Formal Power Series and Algebraic Combinatorics
Séries Formelles et Combinatoire Algébrique
San Diego, California 2006

Combinatorial aspects of elliptic curves

Gregg Musiker

Abstract. Given an elliptic curve C, we study here the number $N_k = \#C(\mathbb{F}_{q^k})$ of points of C over the finite field $\mathbb{F}_{q^k}$. We obtain two combinatorial formulas for N_k. 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.

Résumé. Étant donnée une courbe elliptique C on étudie le nombre $N_k = \#C(\mathbb{F}_{q^k})$ de points de C dans le corps fini $\mathbb{F}_{q^k}$. On obtient deux formules combinatoires pour N_k. En particulier on démontre que $N_k = -\mathcal{W}_k(q,t)|_{t=-N_1}$ oú $\mathcal{W}_k(q,t)$ est une (q,t)-extension du nombre des arbres recouvrants du graphe roue.

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 of 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 looking at field extensions: the infinite sequence of cardinalities is only dependent on a finite set of data. Specifically the number of points over $\mathbb{F}_q$, $\mathbb{F}_{q^2}$, $\dots$, and $\mathbb{F}_{q^g}$ will be sufficient data to determine the number of points on a genus g algebraic curve over any other algebraic field extension. This observation begs the question of how the points over higher field extensions correspond to points over the first g extensions. To see this more clearly, we specialize to the case of elliptic curves, where $g = 1$. Letting N_k equal the number of points on C over $\mathbb{F}_{q^k}$, we find a polynomial formula for N_k in terms of q and N_1. Moreover, the coefficients in our formula have a combinatorial interpretation related to spanning trees of the wheel graph.

2. The Zeta Function of a Curve

The zeta function of a curve C is defined to be the exponential generating function

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

A result due to Weil [**7**] is that the zeta function of an elliptic curve, in fact for any curve, $Z(C,T)$ is rational, and moreover can be expressed as

$$Z(C,T) = \frac{(1-\alpha_1 T)(1-\alpha_2 T)}{(1-T)(1-qT)} = \frac{1-(\alpha_1+\alpha_2)T + \alpha_1\alpha_2 T^2}{(1-T)(1-qT)}.$$

The inverse roots α_1 and α_2 satisfy a functional equation which reduces to

$$\alpha_1\alpha_2 = q$$

in the elliptic curve case.

2000 *Mathematics Subject Classification.* Primary 11G07; Secondary 05C05.
Key words and phrases. elliptic curves, finite fields, zeta functions, spanning trees, Lucas numbers, symmetric functions.
The author thanks the National Science Foundation for their generous support.

Among other things, rationality and the functional equation implies that $N_k = 1 + q^k - \alpha_1^k - \alpha_2^k$, which can be written in plethystic notation as $p_k[1 + q - \alpha_1 - \alpha_2]$. As a special case,

$$\alpha_1 + \alpha_2 = 1 + q - N_1.$$

Hence we can rewrite the Zeta function $Z(C, T)$ totally in terms of q and N_1, hence all the N_k's are actually dependent on these two quantities. The first few formulas are given below.

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

This data gives rise to our first observation.

Theorem 2.1.

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

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

We will prove this in the course of the deriviations in Section 3. Also see [**3**] for a direct proof. This result motivates the combinatorial question: what are the objects that the family of polynomials, $\{P_{i,k}\}$ enumerate?

3. The Lucas Numbers and a (q, t)-analogue

Definition 3.1. We define the $(q, t)-$**Lucas numbers** to be a sequence of polynomials in variables q and t such that $L_n(q, t)$ is defined as

$$L_n(q, t) = \sum_{S \subseteq \{1,2,\ldots,n\} \ : \ S \cap S_1^{(n)} = \phi} q^{\#\text{ even elements in } S}\, t^{\lfloor \frac{n}{2} \rfloor - \#S}. \tag{3.1}$$

Here $S_1^{(n)}$ is the circular shift of set S modulo n, i.e. element $x \in S_1$ if and only if $x - 1 \pmod{n} \in S$. In other words, the 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. [**1**], 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 3.2.

$$1 + q^k - N_k = L_{2k}(q, t)\Big|_{t=-N_1} \tag{3.2}$$

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 (3.1) 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 (3.2) satisfy the recursion

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

for $k \geq 1$.

Proposition 3.1. *For the $(q,t)-$Lucas Numbers $L_k(q,t)$ defined as above,*

$$L_{2k+2}(q,t) = (1+q+t)L_{2k}(q,t) - qL_{2k-2}(q,t). \tag{3.3}$$

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 (3.3) for the L_{2k}'s translates into

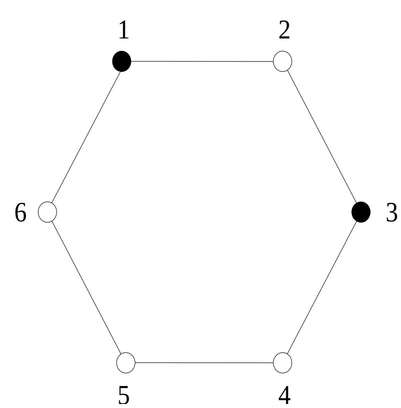
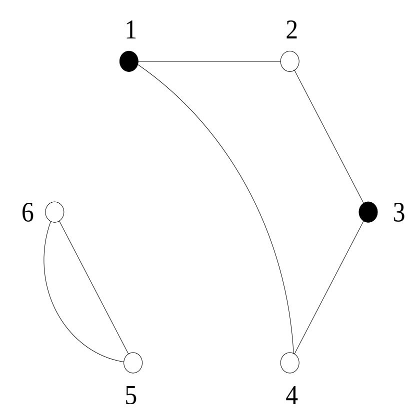
$$\tilde{L}_{2k+2}(q,t) = (1+t+qt)\tilde{L}_{2k}(q,t) - qt^2\tilde{L}_{2k-2}(q,t) \tag{3.4}$$

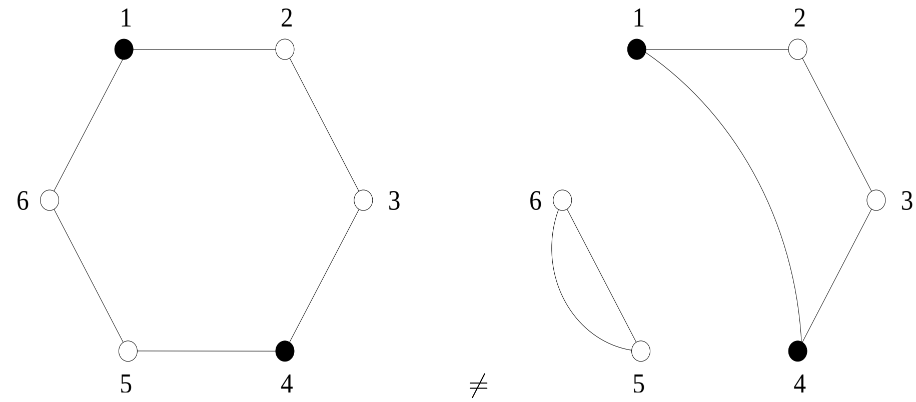
for the $\tilde{L}_{2k}$'s. The $\tilde{L}_{2k}$'s happen to have a nice combinatorial interpretation also, namely

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

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. For example, if $k=2$:

Decomposable  = 



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 3.3. The (twisted) $(q,t)-$Fibonacci polynomials, denoted as $\tilde{F}_n(q,t)$, will be defined as

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

The summands here are subsets of $\{1,2,\dots,k-1\}$ 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. 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 3.4. *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. □

Lemma 3.5. *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 the 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 3.6. *The difference of the quantity referred to in Lemma 3.5 from the quantity in Lemma 3.4 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

$$qt^2\tilde{L}_{2k-2}(q,t) = qt^2\tilde{F}_{2k-2}(q,t) + q^2t^3\tilde{F}_{2k-4}(q,t) \tag{3.5}$$

since

$$\tilde{F}_{2k-2}(q,t) = qt\tilde{F}_{2k-4}(q,t) + \tilde{F}_{2k-3}(q,t). \tag{3.6}$$

Note that identity (3.6) simply comes from the fact that the $(2k-2)$nd bead can be black or white. Finally we prove (3.5) 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 3.2. *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

$$\alpha_1^{k+1} + \alpha_2^{k+1} = (\alpha_1 + \alpha_2)(\alpha_1^k + \alpha_2^k) - q(\alpha_1^{k-1} + \alpha_2^{k-1}). \tag{3.7}$$

Noting that $q = \alpha_1\alpha_2$ we obtain this Proposition after expanding out algebraically the right-hand-side of (3.7). □

With the proof of Proposition 3.1 and 3.2, we have proven Theorem 3.2.

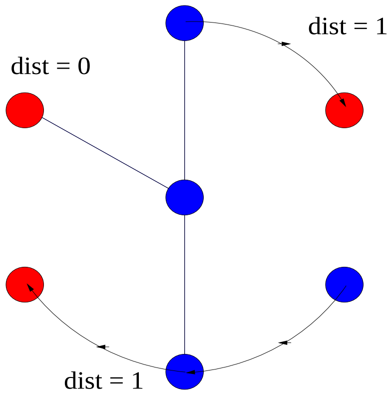
4. $(q,t)-$Wheel Numbers

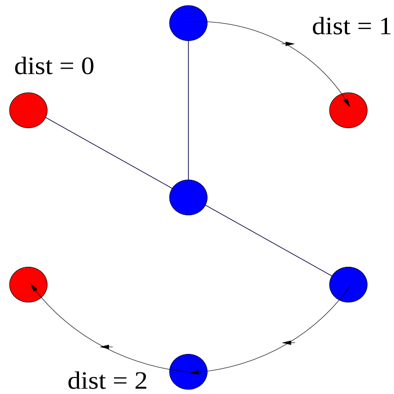
Given that we found the Lucas numbers are related to the polynomial formulas $N_k(q, N_1)$, a natural question concerns how alternative interpretations of the Lucas numbers can help us better understand N_k. As noted in [**1**], [**4**], and [**5**, Seq. A004146], the sequence $\{L_{2n} - 2\}$ counts 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. This combinatorial interpretation motivates the following definition.

Definition 4.1.

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

Here the exponent of t counts the number of edges emanating from the central vertex, and the exponent of q requires further explanation. 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, each spoke will intersect exactly one arc. (An isolated vertex is considered to be an arc of length 1.) 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 $q^{\text{ number of edges between the spoke and the tail}}$. 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 definition actually provides exactly the generating function that we desired, namely we have

Theorem 4.2 (Main Theorem).

$$N_k = -\mathcal{W}_k(q,t)\big|_{t=-N_1}$$

for all $k \geq 1$.

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

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

We will prove this Theorem in two different ways. The first method will utilize Theorem 3.2 and an analogue of the bijection given in [**1**] 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 bases 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 p_k into a linear combination of h_λ's, as described in Section 6.

5. First Proof: Bijective

There is a simple bijection between subsets of $[2n]$ 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 4.2. The bijection is as follows:

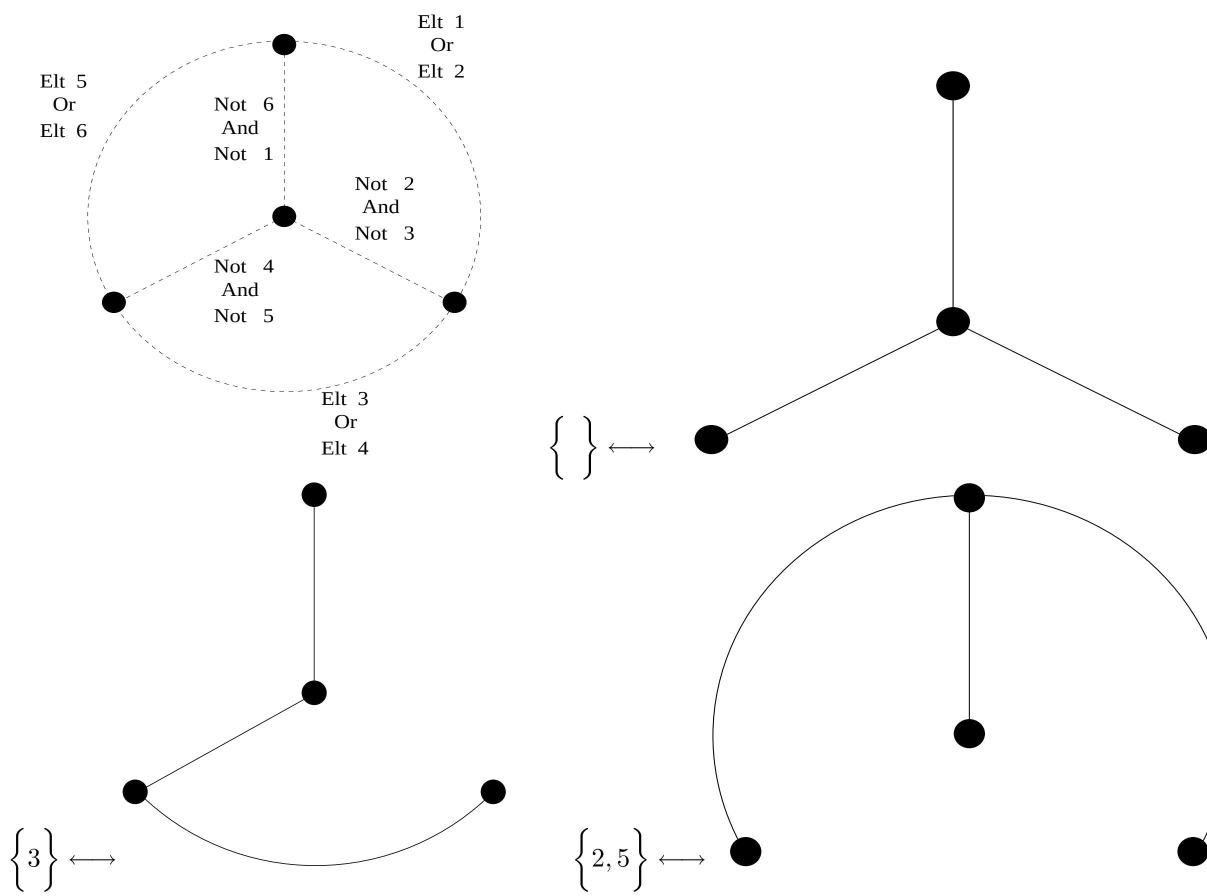
Given a subset S of the set $\{1, 2, \ldots, 2n-1, 2n\}$ with no circularly consecutive elements, we define the corresponding spanning tree T_S of W_n (with the correct q and t weight) in the following way:

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 5.1. *Given this construction, T_S is in fact a spanning tree of W_n and further, tree T_S has the same $q-$ 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 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, $0 \leq k \leq n-1$ and thus 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 $S' = S \cup \{\text{odd}\}$ 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 $S'' = S \cup \{\text{even}\}$ 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. $\square$

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 4.2.

6. Brick Tabloids and Symmetric Function Expansions

Recall that we wrote N_k plethystically as $p_k[1+q-\alpha_1-\alpha_2]$. One advantage of plethystic notation is that we can exploit the following symmetric function identity [**6**, pg. 21]:

$$\sum_{n=0}^{\infty} h_n T^n = \prod_{k\in\mathcal{I}} \frac{1}{1-t_kT} = \exp\left(\sum_{n=1} p_n \frac{T^n}{n}\right) \tag{6.1}$$

where h_n and p_n are symmetric functions in the variables in $\mathcal{I}$. We note that $Z(C,T)$ resembles the right-hand-side of this identity, and consequently, if we had written $Z(C,T)$ as an ordinary power series

$$Z(C,T) = \sum_{k\geq 0} H_k T^k$$

we obtain that $H_k = h_k[1+q-\alpha_1-\alpha_2]$, where h_k denotes the kth homogeneous symmetric function.

