curve case. Since the maps $Cyc_{d_1}(\rho)$ and $Cyc_{d_2}(\rho)$ are group homomorphisms, we get

$$|\text{Ker } Cyc_{d_1}(\rho)\ Cyc_{d_2}(\rho)| = |\text{Ker } Cyc_{d_1}(\rho)| \cdot |\text{Ker } Cyc_{d_2}(\rho)|$$

and the rest of the proof follows as in Chapter 4. $\square$

Thus we identify shift map ρ as being the analogue of the Frobenius map π on elliptic curves. In addition to ρ's appearance in Theorem 6.11, two other comparisons with π are highlighted below.

1.

$$\begin{aligned} \mathcal{C}(W_k(q,t)) &\cong Ker(1-\rho^k) : \mathcal{C}(\overline{W}(q,t)) \to \mathcal{C}(\overline{W}(q,t)) \quad \text{just as} \\ E(\mathbb{F}_{q^k}) &= Ker(1-\pi^k) : E(\overline{\mathbb{F}_q}) \to E(\overline{\mathbb{F}_q}). \end{aligned}$$

2. We get the equation

$$\rho^2 - (1+q+t)\rho + q = 0,$$

which can be read off from matrix M_k and the configuration vectors' images under clockwise and counter-clockwise rotation. This is a simple analogue of the characteristic equation

$$\pi^2 - (1+q-N_1)\pi + q = 0$$

of the Frobenius map π.

6.3 Characterization of critical configurations

In this section we completely characterize critical configurations of the (q,t)-wheel graph. Furthermore, we will shortly see a deterministic finite automaton which admits such critical configurations. As an added bonus, we can construct a zeta function of such a system which is intimately connected to the zeta function of the elliptic curve.

This new characterization of critical configurations also proves Theorem 6.4, giving a bijection between critical configurations and spanning trees.

Proposition 6.12. *A configuration $C = [c_1, c_2, \ldots, c_k]$ of the wheel graph $W_k(q,t)$ is stable if and only if $0 \le c_i \le q+t$ for all $1 \le i \le k$.*

Proof. It is clear that we disallow $c_i < 0$ as a legal configuration by our definition. If such a configuration were to come up, we could add t to every value c_i, simulating the firing of the central vertex. If on the other hand, there exists $c_i \geq 1+q+t$, with all other $c_i \geq 0$, then vertex v_i can fire resulting in a new nonnegative configuration. Otherwise, if all c_i are in the specified range, we have a stable configuration where no vertex except the hub can fire. □

We recall that any stable configuration C is **critical** if and only if it is recurrent, meaning that after adding t to every c_i and applying the chip-firing rules, we arrive back at stable configuration C.

Proposition 6.13. *There exists a unique critical configuration* reachable *from a given stable configuration in the case of the (q,t)-wheel graph.*

Proof. This is a corollary of Proposition 6.1 but we will give the details of the proof for this special case. □

Lemma 6.14. *Let C be a stable configuration, with $\sum_{i=1}^{k} c_i = N$. If C is reachable from some configuration C' (which is not necessarily stable) with $\sum_{i=1}^{k} c'_i > N$, then C is actually critical.*

Proof. We need only check that if we add t to all values c_i and apply the chip-firing rules, we will reach C again. Given the sum of the rows of the Laplacian matrix, there will be some firing sequence such that every vertex will fire, and thus the result being the subtraction of t from every c_i, thus we obtain C again. See [Big99] for more details in the case of a general graph. □

Lemma 6.15. *While we apply the chip-firing rules, every stage will decrease the $\sum_{i=1}^{k} c_i$ by t. In particular, if there are two stable configurations which are equivalent, we will reach the configuration with the biggest $\sum_{i=1}^{k} c_i$ first. Thus, this vector will be the critical configuration out of this equivalence class.*

Proof. This claim follows from the definition of the Laplacian and Lemma 6.14. □

Thus we have proven Proposition 6.13 for the case of the (q,t)-wheel graph. For a more general proof, see [Big99].