REMARK 6.1. In fact H_k has an algebraic geometric interpretation also, just as the N_k's did. H_k equals the number of positive divisors of degree k on curve C.

For a general curve we can thus, by plethysm, write cardinalities N_k in terms of H_1 through H_k, using the same coefficients as those that appear in the expansion of p_k in terms of h_1 through h_k:

$$N_k = \sum_{\lambda\vdash k} c_\lambda H_{\lambda_1} H_{\lambda_2} \cdots H_{\lambda_{|\lambda|}} \tag{6.2}$$

where the c_λ can be written down concisely as

$$c_\lambda = (-1)^{l(\lambda)-1} \frac{k}{l(\lambda)} \binom{l(\lambda)}{d_1, d_2, \ldots, d_k} \tag{6.3}$$

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 Egecioglu and Remmel's interpretation involving weighted brick tabloids [**2**]. We will give another proof of this, involving a *possibly new* combinatorial interpretation for these coefficients, further on, in Section 7.

A brick tabloid [**2**] 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. We 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 6.1 (Egecioglu-Remmel 1991, [**2**]).

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

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 divide these tabloids into classes based on the size of the last brick, we obtain, by counting the number of rearrangements, that there are

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

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 (by abusing multinomial notation slightly)

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

Thus after comparing signs, we obtain that c_λ equals exactly the desired expression.

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
$$Z(C,T) = \frac{1 - (1 + q - N_1)T + qT^2}{(1-T)(1-qT)}$$
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 (6.2), and using (6.3), we get polynomial formulas for N_k in terms of q and N_1, which in fact are an alternative expression for the formulas found in section 2.

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

Thus using these alternative expressions for N_k, we have that Theorem 4.2 is equivalent to the statement

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

7. Second Proof: Via Symmetric Functions

For our second proof of Theorem 4.2, we start with the observation that the sequence of lengths of all disjoint arcs on the rim of W_n corresponds to a partition of n. We will construct a spanning tree of W_n from the following choices:

First we choose a partition $\lambda = 1^{d_1} 2^{d_2} \cdots k^{d_k}$ of n. 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 $l(\lambda)$ arcs on the circle. After picking one to start with, without loss of generality since we are on a circle, we have

$$\frac{1}{l(\lambda)} \binom{l(\lambda)}{d_1,\ d_2,\ \dots\ d_k}$$

choices for such an arrangement.

Third, we pick which vertex w_i of the rim to start with. There are n such choices.

Fourth, we pick where the $l(\lambda)$ spokes actually intersect the arcs. There will be $|\text{arc}|$ choices for each arc, and the q−weight of this sum will be $(1 + q + q^2 + \cdots + q^{|\text{arc}|})$ for each arc.

Summing up all the possibilities yields

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

As noted in Section 6, these coefficients are exactly the correct expansion coefficients by identities (6.1), (6.3), and plethysm. Thus we have given a second proof of Theorem 4.2.

Remark 7.1. We note that in the course of this second proof we have obtained a combinatorial interpretation for the c_λ's that is distinct from the one given in Egecioglu and Remmel's paper [**2**]. In particular this interpretation does not require weighted counting, only signed counting. Instead of defining c_λ as $(-1)^{l(\lambda)-1} w(B_{\lambda,(k)})$, we could define it as

$$(-1)^{l(\lambda)-1} |CB_{\lambda,(k)}|$$

where we define a new combinatorial class of **circular brick tabloids** which we denote as $CB_{\lambda,\mu}$. We define this for the case of $\mu = (k)$ just as we defined the usual brick tabloids, except we are not filling a $k \times 1$ rectangle, but are filling an annulus of circumference k and width 1 with curved bricks of sizes designated by λ. In this way we mimic our construction of the spanning trees.

Additionally, by using the fact that the power symmetric functions are multiplicative, i.e. $p_\lambda = p_{\lambda_1} p_{\lambda_2} \cdots p_{\lambda_r}$, we are able to generalize our definition of circular brick tabloids to allow μ to be any partition. We simply let λ designate what collection of bricks we have to use, and μ determines the filling: we are trying to fill $l(\mu)$ concentric circles where each circle has μ_i spaces. To summarize,

$$p_\mu = \sum_\lambda (-1)^{l(\lambda)-l(\mu)} |CB_{\lambda,\mu}| h_\lambda.$$

Consequently, all identities of [**2**] now involve *cardinalities* of $B_{\lambda,\mu}$, $OB_{\lambda,\mu}$ (Ordered Brick Tabloids), or $CB_{\lambda,\mu}$ and signs depending on $l(\lambda)$ and $l(\mu)$, with no additional weightings needed.

8. Conclusion

The new combinatorial formula for N_k presented in this write-up appears fruitful. It leads one to ask how spanning trees of the wheel graph are related to points on elliptic curves. For instance, is there an involution on (weighted) spanning trees whose fixed points enumerate points on $C(\mathbb{F}_{q^k})$? The fact that the Lucas numbers also enter the picture is also exciting since the Fibonacci numbers and Lucas numbers have so many different combinatorial interpretations, and there is such an extensive literature about them. Perhaps these combinatorial interpretations will lend insight into why N_k depends only on the finite data of N_1 and q for an elliptic curve, and how we can associate points over higher extension fields to points on $C(\mathbb{F}_q)$.

9. Ackowledgements

The author would like to thank Adriano Garsia for his support and many helpful conversations.

References

[1] A. Benjamin and C. Yerger, Combinatorial Interpretations of Spanning Tree Identities, *Bulletin of the Institute for Combinatorics and its Applications*, to appear.

[2] O. Egecioglu and J. Remmel, Brick Tabloids and the Connection Matrices Between Bases of Symmetric Functions, *Disc. Appl. Math.*, **34** (1991), 107-120.

[3] A. Garsia and G. Musiker, *Basics on Hyperelliptic Curves over Finite Fields*, in progress.

[4] B. R. Myers, Number of Spanning Trees in a Wheel, *IEEE Trans. Circuit Theory*, **18** (1971), 280-282.

[5] N. J. A. Sloane, *The On-Line Encyclopedia of Integer Sequences*, http://www.research.att.com/~njas/sequences/index.html.

[6] R. P. Stanley, *Enumerative Combinatorics Vol. 2, volume 62 of Cambridge Studies in Advanced Mathematics*, Cambridge University Press, Cambridge (1999).

[7] A. Weil, *Sur les Courbes Algébriques et les Variétés qui s'en Déduisent*, Hermann, Paris (1948).

Department of Mathematics, UCSD, San Diego, USA, 92037
E-mail address: gmusiker@math.ucsd.edu

Séminaire Lotharingien de Combinatoire **56** (2007), Article B56f

COMBINATORIAL ASPECTS OF ELLIPTIC CURVES

GREGG MUSIKER

ABSTRACT. Given an elliptic curve C, we study here $N_k = \#C(\mathbb{F}_{q^k})$, the number of points of C 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 $N_k = -\mathcal{W}_k(q, -N_1)$, where $\mathcal{W}_k(q,t)$ is a (q,t)-analogue of 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.

CONTENTS

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 of an algebraic equation is a

This work was supported by the NSF, grant DMS-0500557.

discrete subset, but since they satisfy a certain type of algebraic equation this imposes a lot of extra structure beneath the surface. One of the ways to detect this additional structure is by looking at field extensions: the infinite sequence of cardinalities is only dependent on a finite set of data. Specifically the number of points over $\mathbb{F}_q$, $\mathbb{F}_{q^2}$, $\ldots$, and $\mathbb{F}_{q^g}$ will be sufficient data to determine the number of points on a genus g algebraic curve over any other algebraic field extension. This observation motivates the question of how the points over higher field extensions correspond to points over the first g extensions.

To see this more clearly, we specialize to the case of elliptic curves, where $g = 1$, and examine the expressions for N_k, the number of points on C over $\mathbb{F}_{q^k}$, as functions of q and N_1. It follows from the well-known rationality of the zeta function that

$$N_k(q, N_1) = 1 + q^k - \alpha_1^k - \alpha_2^k, \tag{1}$$

where α_1 and α_2 are the two roots of the quadratic $1-(1+q-N_1)T+qT^2$. Additionally, we observe, see Theorem 1, that

(2) $N_k(q, N_1)$ are integral polynomials whose coefficients alternate in sign.

In this paper, we use formulas arising from (1) and (2) to connect elliptic curves to several different areas of combinatorics. Specifically, (1) implies that the family of polynomials $1 + q^k - N_k$ are Chebyshev polynomials of the first kind, a well-studied example of orthogonal polynomials. In Section 4, we describe this perspective in further detail. Alternatively, we can interpret statement (1) as the plethystic expression $N_k = p_k[1+q-\alpha_1-\alpha_2]$, where the p_k's are the power symmetric functions. In summary, we exploit both the fields of orthogonal polynomials and symmetric functions to illustrate numerous identities involving the N_k's.

Moreover, we find that the polynomial expressions for N_k due to (2) are related to a (q, t)-deformation of the Lucas numbers (Theorem 2), and also lead to a combinatorial interpretation involving spanning trees of the wheel graph (Theorem 3). Thus the aforementioned identities also indicate properties of the Lucas numbers and spanning trees as well.

Using these new combinatorial interpretations for N_k, we develop further properties of this sequence, obtaining determinantal formulas (Theorem 5), as well as formulas involving a certain bivariate version of the Fibonacci polynomials (Theorem 4). Another surprising by-product of our analysis is a factorization of N_k into a new sequence of polynomials, which we refer to as elliptic cyclotomic polynomials. Both of these families of polynomials are interesting in their own right and have numerous properties which justify their names. We give a geometric interpretation of the elliptic cyclotomic polynomials as Theorem 7 and close with some combinatorial identities involving this new family of expressions.

2. N_k AS AN ALTERNATING SUM

The zeta function of a curve C is defined to be the exponential generating function

$$Z(C,T) = \exp\left(\sum_{k\geq 1} N_k \frac{T^k}{k}\right). \tag{3}$$

A result due to Weil [22] is that the zeta function of a curve is rational with specific formula given as

$$Z(C,T) = \frac{(1-\alpha_1 T)(1-\alpha_2 T)\cdots(1-\alpha_{2g}T)}{(1-T)(1-qT)}. \tag{4}$$

Here g is the genus of curve C, and the numerator is sometimes written as $L(C,T)$, a degree $2g$ polynomial with integer coefficients. Moreover when E is an elliptic curve, $Z(E,T)$ can be expressed as

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

The zeta function of a curve also satisfies a functional equation which in the elliptic case is simply equivalent to

$$\alpha_1\alpha_2 = q.$$

Among other things, (3) and (4) imply that $N_k = 1+q^k-\alpha_1^k-\alpha_2^k-\cdots-\alpha_{2g}^k$, which can be written in plethystic notation as $p_k[1+q-\alpha_1-\alpha_2]$. We describe symmetric functions and plethystic notation in more depth in Section 3. In the case that E is a curve of genus one and $k=1$ we get

$$\alpha_1+\alpha_2 = 1+q-N_1.$$

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. The first few formulas are given below:

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

This data gives rise to the following observation of Adriano Garsia.

Theorem 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 [7, Chap. 7]. This result motivates the combinatorial question: what are the objects that the family of polynomials, $\{P_{i,k}\}$, enumerate? We answer this question in due course in multiple ways, thus providing an alternate, combinatorial, proof of Theorem 1.

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

Definition 1. 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 polynomials** to be the sequence of polynomials in variables q and t

$$L_n(q,t) = \sum_{S \subseteq \{1,2,\ldots,n\} \,:\, S \cap S_1^{(n)} = \emptyset} q^{\#\text{even elements in } S} \quad t^{\lfloor \frac{n}{2} \rfloor - \#S}. \tag{5}$$

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

These polynomials are a generalization of the sequence of Lucas polynomials 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. [1], 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 prove the following theorem, which relates our newly defined (q,t)-Lucas polynomials to the polynomials of interest, namely the N_k's.

Theorem 2. *We have*

$$1 + q^k - N_k = L_{2k}(q, -N_1) \tag{6}$$

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

$$L_2(q,t) = 1 + q + t$$

and

$$L_4(q,t) = 1 + q^2 + (2q+2)t + t^2,$$

we have proven that the initial conditions agree. Note that the sets of (5) yielding the terms of these sums are respectively

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

It remains to prove that both sides of (6) satisfy the recursion

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

for $k \geq 1$.

Proposition 1. *For the (q,t)-Lucas polynomials $L_k(q,t)$ defined as above,*

$$L_{2k+2}(q,t) = (1+q+t)L_{2k}(q,t) - qL_{2k-2}(q,t). \tag{7}$$

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 (7) for the L_{2k}'s translates into

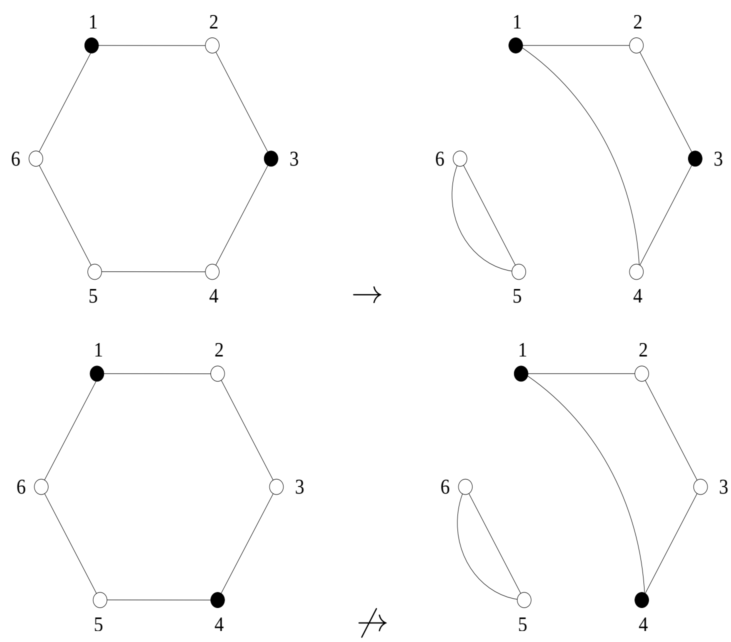
$$\tilde{L}_{2k+2}(q,t) = (1+t+qt)\tilde{L}_{2k}(q,t) - qt^2\tilde{L}_{2k-2}(q,t) \tag{8}$$

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 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. For example, if $k=2$:



Decomposable

Not Decomposable

In these figures, 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 are Fibonacci numbers.

Definition 2. 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. 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 1. *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. □

Lemma 2. *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 3. *The difference of the quantity referred to in Lemma 2 from the quantity in Lemma 1 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

$$qt^2\tilde{L}_{2k-2}(q,t) = qt^2\tilde{F}_{2k-2}(q,t) + q^2t^3\tilde{F}_{2k-4}(q,t), \tag{9}$$

since

$$\tilde{F}_{2k-2}(q,t) = qt\tilde{F}_{2k-4}(q,t) + \tilde{F}_{2k-3}(q,t). \tag{10}$$

Note that identity (10) simply comes from the fact that the $(2k-2)$nd bead can be black or white. Finally we prove (9) 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 2. *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

$$\alpha_1^{k+1} + \alpha_2^{k+1} = (\alpha_1 + \alpha_2)(\alpha_1^k + \alpha_2^k) - q(\alpha_1^{k-1} + \alpha_2^{k-1}). \tag{11}$$

Noting that $q = \alpha_1\alpha_2$ we obtain this proposition after expanding out algebraically the right-hand-side of (11). □

With the proof of Propositions 1 and 2, we have proven Theorem 2.