Lemma 6.16. *Any critical configuration $[c_1, \ldots, c_k]$ will have at least one element $c_i = B$ such that $B \in \{1+q, \ldots, q+t\}$.*

Proof. Assume otherwise. Then $c_i \in \{0, 1, \ldots, q\}$ for all $1 \leq i \leq k$. Consequently, we may add t to every c_i and still obtain a stable configuration. Thus the initial configuration is smaller and cannot be critical. $\square$

Theorem 6.17. *Any configuration C is critical if and only if it consists of a circular concatenation of blokcs of the form*

$$B, M_1, \ldots, M_r$$

with the properties (1) $B \in \{q+1, \ldots, q+t\}$, (2) $M_i \in \{0, 1, \ldots, q\}$, *and* (3) *if* $M_j = 0$, *then* $M_{j+1} = \cdots = M_r = q$.

Proof. We have already shown that there exists at least one $c_i = B$ with $B > q$. Thus we prove this Theorem by induction on n, the number of such elements. Consider such a block in context, and presume it is of form

$$\cdots, M_n^{k_n} \mid B_1, M_1^1, M_1^2, \ldots, M_1^{k_1} \mid B_2, \cdots$$

where $M_p^i \in \{0, 1, \ldots, q\}$ and $B_p \in \{1+q, \ldots, q+t\}$. Here $M_n^{k_n}$ and B_2 represent the end of the previous block and the beginning of the next block, respectively. The heart of the proof is the verification of the following claim.

Claim 6.18. Such a configuration cannot be recurrent unless $M_p^{j_p} = 0$ implies that the remaining M_p^i's, i.e. $M_p^{j_p+1}$ through $M_p^{k_p}$, are equal to q.

Without loss of generality, we will work with $p = 1$ and let $j_1 = j$, $k_1 = k$, $M_n^{k_n} = M_0$. Assume that M_1^1 through $M_1^{j-1} \in \{1, 2, \ldots q\}$. We add t to every element of C, getting $C + [t]$, and then reduce via the chip-firing rules whenever we encounter an element with value greater or equal to $1 + q + t$. Configuration $C + [t]$ contains element $B_1 + t$, with value $\geq 1 + q + t$, but all other elements of the block are $< 1 + q + t$. Once we replace $B_1 + t$ with $B_1 - 1 - q$, and its neighbors with $M_0 + t + 1$ and $M_1^1 + q + t$, respectively, we reduce $M_1^1 + q + t$ since its entry is now $\geq 1 + q + t$. We continue inductively until we reach $M_1^j + q + t$ which is less than $1 + q + t$ since $M_1^j = 0$ by assumption. At this point, the block looks like

$$M_0 + t + 1 \mid B_1 - q, M_1^1, \ldots, M_1^{j-1} - 1, q + t, M_1^{j+1} + t, \ldots, M_1^k + t \mid B_2 + t.$$

Since $B_2 + t \geq 1 + q + t$, we can reduce this block further as

$$M_0 + t + 1 \mid B_1 - q, M_1^1, \ldots, M_1^{j-1} - 1, q + t, M_1^{j+1} + t, \ldots, M_1^k + t + 1 \mid B_2 - 1 - q.$$

By propagating the same reductions to the rest of the configuration, we reduce to a configuration C' which is made up of blocks of the form

$$B_p - q, M_p^1, \ldots, M_p^{j_p-1} - 1, q + t, M_p^{j_p+1} + t, \ldots, M_p^{k_p} + t + 1$$

in lieu of

$$B_p, M_p^1, \ldots, M_p^{j_p-1}, 0, M^{j_p+1}, \ldots, M^{k_p}.$$

Since $M_p^i \leq q$, all elements of C' are less than $1 + q + t$ except possibly for the last elements of each block, e.g. $M_p^k + t + 1$. If all of the M_p^k's are less than q, then C' is stable, and thus the original configuration C is not recurrent, nor critical as assumed.

Thus, without loss of generality, assume that $M_1^k = q$. We then can reduce block

$$\left| \; B_1 - q, M_1^1, \ldots, M_1^{j-1} - 1, q + t, M_1^{j+1} + t, M_1^{j+2} + t, \ldots, M_1^{k-1} + t, q + t + 1 \; \right| \; B_2 - 1 - q$$

on the right-hand-side and obtain

$$\left| \; B_1 - q, M_1^1, \ldots, M_1^{j-1} - 1, q + t, M_1^{j+1} + t, M_1^{j+2} + t, \ldots, M_1^{k-1} + t + 1, 0 \; \right| \; B_2 - 1.$$

By analogous logic, we must have that $M_1^{k-1} = q$ and continuing iteratively, we reduce to

$$M_0 + t + 1 \; \left| \; B_1 - q, M_1^1, \ldots, M_1^{j-1} - 1, q + t + 1, 0, q, \ldots, q, q \; \right| \; B_2 - 1$$

which is equivalent to

$$M_0 + t + 1 \; \left| \; B_1 - q, M_1^1, \ldots, M_1^{j-1}, 0, q, q, \ldots, q, q \; \right| \; B_2 - 1.$$

Finally, $M_0 = M_n^{k_n}$ so we indeed obtain

$$q \; \left| \; B_1, M_1^1, \ldots, M_1^{j-1}, 0, q, q, \ldots, q, q \; \right| \; B_2$$

after iterating over all the blocks to the right and wrapping around. $\square$

Considering these as elements of $\mathcal{C}(W_k(q,t)) \subset \mathcal{C}(\overline{W}(q,t))$, we identity $C_1, \ldots, C_k$ with periodic string

$$\ldots C_k, C_1, C_2, \ldots C_{k-1}, C_k, C_1, \ldots.$$

Thus we have in fact simultaneously given criteria for testing criticality in $\mathcal{C}(W_k(q,t))$ for length arbitrary length k, as well as for an element in direct limit $\overline{\mathcal{C}}(W_k(q,t))$.

6.4 Connections to deterministic finite automata

A deterministic finite automaton (DFA) is a finite state machine M built to recognize a given language L, i.e. a set of words in a specific alphabet. To test whether a given word ω is in language L we write down ω on a strip of tape and feed it into M one letter at a time. Depending on which state the machine is in, it will either accept or reject the character. If the character is accepted, then the machine's next state is determined by the previous state and the relevant character on the strip. As the machine changes states accordingly, and the entire word is fed into the machine, if all letters of ω are accepted, then ω is an element of language L.

For our purposes we consider an automaton M_G with three states, which we label as A, B, and C. In state A we either accept a character in $\{1+q, \ldots, q+t\}$ and return to state A, accept a character in $\{1, \ldots, q\}$ and move to state B, or accept the character 0 and move to state C.

On the other hand, in state B we either accept a character in $\{1+q, \ldots, q+t\}$ and move to state A, accept a character in $\{1, \ldots, q\}$ and return to state B, or accept character 0 and move to state C.

Finally, in state C we either accept a character in $\{1+q, \ldots, q+t\}$ and move to state A, or accept character q and return to state C. A character in $\{1, \ldots, q\}$ is not accepted while in state C. This DFA is illustrated here, with its transition matrix also given.