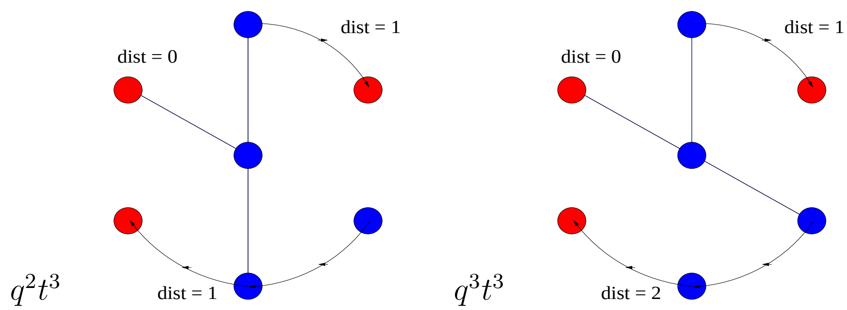
2.2. (q,t)**-Wheel polynomials.** Given that the Lucas numbers are related to the polynomial formulas $N_k(q, N_1)$, a natural question concerns how alternative interpretations of the Lucas numbers can help us better understand N_k. As noted in [1], [14], and [18, Seq. A004146], the sequence $\{L_{2n} - 2\}$ 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.

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 intersects 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 are 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 3.

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

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



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

Theorem 3.

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

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 prove this theorem in two different ways. The first method utilizes Theorem 2 and an analogue of the bijection given in [1] which relates perfect and imperfect matchings of the circle of length $2k$ and spanning trees of W_k. Our second proof uses the observation that we can categorize the spanning trees based on the sizes of the various connected arcs on the rims. Since this categorization corresponds to partitions, this method exploits formulas for decomposing power symmetric function p_k into a linear combination of h_λ's, as described in Section 2.4.

2.3. **First proof of Theorem 3: Bijective.** There is a simple bijection between subsets of $\{1, 2, \ldots, 2n\}$ with size at most $n-1$ as well as no two elements circularly consecutive and spanning trees of the wheel graph W_n. We use this bijection to give our first proof of Theorem 3. The bijection is as follows:

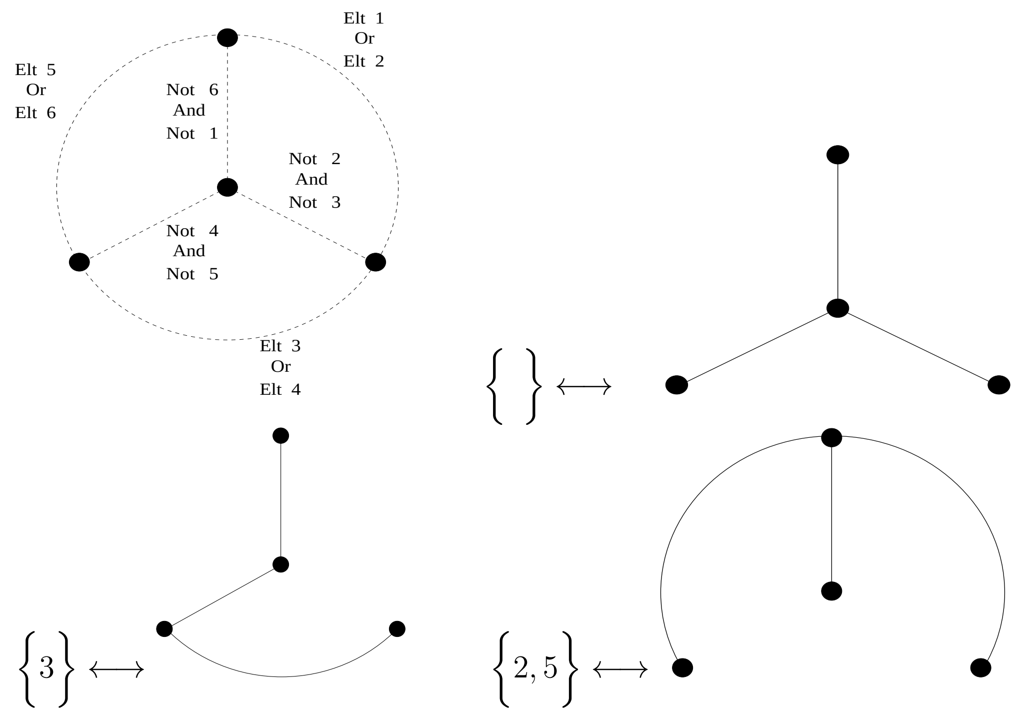
Given a subset S of the set $\{1, 2, \ldots, 2n-1, 2n\}$ with no circularly consecutive elements, we define the corresponding spanning tree T_S of W_n (with the correct q and t weight) in the following way:

1) We 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 exclude the two subsets which consist of all the odds or all the evens from this bijection. Thus we only look 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 3. *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 are $(n-k)$ 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 T_S consists 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 happens 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 3.

2.4. **Second proof of Theorem 3: Via generating function identities.** For our second proof of Theorem 3, we consider writing the zeta function as an ordinary generating function instead, i.e.,

$$Z(C,T) = 1 + \sum_{k\geq 1} H_k T^k. \tag{12}$$

In such a form, the H_k's are positive integers which enumerate the number of effective $C(\mathbb{F}_q)$-divisors of degree k, as noted in several places, such as [13].

Proposition 4.

$$N_k = \sum_{\lambda \vdash k} (-1)^{l(\lambda)-1} \frac{k}{l(\lambda)} \binom{l(\lambda)}{d_1,\ d_2,\ \ldots\ d_m} \prod_{i=1}^{l(\lambda)} H_{\lambda_i}. \tag{13}$$

Proof. Comparing formulas (3) and (12) for $Z(C,T)$ and taking logarithms, we obtain

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

$$= \sum_{m\geq 1} \frac{(-1)^{m-1}\Big(\sum_{n=1}^{k} H_n T^n\Big)^m}{m}\Bigg|_{T^k}.$$

To obtain the coefficient of T^k in

$$\Big(H_1T + H_2T^2 + \cdots + H_kT^k\Big)^m, \tag{14}$$

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_jT^j in the ith term in product (14). Secondly, since the order of these terms does not matter, we include multinomial coefficients. Finally, multiplying through by k yields formula (13) for N_k. □

Remark 1. The same manipulations done above for the generating functions are analogous to identities which relate the power symmetric functions and homogeneous symmetric functions. See for example [5], [12], or [20, pg. 21]. This is no coincidence, and in particular the terminology of plethysm provides a rigorous connection between symmetric functions and the enumeration of points on curves. See Section 3 below, [7], or [15] for more details on plethysm and this connection.

Remark 2. The above algebraic reasoning can also be translated into a combinatorial description of how points on C over $\mathbb{F}_{q^k}$ can be enumerated using inclusion-exclusion, and points over smaller extension fields. See [15, Chap. 4] for more details.

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

$$Z(E,T) = \frac{1-(1+q-N_1)T+qT^2}{(1-T)(1-qT)} = 1 + \frac{N_1T}{(1-T)(1-qT)} \tag{15}$$

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 (13), we get polynomial formulas for N_k in terms of q and N_1

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

Consequently, Theorem 3 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 5.

$$\mathcal{W}_k = \sum_{\lambda\vdash k}\frac{k}{l(\lambda)}\binom{l(\lambda)}{d_1,\ d_2,\ \ldots\ d_k}\Bigg(\prod_{i=1}^{l(\lambda)}(1+q+q^2+\cdots+q^{\lambda_i-1})\Bigg)t^{l(\lambda)}. \tag{16}$$

Proof. We 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

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

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 are $|\mathrm{arc}|$ choices for each arc, and the q-weight of this sum is $(1+q+q^2+\cdots+q^{|\mathrm{arc}|})$ for each arc. Summing up all the possibilities yields (16) as desired. □

Thus we have given a second proof of Theorem 3.

3. 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 consists of two different manifestations of duality.

3.1. Duality between the symmetric functions h_k and e_k.

Given the usefulness of symmetric functions in discovering the identities described by Propositions 4 and 5, we now illustrate further applications of the plethystic view of the zeta function.

The symmetric functions that we utilize in this paper are the power symmetric functions p_k, the complete homogeneous symmetric functions h_k, and the elementary symmetric functions e_k. Given the alphabet $\{x_1, x_2, \dots, x_n\}$, each of these can be written as

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

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 $\{p_\alpha\}_\alpha$. The main example of this technique that we use is $N_k = p_k[1+q-\alpha_1-\alpha_2-\cdots-\alpha_{2g}]$ for a genus g curve.

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

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

where h_n, p_n, and e_n are symmetric functions in the variables $\{t_k\}_{k\in\mathcal{I}}$. [20, pgs. 21, 296] The zeta function $Z(C,T)$ is equal to all of these for a certain choice of $\{t_k\}_{k\in\mathcal{I}}$ and consequently, we get that

$$Z(C,T) = \frac{1}{\sum_{k\geq 0}(-1)^k E_k \cdot T^k}, \tag{17}$$

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

Remark 3. Like the N_k's and H_k's, the E_k's also have an algebraic geometric interpretation, namely E_k equals the *signed* number of positive divisors D of degree k on curve C *such that no prime divisor appears more than once in* D. This follows from the reciprocity between h_k and e_k which is analogous to the reciprocity between *choose* and *multi-choose*, i.e., choice with replacement.

Recall that in Section 2.1, 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. 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. *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.

The expansions for the first several E_k's, i.e., $F_{2k-1}(q,t)$'s, are given below.

$$E_1 = N_1,$$
$$E_2 = -(1+q)N_1 + N_1^2,$$
$$E_3 = (1+q+q^2)N_1 - (2+2q)N_1^2 + N_1^3,$$
$$E_4 = -(1+q+q^2+q^3)N_1 + (3+4q+3q^2)N_1^2 - (3+3q)N_1^3 + N_1^4,$$

$$E_5 = (1+q+q^2+q^3+q^4)N_1 - (4+6q+6q^2+4q^3)N_1^2 + (6+9q+6q^2)N_1^3 - (4+4q)N_1^4 + N_1^5.$$

Before proving Theorem 4 we develop two key propositions.

Proposition 6. *$F_{2n+1}(q,t) = (1+q+t)F_{2n-1}(q,t) - qF_{2n-3}(q,t)$ for $n \geq 2$.*

Proof. This follows the similar logic as the proof of Proposition 1 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 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. □

Proposition 7. *$(-1)^{n+1}E_{n+1} = (1+q-N_1)(-1)^nE_n - q(-1)^{n-1}E_{n-1}$ for $n \geq 2$.*

Proof. One can prove this via plethysm, but it also follows directly from the generating function for the E_n's which is given by

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

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.

Proof of Theorem 4. 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 6 and 7 show that both satisfy the same recurrence relations. Thus we have verified that

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

□

Remark 4. We can utilize plethysm and obtain results of a similar flavor to Proposition 7, for example see Lemma 4 below. With this result in mind, we obtain the following table of symmetric function e_k and h_k in terms of various alphabets.

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	$1+q+\cdots+q^k$	$(-1)^kE_{k+1}/N_1$

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 $T^2-(1+q)T+q$.

Lemma 4. *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 7. 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. □

We apply the above H_k–E_k (i.e., h_k–e_k) duality to 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 use (15) to 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

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

We summarize our results as the following dictionary between elliptic curves and spanning trees accordingly.

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

3.2. **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 the following identity.

Proposition 8. *We have*

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

Proof. We use the identities as above, and the fact that $\frac{1}{Z(E,T)} = \sum_{n\geq 0}(-1)^n E_n T^n$. Thus 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 $-N_1\frac{\partial}{\partial N_1}\Bigg(\sum_{k\geq 1}\frac{N_k}{k}T^k\Bigg)$. Rewriting the N_k's using the polynomial formulas of Theorem 1, we have

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

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

Proposition 8 can also be given a combinatorial proof by the following lemma which contrasts the circular nature of our combinatorial interpretation for the Lucas numbers with the linear nature of the Fibonacci numbers.

Lemma 5. *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*

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

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 (respectively S_2) directly describes the coefficient of q^jt^i in $L_{2k}(q,t)$ (respectively $F_{2k-1}(q,t)$).

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 elements 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. □

Using this relationship, we can derive formulas involving binomial coefficients for $P_{i,k}(q)$ using our combinatorial interpretation for the (q,t)-Lucas polynomials and (q,t)-Fibonacci polynomials.

Proposition 9. *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 [23, Theorem 2.2] or [16, 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 1. *We have*

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

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 5. From the proof in Section 2.4, 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 9 seems to give a combinatorial identity that seems interesting in its own right.

4. Factorizations of N_k

We now introduce a family of k-by-k matrices M_k which, for elliptic curves, yield a determinantal formula for N_k in terms of q and N_1.

Theorem 5. *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*

$$\begin{bmatrix} 1+q-N_1 & -1 & 0 & \dots & 0 & -q \\ -q & 1+q-N_1 & -1 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & -q & 1+q-N_1 & -1 & 0 \\ 0 & \dots & 0 & -q & 1+q-N_1 & -1 \\ -1 & 0 & \dots & 0 & -q & 1+q-N_1 \end{bmatrix}.$$

The sequence of integers $N_k = \#C(\mathbb{F}_{q^k})$ *satisfies the relation*

$$N_k = -\det M_k \text{ for all } k \geq 1.$$

We provide two proofs of this theorem, one which utilizes the three term recurrence from Section 2.1, and one which introduces a new sequence of polynomials which are interesting in their own right.

4.1. **Connection to orthogonal polynomials.** Recall from the zeta function of an elliptic curve, $Z(E,T)$, we derived a three term recurrence relation for the sequence $\{G_k = 1 + q^k - N_k\}$:

$$G_{k+1} = (1 + q - N_1)G_k - qG_{k-1}. \tag{18}$$

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

$$P_{k+1}(x) = (a_k x + b_k)P_k(x) + c_k P_{k-1}(x), \tag{19}$$

where a_k, b_k and c_k are constants that depend on $k \in \mathbb{N}$. Additionally, it is customary 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 (18) is a special case of recurrence (19), therefore $\{P_k(x)\}_{k=1}^{\infty} = \{(1 + q^k - N_k)(N_1)\}_{k=1}^{\infty}$ are a family of orthogonal polynomials. In particular, we plug in the following values for the a_k, b_k, and c_k's:

$$\begin{aligned} a_k &= -1 && \text{for } k \geq 0 \\ b_k &= 1+q && \text{for } k \geq 0, \\ c_1 &= -2q && \text{and} \\ c_k &= -q && \text{for } k \geq 2. \end{aligned}$$

(Note that we take c_1 to be $-2q$ since $G_0 = 1+q^0 - N_0 = 2$, but we wish to normalize so that $P_0(x) = 1$.)