We consider the set of words $\mathcal{L}(q,t)$ which are accepted by M_G with the properties (1) the initial state of M_G is the same as its final state, and (2) M_G is in

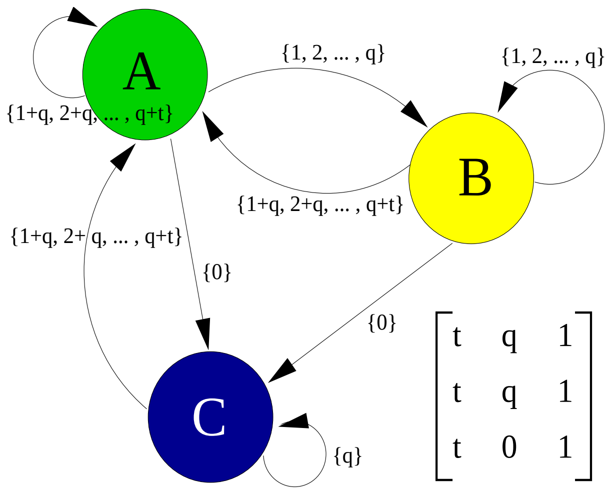


Figure 6.2: Deterministic finite automaton M_G.

state A at some point while verifying ω. Comparing definitions, we observe that the set of such words is in fact the set of critical configurations, as described in Section 6.3. We can in fact characterize this set even more concretely.

Proposition 6.19. *The set $\mathcal{L}(q,t)$ is a* **regular language***, i.e. a set of words which can be described by a DFA $\mathcal{D}_{\mathcal{L}}$. In particular, word ω is in $\mathcal{L}(q,t)$ if and only if ω is admissible by $\mathcal{D}_{\mathcal{L}}$.*

Proof. Regular languages can be built by taking complements, the Kleene star, unions, intersections, images under homomorphisms, and concatenations. Thus we can prove $\mathcal{L}(q,t)$ is regular by decomposing it as the union over all cyclic shifts, a homomorphism, of concatenation of the blocks of form $B, M_1, M_2, \ldots, M_k$. □

More explicitly, we can also use M_G to build a DFA recognizing $\mathcal{L}(q,t)$, thus giving a second proof. First, machine M_G as described is not technically a DFA since we are not specifying which of the three states is the initial state and what state the DFA moves to from state C when it encounters a character in $\{0,1,2,\ldots,q-1\}$. We also have the added restrictions that a word is only admissible if the DFA goes through state A along its path, and that words admitted by closed paths in this DFA.

However, this can be easily rectified. First, we add four additional states: a initial state I, two states $\tilde{B}$ $\tilde{C}$, and a dead state D. Start state I connects to states

A, $\tilde{B}$ and $\tilde{C}$, moving to A if the first letter is $\geq 1+q$, moving to $\tilde{C}$ if the first letter is 0, and moving to $\tilde{B}$ otherwise. Additionally, state $\tilde{B}$ connects to A, $\tilde{B}$, and $\tilde{C}$ just as B connects to A, B, and C; similarly, $\tilde{C}$ connects to A and $\tilde{C}$ just as C connects to A and C. When the machine is in state C or $\tilde{C}$, and a character from $\{0, 1, 2, \ldots, q-1\}$ is read, the machine moves to the dead state D which always loops back to itself. Letting states A, B, and C be the only final/terminal states of this DFA, we now have the property that a word is only admissible if the DFA goes through state A at some point along its path.

We now have to deal with the restriction that a word is admissible only if the word induces a cycle of states in the DFA. To this end, we expand the DFA even further essentially copying it three times and making sure the terminal states correspond to the first state reached, i.e. immediately following the start state.

6.5 Another kind of zeta function

Returning to the original formulation, critical configurations correspond to closed paths in DFA M_G which go through state A. Since a cycle involving both states B and C but not state A is impossible, the only cycles we need to disallow are those containing only state B and those cycles containing only state C. Such words, i.e. the set $\mathcal{L}(q,t)$ is a **cyclic language** since the set is closed under circular shift (more precisely $uv \in \mathcal{L}(q,t)$ if and only if $vu \in \mathcal{L}(q,t)$ for all u, v).

Regular cyclic languages such as $\mathcal{L}(q,t)$ were studied in [BR90], and we can even define a zeta function for them. The zeta function of a cyclic language L is defined as

$$\zeta(L,T) = \exp\left(\sum_{k=1}^{\infty} \mathcal{W}_k \frac{T^k}{k}\right)$$

where $\mathcal{W}_k$ is the number of words of length k. Alternatively, this can be written as

$$\zeta(L,T) \;=\; \exp\left(\sum_{\text{allowed closed paths } P} (\#\text{ words admissible by path } P)\;\; T^k\right).$$

Theorem 6.20 (Berstel and Reutenauer)**.** *The zeta function of a cyclic and regular language is rational.*

Proof. See [BR90] or [Reu97]. □

The **trace** of an automaton $\mathcal{A}$ is the language of words generated by closed paths in $\mathcal{A}$. Such a language is always cyclic and regular by construction, and in fact has a zeta function with an explicit formula.

Proposition 6.21.

$$\zeta(trace(\mathcal{A})) = \frac{1}{\det(I - M \cdot T)},$$

where M encodes the number of directed edges between state i and state j in $\mathcal{A}$.

This matrix is in fact the transition matrix provided above with the example of automaton M_G.

Proof. We omit this proof, again referring the reader to [BR90]. However, we also take this opportunity to mention that the proof is an application of MacMahon's Master Theorem [Mac60] which relates the generating function of traces to a determinantal formula, or more precisely the characteristic polynomial of a matrix. Moreover, analogies between the zeta function of a language and the zeta function of a variety are even clearer since the proof of the Weil conjectures via étale cohomology also involve such determinantal expressions. □

Using this terminology, we can describe the set of critical configurations of (q,t)-W_k as the language obtained by taking the trace of M_G *minus* the trace of cycles only containing state B *minus* the trace of cycles only containing state C. We again note that all other circuits with the same initial and final state necessarily need to contain state A since there are no cycles containing both state B and C but not A. There is no way to go from state C to state B without going through state A first, given the definition of M_G.

Thus the zeta function of this cyclic language is given as

$$\frac{\det([1 - qT]) \det([1 - T])}{\det(I - MT)}$$

where the factor of $\det([1 - qT])$ correspond to the trace of cycles containing state B alone, and $\det([1 - T])$ corresponds to the trace of cycles containing state C

alone. On the other hand, matrix M is the 3-by-3 matrix encoded by the number of directed edges between the various states.

$$\begin{bmatrix} t & q & 1 \\ t & q & 1 \\ t & 0 & 1 \end{bmatrix}$$

Thus we arrive at the following expression for $\zeta(\mathcal{L}(q,t))$, namely

$$\exp\left(\sum_{k=1}^{\infty} \frac{\mathcal{W}_k}{k} T^k\right) = \frac{(1-qT)(1-T)}{1-(1+q+\mathcal{W}_1)T+qT^2}$$

where $\mathcal{W}_k$ equals the number of primitive cycles in M_G, which contain state A but starting at any of the three states.

At this point, we have yet a fourth proof of the Theorem 4.13, which states $N_k = -\mathcal{W}_k(q, -N_1)$. The reasoning being

$$\begin{aligned} \exp\left(\sum_{k\geq 1} \frac{\mathcal{W}_k}{k} T^k\right) &= \frac{(1-qT)(1-T)}{1-(1+q+t)T+qT^2} \\ &= \left(\frac{1-(1+q+t)T+qT^2}{(1-qT)(1-T)}\right)^{-1} \\ &= \left(Z(E,T)|_{N_1=-t}\right)^{-1} \\ &= \exp\left(-\sum_{k\geq 1} \frac{N_k}{k} T^k\right)\Bigg|_{N_1=-t}. \end{aligned}$$

6.6 Conclusions and topics for further research

In this thesis, we have studied the theory of elliptic curves over finite fields with an eye towards combinatorial results. To this end, we have provided symmetric function interpretations of the zeta function, and have given combinatorial interpretations to the coefficients of the polynomial expressions of N_k in terms of q and N_1. In particular, we have illustrated interpretations in terms of Fibonacci numbers, Lucas numbers, and spanning trees; with these in mind, uncovering various identities of a combinatorial flavor.