In fact, the family $\{1+q^k-N_k\}_{k=1}^{\infty}$ can be described in terms of a classical sequence of orthogonal polynomials. Namely $T_k(x)$ denotes 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

$$\alpha = \cos\theta + i\sin\theta$$
$$\beta = \cos\theta - i\sin\theta.$$

Theorem 6. *Considering the $(1 + q^k - N_k)$'s as univariate polynomials in N_1 over the field $\mathbb{Q}(q)$, we obtain*

$$1 + q^k - N_k = 2q^{k/2} T_k\bigg((1 + q - N_1)/2q^{1/2}\bigg).$$

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 6 holds for $k = 1$, and additionally, by Proposition 1, the $\frac{1+q^k-N_k}{2q^{k/2}}$'s satisfy the same recurrence as the $T_k(x)$'s. Namely

$$\begin{aligned} \frac{1 + q^{k+1} - N_{k+1}}{2q^{(k+1)/2}} &= \frac{(1 + q - N_1)(1 + q^k - N_k) - q(1 + q^{k-1} - N_{k-1})}{2q^{(k+1)/2}} \\ &= 2\bigg(\frac{1 + q - N_1}{2q^{1/2}}\bigg)\bigg(\frac{1 + q^k - N_k}{2q^{k/2}}\bigg) - \bigg(\frac{1 + q^{k-1} - N_{k-1}}{2q^{(k-1)/2}}\bigg). \end{aligned}$$

□

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 [2] as described on page 33 of [4]: 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 $\mathcal{W}_k(1,N_1)$ (respectively $N_k(0,N_1)$ and $N_k(1,N_1)$) which are indeed linear transformations of x^k and $T_k(x)$ respectively.

Proposition 10. *We have*

$$N_k(0,N_1) = -(1-N_1)^k + 1, \tag{20}$$
$$N_k(1,N_1) = -2T_k(-N_1/2+1) + 2. \tag{21}$$

Proof. The coefficient of N_1^m in $\mathcal{W}_k(0,N_1)$ is the number of directed rooted 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. Use of 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 6. ☐

4.2. **First proof of Theorem 5: Using orthogonal polynomials.** As an application of Theorem 6, we use the theory of orthogonal polynomials to learn properties of the $(1+q^k-N_k)$'s. For example, 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.

Proposition 11. *We have*

$$1+q^k-N_k$$
$$= \det \begin{bmatrix} 1+q-N_1 & -2q & 0 & 0 & 0 & 0 \\ -1 & 1+q-N_1 & -q & 0 & 0 & 0 \\ 0 & -1 & 1+q-N_1 & -q & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & \cdots & 1+q-N_1 & -q \\ 0 & 0 & 0 & \cdots & -1 & 1+q-N_1 \end{bmatrix}.$$

We denote this matrix as M_k'.

Proof. Given a sequence of orthogonal polynomials satisfying $P_0(x) = 1$, $P_1(x) = a_0x + b_0$ and recurrence (19), we have the formula [10]

$$P_k(x) = \det \begin{bmatrix} a_0x+b_0 & c_1 & 0 & 0 & 0 & 0 \\ -1 & a_1x+b_1 & c_2 & 0 & 0 & 0 \\ 0 & -1 & a_2x+b_2 & c_3 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & \cdots & a_{k-2}x+b_{k-2} & c_k \\ 0 & 0 & 0 & \cdots & -1 & a_{k-1}x+b_{k-2} \end{bmatrix}.$$

Plugging in the a_i, b_i, and c_i's as in Section 4.1 yields the formula. □

Remark 6. Alternatively, we can use symmetric functions and the Newton Identities [20] to obtain these determinant identities, as described in [7, Chap. 7] or [15, Chap. 5].

We can prove Theorem 5 via Proposition 11 followed by an algebraic manipulation of matrix M_k. 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 (respectively $-q$)'s on the diagonal. Given that the sign of such a cyclic permutation is $(-1)^{k-1}$, we obtain $\det(B_k)+\det(C_k) = -q^k - 1$. 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

$$1 + q^k + \det(M_k) = \det(A_k) - q\det(A_{k-2}).$$

After transposing M'_k, by analogous methods we obtain

$$\det M'_k = \det(A_k) - q\det(A_{k-2})$$

and thus the desired formula $\det M_k = -N_k$.

4.3. **Second proof of Theorem 5: 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 refer to them henceforth as **elliptic cyclotomic polynomials**, denoted as $ECyc_d$.

Definition 5. 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 12. *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 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 [9, 17] (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 7. Alternatively, the integrality of the $ECyc_d$'s also 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 $e_1, e_2, \dots$. 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 [11] 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 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.

4.4. **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

$$N_k = \prod_{d|k} ECyc_d(q, N_1) \tag{22}$$

to define elliptic cyclotomic polynomials directly as

$$ECyc_k(q, N_1) = \prod_{d|k} {N_d}^{\mu(k/d)} \tag{23}$$

in addition to the alternative definition

$$ECyc_k(q, N_1) = \prod_{\substack{j=1 \\ \gcd(j,d)=1}}^{k} \left((1 + q - N_1) - q\omega_k^j - \omega_k^{k-j}\right). \tag{24}$$

In particular, $ECyc_1 = N_1$ and $ECyc_p = N_p/N_1$ if p is prime. To get a handle on $ECyc_k$ for k composite, we provide the following table for small values of k:

$$\begin{aligned}
ECyc_4 &= N_1^2 - (2 + 2q)N_1 + 2(1 + q^2) \\
ECyc_6 &= N_1^2 - (1 + q)N_1 + (1 - q + q^2) \\
ECyc_8 &= N_1^4 - (4 + 4q)N_1^3 + (6 + 8q + 6q^2)N_1^2 - (4 + 4q + 4q^2 + 4q^3)N_1 + 2(1 + q^4) \\
ECyc_9 &= N_1^6 - (6 + 6q)N_1^5 + (15 + 24q + 15q^2)N_1^4 - (21 + 36q + 36q^2 + 21q^3)N_1^3 \\
&\quad + (18 + 27q + 27q^2 + 27q^3 + 18q^4)N_1^2 \\
&\quad - (9 + 9q + 9q^2 + 9q^3 + 9q^4 + 9q^5)N_1 + 3(1 + q^3 + q^6) \\
ECyc_{10} &= N_1^4 - (3 + 3q)N_1^3 + (4 + 3q + 4q^2)N_1^2 \\
&\quad - (2 + q + q^2 + 2q^3)N_1 + (1 - q + q^2 - q^3 + q^4) \\
ECyc_{12} &= N_1^4 - (4 + 4q)N_1^3 + (5 + 8q + 5q^2)N_1^2 \\
&\quad - (2 + 2q + 2q^2 + 2q^3)N_1 + (1 - q^2 + q^4)
\end{aligned}$$

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 13. *We have*

$$ECyc_d|_{N_1=0} = C(d)Cyc_d(q) \tag{25}$$

$$ECyc_d|_{N_1=2q+2} = C'(d)Cyc_d(-q), \tag{26}$$

where $C(d)$ and $C'(d)$ are the functions from $\mathbb{Z}_{>0}$ to $\mathbb{Z}_{\geq 0}$ such that

$$C(d) = \begin{cases} 0 \text{ if } d = 1 \\ p \text{ if } d = p^k \text{ for } p \text{ prime} \\ 1 \text{ otherwise} \end{cases}$$

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 [18, 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 $k = p_1^{k'_1} p_2^{k'_2} \cdots p_r^{k'_r}$.

$$\begin{aligned} \frac{1 - q^k}{1 - q} &= 1 + q + q^2 + \cdots + q^{k-1} \\ &= \left(\prod_{j_1=1}^{k'_1} Cyc_{p_1^{j_1}}(q) \right) \left(\prod_{j_2=1}^{k'_2} Cyc_{p_2^{j_2}}(q) \right) \cdots \left(\prod_{j_r=1}^{k'_r} Cyc_{p_r^{j_r}}(q) \right) \\ &\quad \times \left(\prod_{d \text{ is another divisor of } k} Cyc_d(q) \right). \end{aligned}$$

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 (26) 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 [18, Seq. A020513]. □

Proposition 14. *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 7, 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 $\alpha_1^m\alpha_2^0$ (respectively $\alpha_1^0\alpha_2^m$) 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 13 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 notes 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 13. More precisely, the author finds the following empirical evidence for such a claim.

Proposition 15. *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, if we write

$$ECyc_{105}(q, N_1) = Cyc_{105}(1) \cdot Cyc_{105}(q) + \sum_{i=1}^{48} (-1)^i Q_{i,105}(q) N_1^i,$$

where the $Q_{i,105}(q)$'s are univariate polynomials with integer coefficients, then $Q_{2,105}(q)$ through $Q_{48,105}(q)$ indeed have *positive* integer coefficients as expected. However the first univariate polynomial, i.e., the coefficient of $-N_1$ is

$$\begin{aligned} Q_{1,105}(q) = {} & 24q^{47} + 47q^{46} + 69q^{45} + 69q^{44} + 69q^{43} + 50q^{42} + 32q^{41} \\ & - 2q^{40} - 18q^{39} - 33q^{38} - 33q^{37} - 33q^{36} - 21q^{35} - 10q^{34} \\ & + 9q^{32} + 17q^{31} + 24q^{30} + 24q^{29} + 24q^{28} + 20q^{27} + 20q^{26} + 18q^{25} + 18q^{24} \\ & + 18q^{23} + 18q^{22} + 20q^{21} + 20q^{20} + 24q^{19} + 24q^{18} + 24q^{17} + 17q^{16} + 9q^{15} \end{aligned}$$

$$- 10q^{13} - 21q^{12} - 33q^{11} - 33q^{10} - 33q^9 - 18q^8 - 2q^7$$
$$+ 32q^6 + 50q^5 + 69q^4 + 69q^3 + 69q^2 + 47q + 24.$$

Note that there are 46 nonzero coefficients of $Q_{1,105}$ in the expansion of $ECyc_{105}(q, N_1)$, 14 of which have the incorrect 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 given by the following proposition.

Proposition 16.

$$ECyc_{2^m} = 2Cyc_{2^{m-1}}(q) - N_{2^{m-1}}. \tag{27}$$

In particular, for i between 1 and $\phi(2^m) = 2^{m-1}$, we get

$$Q_{i,2^m} = P_{i,2^{m-1}}, \tag{28}$$

where the $P_{i,k}$ are the coefficients of N_k.

Note that in our proof we 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 16 actually extends Proposition 15 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 (28) follows from (27). 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 $C(\mathbb{F}_{q^k})$ where q itself can also be a power of p. Consequently,

$$N_{m\cdot k}(q,\ N_1) = \#C(\mathbb{F}_{q^{m\cdot k}}) = N_m(q^k,\ N_k). \tag{29}$$

While this relation is immediate given our definition of $N_k = \#C(\mathbb{F}_{q^k})$, when we translate this relation in terms of spanning trees, the relation

$$\mathcal{W}_{mk}(q,t) = \mathcal{W}_m\big(q^k, \mathcal{W}_k(q,t)\big) \tag{30}$$

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 4.1, 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.

4.5. **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 7. *We have*

$$ECyc_d = \big|Ker\big(Cyc_d(\pi)\big) :\ C(\overline{\mathbb{F}_q}) \to C(\overline{\mathbb{F}_q})\big|,$$

where π denotes the Frobenius map, and $Cyc_d(\pi)$ is an element of $End(C) = End(C(\overline{\mathbb{F}_q}))$.

Proof. One of the key properties of the Frobenius map is the fact that $C(\mathbb{F}_{q^k}) = Ker(1-\pi^k)$, where $1-\pi^k$ is an element of $End(C)$. See [17] for example. The map $(1-\pi^k)$ factors into cyclotomic polynomials in $End(C)$ 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 $\big|\text{Ker }\big(Cyc_{d_1}Cyc_{d_2}(\pi)\big)\big|$ equals $\big|\text{Ker } Cyc_{d_1}(\pi)\big| \cdot \big|\text{Ker } Cyc_{d_2}(\pi)\big|$. Thus

$$\prod_{d|k} ECyc_d = N_k = \big|\text{Ker } (1-\pi^k)\big| = \Big|\text{Ker}\prod_{d|k} Cyc_d(\pi)\Big| = \prod_{d|k}\big|\text{Ker } Cyc_d(\pi)\big|,$$

and since the last equation is true for all $k \geq 1$, we must have the relations

$$ECyc_d = \big|\text{Ker } Cyc_d(\pi)\big| \tag{31}$$

for all $d \geq 1$. □

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 $WCyc_d(q,t) = -ECyc_d\big|_{N_1 \to -t}$.

This motivates the analogous question, namely does there exist a combinatorial or geometric interpretation of these polynomials? We in fact can answer this in the affirmative and do so in [15, Chap. 6] as well as in a forthcoming paper.

Remark 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.

5. Conclusions and open problems

The new combinatorial formula for N_k presented in this write-up appears fruitful. It leads one to ask how spanning trees of the wheel graph are related to points on elliptic curves. For instance, is there a reciprocity that explains combinatorially why the bivariate integral polynomial formulas for counting points on elliptic curves and counting spanning trees of the wheel graph are equivalent except for the appearance of alternating signs? Such reciprocities occur frequently in combinatorics. For example given the chromatic polynomial $\chi(\lambda)$ of a graph $G = (V,E)$, the expression $(-1)^{|V|}\chi(-1)$ provides a formula for the number of acyclic orientations of G [19].

The fact that the Fibonacci and Lucas numbers also enter the picture is also exciting since these numbers have so many different combinatorial interpretations, and there is such an extensive literature about them. Perhaps these combinatorial interpretations will lend insight into why N_k depends only on the finite data of N_1 and q for an elliptic curve, and how we can associate points over higher extension fields to points on $C(\mathbb{F}_q)$.

The elliptic cyclotomic polynomials provide an additional source of new questions. What is the spanning tree interpretation of $\mathcal{W}_k(q,N_1)$'s factorization? Is there a combinatorial interpretation of $\mathcal{W}_{mk}(q,t) = \mathcal{W}_m(q^k, \mathcal{W}_k(q,t))$? What is a combinatorial interpretation of the integral polynomials $Q_{i,d}$, and what does the fact their coefficients are almost all positive mean? We will tackle some of these problems in a forthcoming paper in which we compare more thoroughly the structures of elliptic curves and spanning trees.

Acknowledgements. The author would like to thank Adriano Garsia for many useful conversations and his invaluable guidance through the author's graduate school. I would also like to thank the referees for their valuable reports in leading to a clearer exposition. In particular, the connection to orthogonal polynomials, including the statement of Theorem 6, was indicated by one of the referees, as were the observations that several of the propositions have direct algebraic proofs. Additionally, Thomas Shemanske brought the significance of the number 105, with respect to ordinary cyclotomic polynomials, to our attention. Section 2 of this paper was presented at FPSAC 2006 and the author would like to thank the conference referees for their edits. This work was supported by the NSF, grant DMS-0500557.

References

[1] A. Benjamin and C. Yerger, Combinatorial Interpretations of Spanning Tree Identities, *Bull. Inst. Combin. Appl.*, to appear.

[2] H. D. Block and H. P. Thielman, Commutative Polynomials. *Quart. J. Math. Oxford Ser. 2* **2** (1951), 241–243.

[3] F. Boesch and H. Prodinger, Spanning Tree Formulas and Chebyshev Polynomials. *Graphs Combin.* **2** (1986), no. 3, 191–200.

[4] P. Borwein and T. Erdélyi, *Polynomials and Polynomial Inequalities, volume 161 of Graduate Texts in Mathematics*, Springer–Verlag, New York (1995).

[5] O. Eğecioğlu and J. Remmel, Brick Tabloids and the Connection Matrices Between Bases of Symmetric Functions, *Discrete Appl. Math.*, **34** (1991), 107–120.

[6] G. Frey, Applications of Arithmetical Geometry to Cryptographic Constructions. *Finite Fields and Applications (Augsburg, 1999)*, Springer, Berlin, (2001), 128–161.

[7] A. Garsia and G. Musiker, *Basics on Hyperelliptic Curves over Finite Fields*, in progress.