As a bonus, as described in Chapter 6, the relationship between elliptic curves and spanning trees appears even more pronounced than one would have guessed from the motivation of Theorem 4.13. Not only do we have formal identities relating the number of spanning trees of wheel graphs and number of points on elliptic curves, but we also have connections between the corresponding group structures of these two families of objects. The connections described here inspire further exploration for connections between these two topics. In addition, future research will consider more techniques from areas such as combinatorics on words and dynamical system and use these to ask or answer questions about elliptic curves.

In Chapter 2, we also discussed combinatorial aspects of algebraic curves in general, using symmetric function theory for the general case. With such techniques in mind, the study of higher genus curves such as hyperelliptic curves, or other classes of abelian varieties will provide many other interesting topics for exploration.

References

[BY06] Arthur T. Benjamin and Carl R. Yerger, *Combinatorial interpretations of spanning tree identities*, Bull. Inst. Combin. Appl. **47** (2006), 37–42.

[Big99] N. L. Biggs, *Chip-firing and the critical group of a graph*, J. Algebraic Combin. **9** (1999), no. 1, 25–45.

[BW] N. L. Biggs and P. Winkler, *Chip-firing and the chromatic polynomial*, CDAM Research Report Series, 97–03.

[Bom74] Enrico Bombieri, *Counting points on curves over finite fields (d'après S. A. Stepanov)*, Séminaire Bourbaki, 25ème année (1972/1973), Exp. No. 430, Springer, Berlin, 1974, pp. 234–241. Lecture Notes in Math., Vol. 383.

[BP86] F. T. Boesch and H. Prodinger, *Spanning tree formulas and Chebyshev polynomials*, Graphs Combin. **2** (1986), no. 3, 191–200.

[BR90] Jean Berstel and Christophe Reutenauer, *Zeta functions of formal languages*, Trans. Amer. Math. Soc. **321** (1990), no. 2, 533–546.

[BT51] H. D. Block and H. P. Thielman, *Commutative polynomials*, Quart. J. Math., Oxford Ser. (2) **2** (1951), 241–243.

[BE95] Peter Borwein and Tamás Erdélyi, *Polynomials and polynomial inequalities*, Graduate Texts in Mathematics, vol. 161, Springer-Verlag, New York, 1995.

[Cas91] J. W. S. Cassels, *Lectures on elliptic curves*, London Mathematical Society Student Texts, vol. 24, Cambridge University Press, Cambridge, 1991.

[CE02] Fan Chung and Robert B. Ellis, *A chip-firing game and Dirichlet eigenvalues*, Discrete Math. **257** (2002), no. 2-3, 341–355, Kleitman and combinatorics: a celebration (Cambridge, MA, 1999).

[Del74] Pierre Deligne, *La conjecture de Weil. I*, Inst. Hautes Études Sci. Publ. Math. (1974), no. 43, 273–307.

[DF91] David S. Dummit and Richard M. Foote, *Abstract algebra*, Prentice Hall Inc., Englewood Cliffs, NJ, 1991.

[Dwo60] Bernard Dwork, *On the rationality of the zeta function of an algebraic variety*, Amer. J. Math. **82** (1960), 631–648.

[EI02] R.B. Ellis III, *Chip-Firing Games with Dirichlet Eigenvalues and Discrete Greens Functions*, Ph.D. thesis, UCSD, 2002.

[ER91] Ömer Eğecioğlu and Jeffrey B. Remmel, *Brick tabloids and the connection matrices between bases of symmetric functions*, Discrete Appl. Math. **34** (1991), no. 1-3, 107–120, Combinatorics and theoretical computer science (Washington, DC, 1989).