[8] M. Golin, X. Yong, Y. Zhang, Chebyshev Polynomials and Spanning Tree Formulas for Circulant and Related Graphs. *Discrete Math.*, **298** (2005), no. 1–3, 334–364.

[9] H. Hasse, Abstrakte Begründung der komplexen Multiplikation und Riemannsche Vermutung in Funktionenkörpern, *Abhandlungen aus dem mathematischen Seminar der Hamburgischen Universität*, **10** (1934), 325–348.

[10] M. E. H. Ismail, H. Prodinger and D. Stanton, Schur's Determinants and Partition Theorems. *Sém. Lothar. Combin.*, **44** (2000), Art. B44a.

[11] N. Loehr, G. Warrington, and H. Wilf, The Combinatorics of a Three-Line Circulant Determinant. *Israel J. Math.*, **143** (2004), 141–156.

[12] I. Macdonald, *Symmetric Functions and Hall Polynomials*, second ed., Oxford Mathematical Monographs, The Clarendon Press Oxford University Press, New York, (1995).

[13] C. Moreno, *Algebraic Curves over Finite Fields*, Cambridge Tracts in Mathematics, vol. 97, Cambridge University Press, Cambridge (1991).

[14] B. R. Myers, Number of Spanning Trees in a Wheel, *IEEE Trans. Circuit Theory*, **18** (1971), 280–282.

[15] G. Musiker, *A Combinatorial Comparison of Elliptic Curves and Critical Groups of Graphs*, Ph.D. thesis, UCSD, 2007.

[16] G. Musiker and J. Propp, Combinatorial Interpretations for Rank-Two Cluster Algebras of Affine Type, *Electron. J. Combin.*, **14** (2007), no. 1, Research Paper 15, 23 pp.

[17] J. Silverman, *The Arithmetic of Elliptic Curves, Volume* 106 *of Graduate Texts in Mathematics*, Springer–Verlag, New York (1986).

[18] N. J. A. Sloane, *The On-Line Encyclopedia of Integer Sequences*, `http://www.research.att.com/~njas/sequences/index.html`.

[19] R. P. Stanley, Acyclic Orientations of Graphs. *Discrete Math.*, **5** (1973), 171–178.

[20] R. P. Stanley, *Enumerative Combinatorics Vol. 2, Volume 62 of Cambridge Studies in Advanced Mathematics*, Cambridge University Press, Cambridge (1999).

[21] L. Washington, *Elliptic Curves: Number Theory and Cryptography. Discrete Mathematics and its Applications*, Chapman & Hall/CRC, Boca Raton, (2003).

[22] A. Weil, *Sur les courbes algébriques et les variétés qui s'en déduisent*, Hermann, Paris (1948).

[23] A. Zelevinsky, Semicanonical Basis Generators of the Cluster Algebra of Type $A_1^{(1)}$, *Electron. J. Combin.*, **14** (2007), no. 1, Note 4, 5 pp.

Mathematics Department, University of California, San Diego
E-mail address: `gmusiker@math.ucsd.edu`

2015

A Combinatorial Exploration of Elliptic Curves

Matthew Lam
Harvey Mudd College

A Combinatorial Exploration of Elliptic Curves

Matt Lam

Nicholas J. Pippenger, Advisor

Michael R. Orrison, Reader

HARVEY MUDD COLLEGE

Department of Mathematics

May, 2015

Abstract

At the intersection of algebraic geometry, number theory, and combinatorics, an interesting problem is counting points on an algebraic curve over a finite field. When specialized to the case of elliptic curves, this question leads to a surprising connection with a particular family of graphs. In this document, we present some of the underlying theory and then summarize recent results concerning the aforementioned relationship between elliptic curves and graphs. A few results are additionally further elucidated by theory that was omitted in their original presentation.

Contents

Chapter 1

Background

1.1 Symmetric Functions

Our approach to the enumeration of points on curves is closely tied to the theory of symmetric functions. We therefore develop a bit of this theory so that we can discuss the enumeration of points in a naturally suited language.

Definition 1.1. *A homogeneous symmetric function of degree n is a formal power series*

$$f(x) = \sum_{\alpha} c_\alpha x^\alpha$$

where $\alpha = (\alpha_1, \alpha_2, \ldots)$ runs over all sequences of nonnegative integers whose sum is n, c_α is a scalar, and x^α represents the monomial $x_1^{\alpha_1} x_2^{\alpha_2} \cdots$. Furthermore, $f(x_{\sigma(1)}, x_{\sigma(2)}, \ldots) = f(x_1, x_2, \ldots)$ for every permutation σ of the positive integers.

It is important to note that, despite their name, symmetric functions should be regarded as purely formal power series and not as actual functions to be evaluated. The set of all homogeneous symmetric functions of degree n is denoted by Λ^n, and the direct sum $\Lambda = \Lambda^0 \oplus \Lambda^1 \oplus \cdots$ is called the *algebra of symmetric functions*. We now describe three important generators of Λ and their relation to one another.

Definition 1.2. *For each positive integer k, we define the elementary symmetric function*

$$e_k = \sum_{i_1 < \cdots < i_k} x_{i_1} \cdots x_{i_k}\ , \quad k \geq 1 \quad (\textit{with } e_0 = 1).$$

In words, e_k is the sum of all distinct products of k distinct variables.

Definition 1.3. *For each positive integer* k*, we define the complete symmetric function*

$$h_k = \sum_{i_1 \leq \cdots \leq i_k} x_{i_1} \cdots x_{i_k}\,, \quad k \geq 1 \quad (\textit{with } h_0 = 1).$$

In words, h_k *is the sum of all distinct products of* k *not-necessarily distinct variables.*

In both types of symmetric functions, the number of actual variables has not been specified. Also note the similarity between the two formal definitions; their relationship is quite analogous to the reciprocity between choose and multi-choose. In fact, taking the two symmetric functions in n variables and evaluating at $x_1 = \cdots = x_n = 1$, we obtain

$$e_k(1,\ldots,1) = \sum_{i_1 < \cdots < i_k} 1 = \binom{n}{k}$$
$$h_k(1,\ldots,1) = \sum_{i_1 \leq \cdots \leq i_k} 1 = \left(\!\!\binom{n}{k}\!\!\right) = (-1)^k \binom{-n}{k}.$$

We expand upon this a bit further.

Proposition 1.1. *Define the endomorphism* $\omega : \Lambda \to \Lambda$ *by* $\omega(e_n) = h_n$. *Then* ω *is an involution, i.e.* ω^2 *is the identity automorphism.*

Proof. Define the auxiliary power series

$$E(t) = \sum_{n \geq 0} e_n t^n, \quad H(t) = \sum_{n \geq 0} h_n t^n.$$

Since the e_n contain products of distinct variables, and the h_n products of not-necessarily distinct variables, we can rewrite these series as

$$E(t) = \prod_i (1 + x_i t), \quad H(t) = \prod_i \frac{1}{1 - x_i t}.$$

For $H(t)$ we have used the closed form expression for a geometric series. It is then clear that $E(t)H(-t) = 1$, hence we can equate coefficients and apply ω to obtain

$$0 = \sum_{k=0}^{n} (-1)^{n-k} h_k \omega(h_{n-k}) = (-1)^n \sum_{k=0}^{n} (-1)^{n-k} \omega(h_k) h_{n-k}.$$

The elementary symmetric functions generate Λ as an algebra, so this does in fact fully define an endomorphism. (For a proof of this, see Stanley (2001).)

The last step involved reindexing the summation $k \to n-k$, i.e. reversing the summation limits. Now consider $n = 0$; it follows that $\omega(h_0) = 1 = e_0$. Consequently, by equating further coefficients the result follows inductively. □

This result generalizes the special case $\binom{n}{k} \to (-1)^k \binom{-n}{k}$ corresponding to $e_k(1,\ldots,1) \to h_k(1,\ldots,1)$ seen above. We now introduce one more class of symmetric functions.

Definition 1.4. *For each positive integer k, we define the power sum symmetric function*

$$p_k = \sum_i x_i^k\,, \quad k \geq 1 \quad (\textit{with } p_0 = 1).$$

Like the homogeneous and elementary symmetric functions, the power sum symmetric functions also generate Λ. However, it is often convenient to work with a linear basis for Λ. This brings us to the following proposition.

Proposition 1.2. *Recall that a partition λ of a positive integer n is a positive sequence $(\lambda_1,\ldots,\lambda_k)$ where $\sum_i \lambda_i = n$. Let $Par := \bigcup_{n\geq 0} Par(n)$, i.e. the set of all partitions of all positive integers. Then*

$$\{h_\lambda = h_{\lambda_1}\cdots h_{\lambda_k}\},\ \{e_\lambda = e_{\lambda_1}\cdots e_{\lambda_k}\},\ \{p_\lambda = p_{\lambda_1}\cdots p_{\lambda_k}\};$$
$$\lambda = (\lambda_1,\ldots,\lambda_k) \in Par$$

are each additive bases for Λ.

See Stanley (2001) for a proof.

1.1.1 Plethysm

Definition 1.5. *Let $f \in \Lambda$ be the sum of monomials $\sum_{i\geq 0} x^{\alpha_i}$. The plethysm $g[f]$ is defined as*

$$g[f] = g(x^{\alpha_1}, x^{\alpha_2}, \ldots).$$

Although this appears to depend on the order that the monomials of f are summed, the operation is in fact well defined because g is symmetric, hence any reordering of the terms in f yield the same result. The plethysm $g[f]$ is sometimes written as $f \circ g$, because in certain contexts the operation really is composition. We will not use that notation. In general, the plethystic expression $g[f]$ is only defined when the number of monomials in f equals the number of variables in g.

Example 1.1. *Consider the power symmetric function p_k and the arbitrary symmetric function f from definition 1.5. Then*

$$f[p_k] = f(x_1^k, x_2^k, \ldots) = \sum_{i \geq 0} x^{\alpha_i k} = p_k[f].$$

It is clear that

$$(af + bg)[h] = af[h] + bg[h] \text{ and}$$
$$(fg)[h] = f[h] \cdot g[h],$$

so with example 1.1 we can define the plethysm for any functions by using the power sum symmetric basis. In particular, if $g = \sum_\lambda c_\lambda p_\lambda$ where the c_λ are scalars, then

$$g[f] = \sum_\lambda c_\lambda p_\lambda[f] = \sum_\lambda c_\lambda \prod_{i=1}^{\ell(\lambda)} f(x_1^{\lambda_1}, x_2^{\lambda_2}, \ldots).$$

1.2 The Zeta Function of a Curve

To study the numbers N_s of points on a curve C over the finite field F_{q^s}, we consider the generating function $\sum_{s \geq 1} N_s u^s$. As with the case of symmetric functions, we will deal with these as formal power series.

Definition 1.6. *The zeta function of a curve C is given by the series*

$$Z_C(u) = \exp\left(\sum_{s=1}^{\infty} \frac{N_s u^s}{s}\right),$$

where we are using the identity $\exp(u) = \sum_{s \geq 0} u^s/s!$.

Although the zeta function essentially encodes the same information as $\sum_{s \geq 1} N_s u^s$, it turns out that the zeta function is more convenient to work with.

Example 1.2. *Consider the circle at infinity in the finite projective plane, i.e. $z = 0$. By definition, this is the set of points $(x, y, 0) \in \mathbb{P}^2(F_{q^s})$, and the number of such* *points equals the number of points in* $\mathbb{P}^1(F_{q^s})$*. Therefore* $N_s = q^s + 1$*, so*

$$\begin{aligned}\sum_{s=1}^{\infty} \frac{N_s u^s}{s} &= \sum_{s=1}^{\infty} \frac{(q^s+1)u^s}{s} \\ &= \left(\sum_{s=1}^{\infty} \frac{u^s}{s}\right) + \left(\sum_{s=1}^{\infty} \frac{q^s u^s}{s}\right) \\ &= -\ln(1-u) - \ln(1-qu) \\ &= -\ln((1-u)(1-qu)).\end{aligned}$$

The third line follows from the identity $\sum_{s\geq 1} w^s/s = -\ln(1-w)$*. Hence*

$$Z_z(u) = \frac{1}{(1-u)(1-qu)}.$$

A key feature of this function is that it is rational with integer coefficients. This leads us to an important theorem regarding the enumeration of points on elliptic curves.

Theorem 1.1. (*Weil*) *Let* $f(x,y,z) \in F_{q^s}[x,y,z]$ *be a nonzero, nonsingular homogeneous polynomial. Then*

$$Z_f(u) = \frac{P(u)}{(1-u)(1-qu)}$$

where $P(u)$ *is a polynomial with integer coefficients, with degree equaling twice the genus of the curve, and* $P(0) = 1$*.*

Recall that a polynomial is nonsingular if the curve it defines has a unique tangent line at every point, and also that the genus of a curve is defined as $g = \frac{1}{2}(d-1)(d-2)$.

Corollary. *If E is an elliptic curve, then*

$$Z_E(u) = \frac{1-(\alpha_1+\alpha_2)u + \alpha_1\alpha_2 u^2}{(1-u)(1-qu)}.$$

Chapter 2

Enumerating Points on Elliptic Curves

A motivating result for our forthcoming investigation is that, for an algebraic curve of genus g, the number of points over the finite fields $F_q, F_{q^2}, \ldots, F_{q^g}$ is sufficient data to determine the number of points over any higher field extension. This leads one to question how exactly the points over these higher field extensions relate to those over the first g. In the remaining discussion we will focus on the case $g = 1$.

2.1 Preliminary Results

In the background section we have already begun to touch upon the enumeration of points on algebraic curves. Combining definition 1.6 with the corollary of theorem 1.1, it is seen by equating coefficients that $N_k = 1+q^k-\alpha_1^k-\alpha_2^k$. Or, in plethystic notation, $N_k = p_k[1+q-\alpha_1-\alpha_2]$. Since the case $k = 1$ yields the relation $\alpha_1 + \alpha_2 = 1 + q - N_1$, it follows that q and N_1 fully determine all N_k.

Theorem 2.1 (Garsia)**.**

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

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

The $P_{i,k}$ in fact relate directly to wheel graphs and their spanning trees. In particular, the quantity $1 + q^k - N_k$ can be shown to satisfy the same recurrence relation as a generalization of the Lucas numbers. A bijection is then established between the generalized Lucas numbers and the spanning trees of a wheel graph. See Musiker (2007) for details.

Other aspects of the zeta function yield combinatorial identities as well. For instance, using the symmetric function identity

$$\exp\left(\sum_{k\geq 1}\frac{p_k u^k}{k}\right) = \frac{1}{\sum_{k\geq 0}(-1)^k e_k u^k}$$

and the fact that $N_k = p_k[1 + q - \alpha_1 - \alpha_2]$, we can write the zeta function as

$$Z_E(u) = \frac{1}{\sum_{k\geq 0}(-1)^k E_k u^k}$$

where $E_k = e_k[1 + q - \alpha_1 - \alpha_2]$. It turns out that the E_k can be obtained by evaluating a bivariate polynomial generalization of the Fibonacci numbers at the point $(q, -N_1)$. So like the N_k, the E_k also have a natural formula in terms of q and N_1. A proof of this is also given in Musiker (2007). Lastly, consider the symmetric function identity

$$\exp\left(\sum_{k\geq 1}\frac{p_k u^k}{k}\right) = \sum_{k\geq 0} h_k u^k.$$

Following the above reasoning, the zeta function has yet another form

$$Z_E(u) = \sum_{k\geq 0} H_k u^k$$

where $H_k = h_k[1 + q - \alpha_1 - \alpha_2]$.