[Fre01] Gerhard Frey, *Applications of arithmetical geometry to cryptographic constructions*, Finite fields and applications (Augsburg, 1999), Springer, Berlin, 2001, pp. 128–161.

[Ful89] William Fulton, *Algebraic curves*, Advanced Book Classics, Addison-Wesley Publishing Company Advanced Book Program, Redwood City, CA, 1989, An introduction to algebraic geometry, Notes written with the collaboration of Richard Weiss, Reprint of 1969 original.

[Gan] Wee Tak Gan, *Lecture notes*, UCSD 2005.

[GM] Adriano Garsia and Gregg Musiker, *Basics on hyperelliptic curves over finite fields*, Mongraphies du LaCIM, To appear.

[Har77] Robin Hartshorne, *Algebraic geometry*, Springer-Verlag, New York, 1977, Graduate Texts in Mathematics, No. 52.

[Has34] H. Hasse, *Abstrakte Begründung der komplexen Multiplikation und Riemannsche Vermutung in Funktionenkörpern*, Abh. Math. Sem. Univ. Hamburg **10** (1934), 250–263.

[Hus04] Dale Husemöller, *Elliptic curves*, second ed., Graduate Texts in Mathematics, vol. 111, Springer-Verlag, New York, 2004, With appendices by Otto Forster, Ruth Lawrence and Stefan Theisen.

[IPS00] Mourad E. H. Ismail, Helmut Prodinger, and Dennis Stanton, *Schur's determinants and partition theorems*, Sém. Lothar. Combin. **44** (2000), Art. B44a, 10 pp. (electronic).

[JNR03] Brian Jacobson, Andrew Niedermaier, and Victor Reiner, *Critical groups for complete multipartite graphs and Cartesian products of complete graphs*, J. Graph Theory **44** (2003), no. 3, 231–250.

[Lan78] Serge Lang, *Elliptic curves: Diophantine analysis*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 231, Springer-Verlag, Berlin, 1978.

[Lan82] ______, *Introduction to algebraic and abelian functions*, second ed., Graduate Texts in Mathematics, vol. 89, Springer-Verlag, New York, 1982.

[Len96] H. W. Lenstra, Jr., *Complex multiplication structure of elliptic curves*, J. Number Theory **56** (1996), no. 2, 227–241.

[Lor00] Dino Lorenzini, *Arithmetical properties of Laplacians of graphs*, Linear and Multilinear Algebra **47** (2000), no. 4, 281–306.

[LP01] Matthieu Latapy and Ha Duong Phan, *The lattice structure of chip firing games and related models*, Phys. D **155** (2001), no. 1-2, 69–82.

[LWW04] Nicholas A. Loehr, Gregory S. Warrington, and Herbert S. Wilf, *The combinatorics of a three-line circulant determinant*, Israel J. Math. **143** (2004), 141–156.

[Mac60] Percy A. MacMahon, *Combinatory analysis*, Two volumes (bound as one), Chelsea Publishing Co., New York, 1960.

[Mac95] I. G. Macdonald, *Symmetric functions and Hall polynomials*, second ed., Oxford Mathematical Monographs, The Clarendon Press Oxford University Press, New York, 1995, With contributions by A. Zelevinsky, Oxford Science Publications.

[Max06] Molly Maxwell, *Enumerating bases of self-dual matroids*, 2006.

[Mer05] Criel Merino, *The chip-firing game*, Discrete Math. **302** (2005), no. 1-3, 188–210.

[Mil06] J. S. Milne, *Elliptic curves*, BookSurge Publishers, Charleston, SC, 2006.

[Mor91] Carlos Moreno, *Algebraic curves over finite fields*, Cambridge Tracts in Mathematics, vol. 97, Cambridge University Press, Cambridge, 1991.

[MOV93] Alfred J. Menezes, Tatsuaki Okamoto, and Scott A. Vanstone, *Reducing elliptic curve logarithms to logarithms in a finite field*, IEEE Trans. Inform. Theory **39** (1993), no. 5, 1639–1646.