2.2 (q, t)-Wheel Graphs

As mentioned above, the equation in Theorem 2.1 leads to a connection with wheel graphs. We define the (q, t)-wheel graph on $k + 1$ vertices by the following construction. Begin with the cycle graph on k vertices, with edges directed counter-clockwise. Then include an additional central vertex, which is attached by t bidirectional spokes to each rim vertex. Lastly, attach q clockwise edges between each pair of adjacent rim vertices. This construction will be denoted by $W_k(q, t)$. See Fig. 2.1 for an example.

With this construction, Theorem 2.1 may be restated very cleanly.

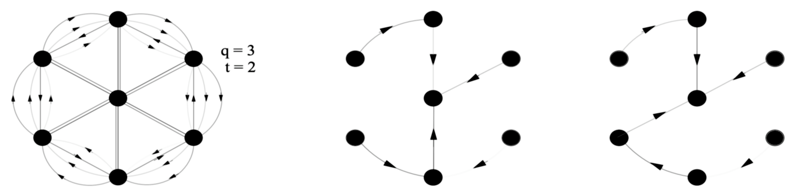


Figure ▯.▯ The (q, t)-wheel graph $W_6(3, 2)$ and two directed spanning trees with central roots.

Theorem 2.2. *Let $\mathcal{W}_k(q, t)$ denote the number of directed spanning trees of $W_k(q, t)$ with all edges directed towards the central vertex. Then*

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

The expression $-\mathcal{W}_k(q, -N_1)$ does not admit the same enumerative interpretation as given above, because it would refer to a graph containing negative numbers of edges. This result is in similar spirit to a theorem due to R.P. Stanley, which gives a combinatorial interpretation to the evaluation of a graph's chromatic polynomial on negative integers. In particular, $|\chi(-1)|$ equals the number of acyclic orientations of the graph. See Stanley (1972) for details.

In light of the generating function identities obtained in section 2.1, Theorem 2.2 suggests an analogous investigation of the generating function

$$W_{q,t}(u) = \exp\left(\sum_{k=1}^{\infty} \frac{\mathcal{W}_k(q, t)u^k}{u}\right).$$

Using Theorem 2.2 and the corollary of Theorem 1.1,

$$\begin{aligned} W_{q,N_1}(u) &= \frac{1}{\exp\left(\sum_{k=1}^{\infty} \frac{N_k|_{N_1 \to -N_1} u^k}{u}\right)} \\ &= \frac{(1-u)(1-qu)}{1-(1+q+N_1)u+qu^2} \end{aligned}$$

where in the second step N_k is to be viewed as a function of N_1. Factoring the denominator as $1-(1+q+N_1)u+qu^2 = (1-\beta_1 u)(1-\beta_2 u)$, we can expand the generating function as a product of two geometric series. Matching coefficients then yields

$$\mathcal{W}_k(q, N_1) = (-1)^k + (-q)^k + \beta_1^k + \beta_2^k = p_k[-1-q+\beta_1+\beta_2];$$

note the similarity compared to $N_k = p_k[1+q-\alpha_1-\alpha_2]$. The same generating function identities used in section 2.1 can also be applied to express E_k and H_k in terms of this "alphabet" $-1-q+\beta_1+\beta_2$. The results are compiled in the following table.

	Elliptic Curves	(q,t)-Wheel Graphs
Exponential generating function	$\frac{1-(1+q-N_1)u+qu^2}{(1-u)(1-qu)}$	$\frac{(1-u)(1-qu)}{1-(1+q+N_1)u+qu^2}$
Alphabet	$1+q-\alpha_1-\alpha_2$	$-1-q+\beta_1+\beta_2$
N_k ($\mathcal{W}_k$ for wheel graphs)	$p_k[1+q-\alpha_1-\alpha_2]$	$p_k[1+q-\alpha_1-\alpha_2]$
H_k	$h_k[1+q-\alpha_1-\alpha_2]$	$(-1)^{k-1}e_k[1+q-\alpha_1-\alpha_2]$
E_k	$e_k[1+q-\alpha_1-\alpha_2]$	$(-1)^k h_k[1+q-\alpha_1-\alpha_2]$

Chapter 3

Elliptic Curves and Critical Groups of Wheel Graphs

The previous chapter connected elliptic curves to wheel graphs by equating their numbers of points and spanning trees, respectively. Recall, however, that to establish this equivalence we had to construct "graphs" with negative numbers of edges. Thus the most basic description using a vertex set and edge set does not apply, but matrix representations are perfectly admissible. This chapter is largely devoted to the relation between these matrix representations and elliptic curves.

3.1 Elliptic Curve Hierarchy

Several key properties of elliptic curves have analogous statements concerning critical groups (as yet undefined) of wheel graphs. These properties of elliptic curves are now briefly described.

Let E be an elliptic curve (over an unspecified field) and $q = p^k$ for some prime p and positive integer k. Recall that there is an inclusion of fields

$$\mathbb{F}_q \subset \mathbb{F}_{q^{k_1}} \subset \mathbb{F}_{q^{k_2}} \subset \cdots \subset \overline{\mathbb{F}_p}$$

whenever the divisibilities $k_i | k_{i+1}$ hold, and where $\overline{\mathbb{F}_p}$ is the algebraic closure of $\mathbb{F}_q$. This implies a subgroup series

$$E(\mathbb{F}_q) \subset E(\mathbb{F}_{q^{k_1}}) \subset E(\mathbb{F}_{q^{k_2}}) \subset \cdots \subset E(\overline{\mathbb{F}_p})$$

with the same divisibility constraints $k_i | k_{i+1}$. One more important feature is the Frobenius endomorphism $\phi : x \mapsto x^q$, which is an element of (and in fact generates) the Galois group $\mathrm{Gal}(\mathbb{F}_{q^\ell}/\mathbb{F}_q)$. Since ϕ must fix the ground field, extending this map to act on points by $\phi : P = (x, y) \mapsto (x^q, y^q)$ has the property that

$$\phi^k(P) = P \text{ if and only if } P \in E(\mathbb{F}_{q^k}).$$

In other words, each group $E(\mathbb{F}_{q^k})$ can be defined by $\ker(1 - \phi^k)$. See Silverman (2009) for a more detailed discussion.

3.2 Critical Groups

Let $G = (V, E)$ be a (possibly) directed graph, $A(G)$ be the usual adjacency matrix of G, and $\Delta(G)$ be the diagonal matrix such that $\Delta(G)_{vv}$ is the out-degree of vertex v. ($\Delta(G)$ is usually referred to as the degree matrix of G.) The Laplacian is defined as $Q(G) = \Delta(G) - A(G)$.

Definition 3.1. *The critical group $\mathcal{K}(G)$ of a graph G is the cokernel of the transpose of the Laplacian $Q(G)$ acting on $\mathbb{Z}^{|V|}$:*

$$\mathcal{K}(G) := \operatorname{coker} Q(G)^T = \mathbb{Z}^{|V|}/Q(G)^T\mathbb{Z}^{|V|}.$$

The dual critical group of G is similarly defined to be $\mathcal{K}^*(G) :=$ coker $Q(G)$.

Definition 3.2. *The reduced critical group $K(G)$ of G is the torsion subgroup of $\mathcal{K}(G)$. Similarly, the reduced dual critical group $K^*(G)$ is the torsion subgroup of $\mathcal{K}^*(G)$.*

The relationship between a graph and the structure of its critical group is, in general, not well understood. There exist, however, well-behaved examples. See Biggs (1999) for a complete classification in the case of simple wheel graphs.

Definition 3.3. *Let $G = (V, E)$ be a digraph and $\pi = (\pi_1, \ldots, \pi_p)$ be an ordered partition of V. The partition π is said to be* equitable *for G if there exist nonnegative integers F_{ij} and R_{ij} for all $1 \le i, j \le p$ such that every vertex in π_i is the initial vertex of exactly F_{ij} edges having terminal vertices in π_j, and every vertex in π_j is the terminal vertex of exactly R_{ij} edges with initial vertex in π_i.*

Definition 3.4. *Let G be a digraph and π be an equitable partition of G. The quotient of G by π, denoted G/π, is the graph whose adjacency matrix is given by F_{ij}.*

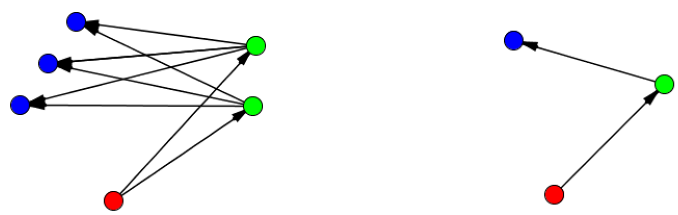


Figure 3.1 A digraph (left) and its quotient (right) by an equitable partition marked by vertex color.

Figure 3.1 illustrates an example of a directed graph and its quotient by an equitable partition. Not surprisingly, the critical groups of a graph and one of its quotients are related. The following theorem makes this precise.

Theorem 3.1 (Wagner)**.** *Let $G = (V, E)$ be a strongly connected graph and $\pi = \{\pi_1, \ldots, \pi_p\}$ be an equitable partition of G. Then there exists a natural injective homomorphism $\psi : \mathcal{K}(G/\pi) \to \mathcal{K}(G)$ given by $\psi(x) = Px$, where P is the $V \times \{1, \ldots, p\}$ matrix with entries*

$$P_{vi} = \begin{cases} 1 & \text{if } v \in \pi_i, \\ 0 & \text{if } v \notin \pi_i. \end{cases}$$

The restriction of ψ to the torsion subgroup $\psi|_{tor} : K(G/\pi) \to K(G)$ is also injective.

Under the hypotheses of Theorem 3.1, $\mathcal{K}(G/\pi)$ may be regarded as a subgroup of $\mathcal{K}(G)$ with the inclusion ψ.

3.3 Critical Groups of (q, t)-Wheel Graphs

Throughout this section, we will denote $K(W_k(q, t))$ by $K(k, q, t)$. Our general aim is to map properties of the sequence $\{K(k, q, t)\}_{k \geq 1}$ onto those of elliptic curves given above.

In Musiker (2009), Musiker defines the critical group of a wheel graph using not the Laplacian Q, but the reduced Laplacian Q_0 obtained by deleting the row and column corresponding to the central vertex. This structure is manifestly not the critical group of any graph, but is in fact isomorphic to the reduced critical group.

Proposition 3.1. *The "critical group" defined using the reduced Laplacian Q_0 of a wheel graph $W_k(q,t)$ is isomorphic to the reduced critical group $K(k,q,t)$.*

$$\operatorname{coker}(Q_0^T(W_k(q,t))) \cong K(k,q,t).$$

Proof. For a wheel graph $W_k(q,t)$, the Laplacian is given by the $(k+1)\times(k+1)$ matrix

$$Q = \begin{bmatrix} 1+q+t & -q & 0 & \cdots & 0 & -1 & -t \\ -1 & 1+q+t & -q & 0 & \cdots & 0 & -t \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & -t \\ 0 & \cdots & -1 & 1+q+t & -q & 0 & -t \\ 0 & \cdots & 0 & -1 & 1+q+t & -q & -t \\ -q & 0 & \cdots & 0 & -1 & 1+q+t & -t \\ -t & -t & -t & \cdots & -t & -t & kt \end{bmatrix},$$

where the last row and column correspond to the central vertex. By adding the first k rows to the last row, and then the first k columns to the last column, we obtain the matrix

$$Q' = \begin{bmatrix} 1+q+t & -q & 0 & \cdots & 0 & -1 & 0 \\ -1 & 1+q+t & -q & 0 & \cdots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & \cdots & -1 & 1+q+t & -q & 0 & 0 \\ 0 & \cdots & 0 & -1 & 1+q+t & -q & 0 \\ -q & 0 & \cdots & 0 & -1 & 1+q+t & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 \end{bmatrix}.$$

Since Q and Q' are related by invertible matrices, $\operatorname{coker}(Q^T) \cong \operatorname{coker}(Q'^T)$. Note also that the image of Q'^T is isomorphic to the image of Q_0^T, the only formal difference being the presence of a zero in the last entry of every element in Im Q'^T. It follows that $\operatorname{coker}(Q'^T) \cong \operatorname{coker}(Q_0^T) \oplus \mathbb{Z}$, thus

$$\mathcal{K}(W_k(q,t)) \cong \operatorname{coker}(Q_0^T) \oplus \mathbb{Z}.$$

In Musiker (2009) it is shown that $\operatorname{coker}(Q_0^T)$ has finite order, and therefore must be isomorphic to the reduced critical group. □

Proposition 3.2. *Let E be an elliptic curve and N_k be the number of points of E over the finite field $\mathbb{F}_{q^k}$. Then Theorem 2.2 is equivalent to $N_k = |K(k,q,-N_1)|$.*

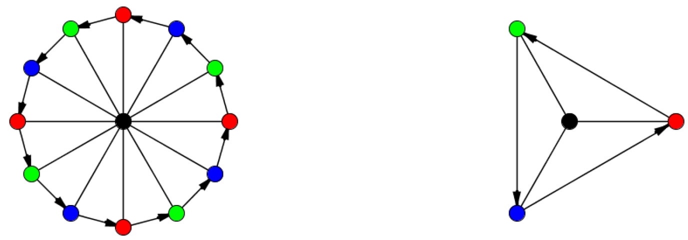


Figure ⊠.⊠ A wheel graph $K(12, 0, 1)$ (left) and its quotient $K(3, 0, 1)$ (right) by a partition using $l = 3$ parts marked by color.

Proof. The Matrix-Tree theorem tells us that for a connected, undirected graph G, the order of $K(G)$ equals the number of spanning trees of G. A slight extension of this theorem shows that for a digraph G and reduced Laplacian $Q_v(G)$ obtained by deleting the row and column associated with vertex v, the order of $\mathrm{coker}(Q_v^T)$ is equal to the number of directed spanning trees of G with sink v.

From Prop. 3.1 it follows that $|\,\mathrm{coker}(Q_0^T)| = |K(k, q, t)|$, and then substituting $t \to -N_1$ yields the desired result. □

We now turn our attention to the structure of $\{K(k, q, t)\}_{k \geq 1}$ itself. Recall Theorem 3.1, which provides an injective homomorphism between critical groups whenever one of their associated graphs is a quotient of the other. There are in fact very natural equitable partitions on wheel graphs; in particular, if l divides the number of rim vertices k, then we can form parts by walking around the rim and repeatedly counting off to l so that each part contains k/l vertices. The hub vertex comprises its own part. Taking the quotient by this partition yields another wheel graph with l rim vertices. See Figure 3.2 for an example.

This implies the existence of an injective homomorphism

$$\psi_{k_2,k_1} : K(k_1, q, t) \to K(k_2, q, t)$$

whenever k_1 divides k_2. As mentioned above, we may thus view $K(k_1, q, t)$ as a subgroup of $K(k_2, q, t)$ when $k_1 \mid k_2$. This is the exact partial ordering as given for elliptic curves, where $E(\mathbb{F}_{q_1^k}) \leq E(\mathbb{F}_{q_2^k})$ precisely when $k_1 \mid k_2$.

By definition of the map ψ given in Theorem 3.1, we have $\psi_{k_3,k_2} \circ \psi_{k_2,k_1} = \psi_{k_3,k_1}$. Therefore we can form the direct limit

$$\overline{K}(q,t) := \varinjlim_{k \geq 1} \{K(k,q,t)\}$$

so that every critical group $K(k,q,t)$ may be naturally identified with a subgroup of $\overline{K}(q,t)$. This direct limit is analogous to the field $\overline{\mathbb{F}_p}$. Closer examination of the maps ψ_{k_2,k_1} reveals that they simply repeat the input vector k_2/k_1 times. One is then lead to view $\overline{K}(q,t)$ as the set of all periodic vectors $w = (\ldots, w_{-1}, w_0, w_1, \ldots)$, so that the subgroup of $\overline{K}(q,t)$ isomorphic to $K(k,q,t)$ is all the vectors of period k.