[MP07] G. Musiker and J. Propp, *Combinatorial Interpretations for Rank-Two Cluster Algebras of Affine Type*, the electronic journal of combinatorics **14** (2007), no. R15, 1.

[Mye71] B. Myers, *Number of spanning trees in a wheel*, Circuits and Systems, IEEE Transactions on [legacy, pre-1988] **18** (1971), no. 2, 280–282.

[Pro] Jim Propp, *Somos sequence website*, `http://www.math.wisc.edu/~propp/somos.html`.

[Reu95] Christophe Reutenauer, *On symmetric functions related to Witt vectors and the free Lie algebra*, Adv. Math. **110** (1995), no. 2, 234–246.

[Reu97] ______, **N**-*rationality of zeta functions*, Adv. in Appl. Math. **18** (1997), no. 1, 1–17.

[Sil92] Joseph H. Silverman, *The arithmetic of elliptic curves*, Graduate Texts in Mathematics, vol. 106, Springer-Verlag, New York, 1992, Corrected reprint of the 1986 original.

[Slo] N.J.A. Sloane, *The on-line encyclopedia of integer sequences*, `http://www.research.att.com/∼ njas/sequences/index.html`.

[Sta73] H. M. Stark, *On the Riemann hypothesis in hyperelliptic function fields*, Analytic number theory (Proc. Sympos. Pure Math., Vol. XXIV, St. Louis Univ., St. Louis, Mo., 1972), Amer. Math. Soc., Providence, R.I., 1973, pp. 285–302.

[Sta97] Richard P. Stanley, *Enumerative combinatorics. Vol. 1*, Cambridge Studies in Advanced Mathematics, vol. 49, Cambridge University Press, Cambridge, 1997, With a foreword by Gian-Carlo Rota, Corrected reprint of the 1986 original.

[Sta99] Richard P. Stanley, *Enumerative combinatorics. Vol. 2*, Cambridge Studies in Advanced Mathematics, vol. 62, Cambridge University Press, Cambridge, 1999, With a foreword by Gian-Carlo Rota and appendix 1 by Sergey Fomin.

[Swa] C. Swart, *Elliptic curves and related sequences*, Ph.D. thesis, PhD Thesis, Royal Holloway and Bedford New College, University of London, 2003.

[VDPS06] A.J. Van Der Poorten and C.S. Swart, *Recurrence relations for elliptic sequences: every Somos* 4 *is a Somos*$_k$, Bulletin of the London Mathematical Society **38** (2006), no. 04, 546–554.

[Wag00] D.G. Wagner, *The critical group of a directed graph*, `arXiv:math.CO/0010241` (2000).

[War48] Morgan Ward, *Memoir on elliptic divisibility sequences*, Amer. J. Math. **70** (1948), 31–74.

[Was03] Lawrence C. Washington, *Elliptic curves*, Discrete Mathematics and its Applications (Boca Raton), Chapman & Hall/CRC, Boca Raton, FL, 2003, Number theory and cryptography.

[Wei48] André Weil, *Sur les courbes algébriques et les variétés qui s'en déduisent*, Actualités Sci. Ind., no. 1041 = Publ. Inst. Math. Univ. Strasbourg **7** (1945), Hermann et Cie., Paris, 1948.

[Wit01] Christian Wittmann, *Group structure of elliptic curves over finite fields*, J. Number Theory **88** (2001), no. 2, 335–344.

[Zel07] A. Zelevinsky, *Semicanonical basis generators of the cluster algebra of type*, the electronic journal of combinatorics **14** (2007), no. 4, 1.

[ZYG05] Yuanping Zhang, Xuerong Yong, and Mordecai J. Golin, *Chebyshev polynomials and spanning tree formulas for circulant and related graphs*, Discrete Math. **298** (2005), no. 1-3, 334–364.