Define the shift map $\rho : \overline{K}(q,t) \to \overline{K}(q,t)$ by

$$\rho(\ldots, w_{i-1}, w_i, w_{i+1}, \ldots) = (\ldots, w_{i-2}, w_{i-1}, w_i, \ldots).$$

Then a theorem due to Musiker states that for all $k \geq 1$, $q \geq 0$, and $t \geq 1$,

$$K(k,q,t) \cong \mathrm{Ker}(1 - \rho^k).$$

The results of this chapter are summarized by the following correspondences:

$$\begin{array}{rcl} K(k,q,t) & \longleftrightarrow & E(\mathbb{F}_{q^k}) \\ \overline{K}(q,t) & \longleftrightarrow & E(\overline{\mathbb{F}_p}) \\ \text{Frobenius map } \pi & \longleftrightarrow & \text{shift map } \rho \end{array}$$

where $K(k_1,q,t) \leq K(k_2,q,t)$ if and only if $E(\mathbb{F}_{q^{k_1}}) \leq E(\mathbb{F}_{q^{k_2}})$, and $K(k,q,t) \leq \overline{K}(q,t)$, $E(\mathbb{F}_{q^k}) \leq E(\overline{\mathbb{F}_p})$ for all $k \geq 1$.

The reader may be left wondering what can be said regarding the internal structure of $K(k,q,t)$ as compared to that of $E(\mathbb{F}_{q^k})$. We end the chapter with a theorem addressing a only special case of this question, albeit in a very satisfying way.

Theorem 3.2. *Let E be an elliptic curve with endomorphism ring $End(E) \cong \mathbb{Z}[\pi]$, where π is the Frobenius map. As before, $N_1 = |E(\mathbb{F}_q)|$. Then*

$$E(\mathbb{F}_{q^k}) \cong K(k,q,-N_1).$$

For a proof of this, see Musiker (2009).

Bibliography

Biggs, N.L. 1999. Chip-firing and the critical group of a graph. *Journal of Algebraic Combinatorics* 9.

Ireland, Kenneth, and Michael I. Rosen. 1972. *Elements of Number Theory: Including an Introduction to Equations over Finite Fields*. New York: Bogden and Quigley, Inc.

Musiker, Gregg. 2007. Combinatorial aspects of elliptic curves. *Seminaire Lotharingien de Combinatoire* 56.

———. 2009. The critical groups of a family of graphs and elliptic curves over finite fields. *Journal of Algebraic Combinatorics* 30.

Silverman, Joseph H. 2009. *The Arithmetic of Elliptic Curves*. New York: Springer.

Silverman, Joseph H., and John Tate. 2010. *Rational Points on Elliptic Curves*. New York: Springer.

Stanley, R. P. 1972. Acyclic orientations of graphs. *Discrete Mathematics* 5.

Stanley, Richard P. 2001. *Enumerative Combinatorics*, vol. 2. Cambridge: Cambridge University Press.

Wagner, David G. 2000. The critical group of a directed graph. *ArXiv Mathematics e-prints* math/[illegible]

Alma Mater Studiorum · Università di Bologna

SCUOLA DI SCIENZE
Corso di Laurea in Matematica

ZETA AND L-FUNCTIONS OF ELLIPTIC CURVES

Tesi di Laurea in Teoria dei Numeri

Relatore:
Chiar.mo Prof.
Luca Migliorini

Presentata da:
Matteo Tamiozzo

Correlatore:
Chiar.mo Prof.
Jan Nekovar

II Sessione
Anno Accademico 2013-2014

Introduction

Elliptic curves are extremely rich and fascinating objects, whose study involves techniques borrowed from many different areas of mathematics: from number theory to complex analysis, from algebraic geometry to the theory of Riemann surfaces. Our point of view in this thesis will be mainly algebraic: we will study elliptic curves within the context of plane algebraic curves, and we will be mainly interested in the properties of elliptic curves defined over finite fields and over $\mathbb{Q}$.
However, the seemingly purely algebraic problem of studying elliptic curves defined over finite fields and over $\mathbb{Q}$ will naturally lead us to introduce two types of analytic functions attached to an elliptic curve, the so called Zeta functions and L-functions. Our main aim will be to study some fundamental properties of these functions, investigating how the analytic properties of the Zeta and L-function of an elliptic curve reflect on the arithmetic properties of the curve. The idea of associating analytic functions to algebraic objects, using analysis to shed new light on number theoretic problems, is widespread in modern number theory and very far reaching. We will analyse in this thesis only some basic, yet very significant examples of this general approach.
Many of the most important results that we are going to describe are particular instances of general theorems whose proof requires advanced tools. However, it is often possible to prove special cases of these results using quite elementary means. Thus we will usually give complete proofs of particular cases of many general results (such as the Riemann hypothesis for elliptic curves, which is proved for two particular families of curves using only the properties of Jacobi sums) and then state the corresponding general theorems without proof. This has the advantage of allowing the reader to familiarize with some deep theorems, to see how far in their proof one can get using only elementary tools and to experience "concretely" the need for more abstract and powerful theories to achieve a complete proof. After all, this is how mathematics develops.
Along the way, we sometimes describe additional results which can be easily proved with the techniques developed (e. g. the quadratic reciprocity law and a counterexample to Hasse-Minkowski local-global principle for cubic equations) and hint at more advanced theorems and conjectures closely related to the topics we deal with (such as Birch and Swinnerton-Dyer conjecture, Weil conjectures, Sato-Tate conjecture). This allows the reader to have a look at some further applications of the results described in the thesis, and at the many fascinating open problems which naturally arise from our discussion and are at the hearth of contemporary mathematical research.

The organisation of the thesis is as follows: in Chapter 1 we briefly recall some preliminary results which will be used throughout the thesis; the reader can just skip through this

Chapter, and come back to it later when its results are cited. Chapter 2 deals with the basic theory of algebraic curves, with a great emphasis on the link between the geometric properties of curves and the algebraic properties of the associate coordinate ring. We also state Riemann-Roch theorem, an essential tool in the sequel. In Chapter 3 we introduce Gauss and Jacobi sums and study some of their properties. Their interest for us lies in the fact that they are among the most effective elementary tools which can be used to count the number of points on curves defined over finite fields. In Chapter 4 we study the Zeta function of affine and projective curves. In particular, we prove the rationality of the Zeta function of a projective plane curve. Chapter 5 introduces elliptic curves and their L-functions. We study in detail elliptic curves with affine equation $y^2 = x^3 + D$ and $y^2 = x^3 - Dx$, proving Riemann hypothesis for these curves and then facing the problem of the existence of an analytic continuation of the associated L-function to the whole complex plane. To solve this problem we define Hecke characters and the associated L-functions, and we hint at the theory of elliptic curves with complex multiplication.
The prerequisites for reading this thesis are quite modest. The reader is assumed to have a basic knowledge of abstract algebra, commutative algebra and complex analysis. Some familiarity with projective geometry and algebraic number theory could be useful but is not strictly necessary: all the results that are used are collected in Chapter 1.

I want to express my sincere gratitude to Gabriele Rembado for his constant support, for the long time he spent discussing with me about mathematics and, above all, for conveying to me his enthusiasm and teaching me the importance of hard work and perseverance.

Introduzione

Le curve ellittiche sono oggetti estremamente ricchi e affascinanti, il cui studio coinvolge tecniche provenienti da diverse aree della matematica: dalla teoria dei numeri all'analisi complessa, dalla geometria algebrica alla teoria delle superfici di Riemann. In questa tesi adotteremo un punto di vista prevalentemente algebrico: studieremo le curve ellittiche inserite nel contesto delle curve algebriche piane, con particolare interesse per le proprietà delle curve ellittiche definite su campi finiti e su $\mathbb{Q}$.
Tuttavia il problema, in apparenza puramente algebrico, di studiare le curve ellittiche definite su campi finiti e su $\mathbb{Q}$ ci porterà naturalmente ad introdurre due tipi di funzioni analitiche associate alle curve ellittiche, le cosidette funzioni Zeta e L. Il nostro principale scopo sarà studiare le proprietà fondamentali di queste funzioni, analizzando come le proprietà analitiche delle funzioni Zeta e L di una curva ellittica si riflettano sulle proprietà aritmetiche della curva. L'idea di associare funzioni analitiche ad oggetti algebrici, utilizzando l'analisi per gettare nuova luce su problemi di teoria dei numeri, è assai diffusa nella moderna teoria dei numeri ed è di portata assai ampia. In questa tesi analizzeremo solo alcuni esempi basilari, ma molto significativi, di questo approccio generale.
Molti dei risultati più importanti che descriveremo sono casi particolari di teoremi generali la cui dimostrazione richiede strumenti avanzati. Tuttavia, è spesso possibile dimostrare questi teoremi in alcuni casi più semplici utilizzando mezzi elementari. Daremo quindi dimostrazioni complete di casi particolari di molti risultati generali (ad esempio, l'ipotesi di Riemann per le curve ellittiche è dimostrata per due famiglie specifiche di curve usando solamente le proprietà delle somme di Jacobi); enunceremo poi i corrispondenti teoremi generali senza dimostrarli. Ciò ha il vantaggio di permettere al lettore di familiarizzare con alcuni risultati profondi, di comprendere quanto lontano ci si possa spingere nella loro dimostrazione usando solo strumenti elementari e di sperimentare "concretamente" la necessità di strumenti teorici più astratti e potenti per ottenere una prova completa. Dopotutto, è così che la matematica si sviluppa.
Nel corso della trattazione sono descritti a volte risultati aggiuntivi la cui dimostrazione si ottiene facilmente con le tecniche sviluppate (ad esempio la legge di reciprocità quadratica e un controesempio al principio di Hasse-Minkowski per equazioni di terzo grado) e sono accennati alcuni teoremi e congetture di natura più avanzata strettamente correlati agli argomenti trattati (come la congettura di Birch e Swinnerton-Dyer, le congetture di Weil, la congettura di Sato-Tate). Questo permette al lettore di esplorare ulteriori applicazioni dei risultati descritti nella tesi, e molti problemi aperti che sorgono naturalmente dalla nostra esposizione e sono al centro della moderna ricerca matematica.

La tesi è organizzata come segue: nel Capitolo 1 sono richiamati brevemente alcuni risul-

tati preliminari utilizzati in tutta la tesi; il lettore può sfogliare rapidamente questo Capitolo, per ritornarci più tardi quando i risultati ivi esposti sono citati. Il Capitolo 2 espone la teoria di base delle curve algenriche, con particolare enfasi sul legame tra le proprietà geometriche delle curve e le proprietà algebriche dell'anello di funzioni polinomiali associato. Si enuncia inoltre il teorema di Riemann-Roch, strumento essenziale nel seguito. Nel Capitolo 3 sono introdotte le somme di Gauss e Jacobi, di cui si studiano alcune proprietà. Per noi, la loro importanza risiede nel fatto che sono tra gli strumenti elementari più efficaci per contare il numero di punti sulle curve definite su campi finiti. Nel Capitolo 4 si studia la funzione Zeta delle curve affini e proiettive. In particolare, si dimostra che la funzione Zeta di una curva piana proiettiva è una funzione razionale. Nel Capitolo 5 sono introdotte le curve ellittiche e le funzioni L ad esse associate. Si studiano in dettaglio le curve ellittiche di equazione affine $y^2 = x^3 + D$ e $y^2 = x^3 - Dx$, dimostrando l'ipotesi di Riemann per queste curve e affrontando il problema dell'esistenza di un prolungamento analitico al piano complesso delle funzioni L associate. Per risolvere questo problema si introduce la nozione di carattere di Hecke e di funzione L ad esso associata, e si accenna alla teoria delle curve ellittiche con moltiplicazione complessa.
I prerequisiti necessari per leggere questa tesi sono minimi. Si assume che il lettore abbia una conoscenza di base dell'algebra astratta, dell'algebra commutativa e dell'analisi complessa. Una certa familiarità con la geometria proiettiva e la teoria algebrica dei numeri può essere utile, ma non è essenziale: tutti i risultati utilizzati sono richiamati nel Capitolo 1.

Contents

Chapter 1

Some algebraic tools

In this chapter, after having fixed some notations and conventions, we collect some tools and results that will be used throughout the thesis. We will recall some fundamental definitions and theorems, and give references for most of the proofs.
The reader who is not already familiar with the topics we will deal with in this chapter can read for example [2], [8], [10], [11], [13], [16].

1.1 Notations and conventions

- Unless the contrary is explicitly stated, all rings will be assumed to be commutative with unit.
  If K is a field, $\bar{K}$ will denote the algebraic closure of K (see section 1.2 for a discussion of this concept).

- If R is a ring, $Spec(R)$ denotes the set of all prime ideals of R, while $Max(R)$ denotes the set of its maximal ideals.

- If R is a ring, R^* will denote the multiplicative group of units of R. In particular, if F is a field, $F^* = F \setminus \{0\}$.

- If A, B are rings, $A \subseteq B, b \in B$, $A[b]$ denotes the intersection of all subrings of B containing A and b. Similarly, if K, L are fields, $K \subseteq L$ and $x \in L$, $K(x)$ denotes the intersection of all subfields of L containing K and x. If $P = (a, b) \in \mathbb{A}^2(\bar{K})$, $K(P)$ denotes the field $K(a, b) \subseteq \bar{K}$.

- If R is a ring, and $a \in R$, (a) denotes the principal ideal generated by a. Two elements $a, b \in R$ are *associates* if $(a) = (b)$. We also say that a is associated with b.

- If $r, s \in R$, $I \subseteq R$ is an ideal, $r \equiv s \pmod{I}$ means $r - s \in I$. If $r \equiv s \pmod{(a)}$ we will also write $r \equiv s \pmod{a}$.

- If A is a set, $|A|$ denotes the cardinality of A.

- If H is a subgroup of a group G, $[G : H]$ denotes the index of H in G.

- $\mathbb{F}_{p^m}$ denotes the field with p^m elements.

- If $f \in \mathbb{F}_p[x, y]$, $N(f = 0)$ denotes the number of solutions of the equation $f = 0$ in $\mathbb{A}^2(\mathbb{F}_p)$. If $f \in \mathbb{F}_p[x, y, z]$ is homogeneous, $N(f = 0)$ denotes the number of solutions to the equation $f = 0$ in $\mathbb{P}^2(\mathbb{F}_p)$. (see Chapter 2 for a definition of these spaces).

1.2 Infinite Galois Theory

Good references for the material covered in this section, including proofs of the statements, are [2], [10], [13].
Let K be a field. Recall that there exists an algebraic extension $\bar{K}$ of K, with a field morphism $i : K \to \bar{K}$, characterised by the following universal property: for each algebraic extension L of K and for each field morphism $\phi : K \to L$ there exists a unique morphism $\bar{\phi}$ making the following diagram commute (equivalently, $\bar{\phi}$ is a morphism of $K-$algebras):

$$\begin{tikzcd} L \arrow[r, "\bar{\phi}"] & \bar{K} \\ K \arrow[u, "\phi"] \arrow[ur, "i"'] & \end{tikzcd}$$

Moreover, $\bar{K}$ is unique up to a K-algebra isomorphism. It is called the *algebraic closure* of K. [1]
Now, fix an algebraic closure $\bar{K}$ of K. If K is perfect, the extension $\bar{K}/K$ is separable; clearly, it is also a normal extension. The group

$$Gal(\bar{K}/K) = \{\sigma : \bar{K} \to \bar{K} :\ \sigma(x) = x\ \forall x \in K\}$$

is called the *absolute Galois group* of K. For any intermediate field $\bar{K} \supseteq L \supseteq K$, L is the fixed field of $Gal(\bar{K}/L)$. Moreover, if L/K is a finite extension, $Gal(\bar{K}/L)$ is a subgroup of finite index in $Gal(\bar{K}/K)$, and we have the equality:

$$[Gal(\bar{K}/K) : Gal(\bar{K}/L)] = [L : K]$$

The absolute Galois group $Gal(\bar{K}/K)$ is isomorphic to the inverse limit $\varprojlim(Gal(L/K))$ for L varying in the set of all finite Galois extensions of K, with the morphisms being restrictions.

Example 1.1. $Gal(\bar{\mathbb{F}}_p/\mathbb{F}_p) \cong \varprojlim Gal(\mathbb{F}_{p^n}/\mathbb{F}_p) = \varprojlim \mathbb{Z}/n\mathbb{Z} = \hat{\mathbb{Z}}$. The Frobenius automorphism, $\phi : x \mapsto x^p$ generates a cyclic subgroup which is dense in $Gal(\bar{\mathbb{F}}_p/\mathbb{F}_p)$ (with respect to the Krull topology). Let $x \in \bar{\mathbb{F}}_p$. Then $x \in \mathbb{F}_{p^n} \Leftrightarrow \phi^n(x) = x$.

[1] All this can be proved for an arbitrary field using Zorn Lemma. Anyway, in the sequel we will work with the algebraic closure of a finite field $\mathbb{F}_p$, which can also described explicitly: it is $\cup_{n\geq 1}\mathbb{F}_{p^n}$.

1.3 Commutative algebra

Definition 1.1. Let R be a ring. A sequence of prime ideals $P_n \subset P_1 \subset \ldots \subset P_0$, where all inclusions are proper, is called a *chain* of prime ideals. The *height* of a prime ideal P, denoted by $h(P)$, is the supremum of the lengths of all chains of prime ideals with $P_0 = P$. The *Krull dimension* of R is the supremum of the heights of all prime ideals in R. It will be denoted by $dim(R)$.

Example 1.2. If K is a field, $K[x_1, \ldots, x_n]$ is a ring of dimension n. A proof of this fact for the case $n = 2$, which will be enough for us, can be found in [11]. For the general proof, see [15].

Observe that an integral domain R has dimension 1 if and only if every non zero prime ideal of R is maximal. Hence principal ideal domains that are not fields have dimension 1.

1.3.1 Localization

Definition 1.2. Let R be a ring. A subset $S \subset R$ is called *multiplicative* if:

1. $1 \in S$;

2. $a \in S, b \in S \Rightarrow ab \in S$

If S is a multiplicative subset of a ring R, consider the set $R \times S$, and the relation: $(a, s) \equiv (b, t) \Leftrightarrow \exists \sigma \in S : \sigma(at - bs) = 0$. $\equiv$ is an equivalence relation.
Let $S^{-1}R$ be the set of equivalence classes of $R \times S$. Denote by $\frac{a}{s}$ the equivalence class of (a, s).
The following operations are well defined and provide $S^{-1}R$ with a ring structure:

$$\frac{a}{s} + \frac{b}{t} = \frac{at + bs}{ts}$$
$$\frac{a}{s}\frac{b}{t} = \frac{ab}{ts}$$

$S^{-1}R$ with this ring structure is called the *localization* of R at S. The map

$$j_S : R \to S^{-1}R,\ j(a) = a/1$$

is a ring homomorphism. It will be denoted simply by j if there is no risk of confusion.
Note that j is injective if R is an integral domain. Hence in this case we can regard R as a subring of $S^{-1}R$.

Example 1.3. 1. Let D be an integral domain, $S = D \setminus \{0\}$. Then $S^{-1}D$ is the field of fractions of D.

2. Let R be a ring, $S = D \setminus P$, where P is a fixed prime ideal of R. Then $S^{-1}R$ is called the localization of R at P, and is denoted by R_P.

Proposition 1.1. *Let R be a ring, $S \subseteq R$ a multiplicative subset. Then, the map j induces a bijection:*

$$\begin{aligned} j^* : Spec(S^{-1}R) &\to \{P \in Spec(R) : P \subset R \setminus S\} \\ P &\mapsto j^{-1}(P) \end{aligned}$$

The inverse of j^* is the map sending $P \in Spec(R),\ P \cap S = \varnothing$ to the ideal in $Spec(S^{-1}R)$ generated by $j(P)$.

Proof. See [11, p. 60]. □

Hence, prime ideals in the localization $S^{-1}R$ correspond bijectively to prime ideals in R which do not intersect S. As a consequence, we have the following:

Corollary 1.1. *Let R be a ring, $P \in Spec(R)$. Then R_P has only one maximal ideal, generated by $j(P)$. Moreover, $dim(R_P) = h(P)$. In particular, if $dim(R) = 1$ and $P \in Max(R)$, then $dim(R_P) = 1$.*

Rings with only one maximal ideal are of the uttermost importance, and they deserve a name:

Definition 1.3. A ring with only one maximal ideal is called a *local ring*.

The above corollary states that the localization of a ring at a prime ideal is a local ring.
Let R be a local ring with maximal ideal M. Take $x \in R$. Reminding that any proper ideal in a ring is contained in a maximal ideal, it's easy to show that $x \in M \Leftrightarrow x$ is not a unit.

Definition 1.4. Let A, B be two rings, $A \subseteq B$. An element $b \in B$ is called *integral* over A if it satisfies one of the following equivalent properties: (see [11, p. 12])

1. b is the zero of a *monic* polynomial in $A[x]$;
2. $A[b]$ is a finitely generated $A-$module;
3. $bM \subseteq M$ for a finitely generated $A-$submodule M of B.

The set of all elements in B which are integral over A is a ring, called the *integral closure* of A in B. An integral domain D is called *integrally closed* if it equals its integral closure in its field of fractions.

Proposition 1.2. *Let D be an integral domain. Then the following properties are equivalent:*

1. *D is integrally closed;*
2. *D_P is integrally closed for all $P \in Spec(D)$;*
3. *D_M is integrally closed for all $M \in Max(D)$.*

Proof. See [11, p. 74]. □

Property (2) means that being integrally closed is a *local* property of a domain: in general, a property of a ring is called local when it is satisfied by the ring if and only if it holds for all the localizations of that ring at its prime ideals. Here are other examples of local properties. For the proof, see [11].

Proposition 1.3. *Noetherianity is a local property; being a PID is a local property.*

Proposition 1.4. *Let D be a noetherian local domain of dimension 1. Then D is a PID if and only if it is integrally closed.*

Proof. It is easy to show that factorial domains are always integrally closed. In particular, $PIDs$ are integrally closed.
Conversely, suppose that D is integrally closed. Let M be the maximal ideal of D. Take $x \in M, x \neq 0$. If $(x) = M$, then all prime ideals of D are principal, hence (see [11]) D is a PID.
If $(x) \neq M$, observe first of all that since D is noetherian the ideal (x) contains a power of M.
In fact, any nonzero ideal in a noetherian ring contains a product nonzero of prime ideals (to show it, suppose it's false, and take a maximal element in the family of ideals not containing any product of nonzero prime ideals and obtain a contradiction).
Therefore, there exists $n \in \mathbb{N}$ such that $M^n \subseteq (x)$, and $M^{n-1} \nsubseteq (x)$. Take $y \in M^{n-1} \setminus (x)$ Then, if K denotes the field of fractions of the domain D, $y/x \in K \setminus D$. Since D is integrally closed y/x is not integral over A, so $(y/x)M \nsubseteq M$, as M is a finitely generated $D-$module (D is noetherian).
By construction, we have $yM \subset M^n \subset (x) \Rightarrow (y/x)M \subset D$. Therefore, $(y/x)M$ is an ideal of D not contained in M, so $(y/x)M = D$. Hence D is a PID. □

Definition 1.5. A noetherian local domain whose maximal ideal is principal is called a *discrete valuation ring* (DVR).

Let D be a discrete valuation ring with maximal ideal $M = (\pi)$ (such a π is called a *uniformizing element* of D), $x \in D \setminus \{0\}$. Then, noetherianity of D implies that there exists an integer $n_x \geq 0$ such that $M^{n_x} \subseteq (x)$, and $M^{n_x - 1} \nsubseteq (x)$. Then, $x = \pi^{n_x} u$, where u is a unit in D. The integer n_x does not depend on the choice of the uniformizing parameter π. Hence, we can define:

$$\begin{aligned} v : D &\to \mathbb{Z} \\ x &\mapsto n_x \ \text{ if } x \neq 0 \\ 0 &\mapsto \infty \end{aligned}$$

This map satisfies the following properties:

1. $v(x) \geq 0$, and $v(x) = \infty \Leftrightarrow x = 0$
2. $v(xy) = v(x) + v(y)$
3. $v(x + y) \geq min\{v(x), v(y)\}$

A map with the above properties is called a *discrete valuation* (sometimes just valuation) on D. It can be extended to the field of fractions K of D by setting: $v(x/y) = v(x) - v(y)$. The valuation defined in this way on a discrete valuation ring D (or on its field of fractions) will be called the *standard valuation* on D.

Example 1.4. Let $p \in \mathbb{Z}$ be a prime. The inverse limit $\mathbb{Z}_p = \varprojlim(\mathbb{Z}/p^n\mathbb{Z})$ (where $n \in \mathbb{N}$ and the morphisms are projections) is called the *ring of p-adic integers*. It is a noetherian local ring, with maximal ideal $p\mathbb{Z}_p$, which is principal. Hence $\mathbb{Z}_p$ is a DVR. Its field of fractions, $\mathbb{Q}_p$, is called the field of *p-adic numbers*. $\mathbb{Q}_p$ has characteristic 0, so $\mathbb{Q} \subseteq \mathbb{Q}_p$. [2].
Let $f(x, y, z) \in \mathbb{Q}[x, y, z]$ be an homogeneous polynomial. If the equation $f(x, y, z) = 0$ has a nontrivial solution $(x_0, y_0, z_0) \neq (0, 0, 0)$ in $\mathbb{Q}^3$, clearly it has a nontrivial solution in $\mathbb{Q}_p^3$ for every prime p, and also a nontrivial solution in $\mathbb{R}^3$. The highly non trivial fact is that in some lucky cases the converse is true:

Theorem 1.1. *(Hasse-Minkowski, local-global principle) Let $f(x, y, z) \in \mathbb{Q}[x, y, z]$ be an homogeneous polynomial of degree 2. If the equation $f(x, y, z) = 0$ has a nontrivial solution in every $\mathbb{Q}_p$ and a nontrivial solution in $\mathbb{R}$, then it has a nontrivial solution in $\mathbb{Q}$.*

A nice account of the proof is given in [17].

Definition 1.6. A *Dedekind domain* is a noetherian, one dimensional, integrally closed domain.

Dedekind domains enjoy the fundamental property of *unique factorization of ideals*: each non trivial ideal I in a Dedekind domain D can be written uniquely (up to order) as a product of prime ideals. This is proved in any introductory book to Algebraic Number Theory, for example [16].

Example 1.5. Let $\mathbb{Q} \subseteq K$ be a finite field extension. Then K is called a *number field*. The set of elements in K which are integral over $\mathbb{Z}$ is a ring, called the *ring of algebraic integers* of K. We will sometimes call it merely the ring of integers of K.
This ring is a Dedekind domain. The study of the properties of rings arising in this way is a major issue of Algebraic Number Theory; we will recall some of its basic results in the following section.

1.4 Algebraic Number Theory

Let $K/\mathbb{Q}$ be a finite field extension of degree n. The integral closure of $\mathbb{Z}$ in K:

$$\mathcal{O}_K = \{x \in K : \exists\, a_0, \ldots, a_{n-1} \in \mathbb{Z} : x^n + a_{n-1}x^{n-1} + \ldots + a_0 = 0\}$$

called the *ring of algebraic integers* of K enjoys the following properties (see [16]):

1. $\mathcal{O}_K$ is a free $\mathbb{Z}-$module of rank $n = [K : \mathbb{Q}]$;

2. $\mathcal{O}_K$ is a Dedekind domain.

[2]See [16] for a detailed study of *p-adic* fields.

Let $\mathcal{O}_K$ be the ring of algebraic integers of a number field K. Take $p \in \mathbb{Z}$ prime. The (non trivial) ideal $P = p\mathcal{O}_K \subseteq \mathcal{O}_K$ can be factored as:

$$P = \prod_i P_i^{e_i}$$

with $P_i \in Spec(\mathcal{O}_K)$. We say that the primes P_i are *over* p. The integer e_i is called the *ramification index* of P_i over p. Clearly $P_i \supseteq P\ \forall i$ and $P_i \cap \mathbb{Z} = p$. Therefore, $\mathcal{O}_K/P_i$ is a finite field extension of $\mathbb{Z}/p\mathbb{Z} \cong \mathbb{F}_p$. The degree of the extension is denoted by f_i and is called the *residual degree* of P_i over p.
The integers e_i, f_i and n are related by the following fundamental relation:

$$\sum_i e_i f_i = n$$

An easy proof of this equality can be found in [8].
Notice that, as a consequence of the above equality, a fixed prime $p \in \mathbb{Z}$ has at most n prime ideals of $\mathcal{O}_K$ above it.
Let $I \subseteq \mathcal{O}_K$ be a non zero ideal, $a \in I$. Then $a\mathcal{O}_K \subseteq I \subseteq \mathcal{O}_K$. As $a\mathcal{O}_K$ and $\mathcal{O}_K$ are free $\mathbb{Z}-$modules of rank n, I is also a free $\mathbb{Z}-$module of rank n. Therefore, there is a $\mathbb{Z}-$basis $e_1, \ldots, e_n$ of $\mathcal{O}_K$ and there are integers $1 < d_1 \mid d_2 \ldots \mid d_n$ such that $d_1 e_1, \ldots, d_n e_n$ is a $\mathbb{Z}-$basis of I. Hence $\mathcal{O}_K/I$ is finite (of cardinality $d_1 \ldots d_n$). We define:

$$N(I) = |\mathcal{O}_K/I|$$

$N(I)$ is called the norm of the ideal I. It can be shown that, if I, J are non zero ideals of $\mathcal{O}_K$, $N(IJ) = N(I)N(J)$. We say that the norm is *multiplicative*.
This relation remains true if we replace $\mathcal{O}_K$ by an arbitrary Dedekind domain D such that the quotient D/I is a finite set for each non zero ideal I of D, so that the definition of the norm of an ideal still makes sense. We call it a Dedekind domain with *finite quotients*. See [11] for a proof of these facts.
If $a \in \mathcal{O}_K, a \neq 0$, we define $N(a) = N((a))$.
Let us study in more detail two rings of algebraic integers which will be very useful later.

1.4.1 The ring $\mathbb{Z}[i]$

The ring $\mathbb{Z}[i] = \{a + bi, a, b \in \mathbb{Z}\}$ is the ring of algebraic integers of the field $\mathbb{Q}(i)$. Besides being a Dedekind domain, it is a PID (actually, it's also a euclidean domain; see [8] for a proof of these facts).[3] Let $x = a + ib \in \mathbb{Z}[i]$. Then $N(x) = x\bar{x} = a^2 + b^2$. In fact, with respect to the basis $(1, i)$ of $\mathbb{Z}[i]$, a basis of (x) is: $(a, b), (-b, a)$. It follows that $|\mathbb{Z}[i]/x\mathbb{Z}[i]| = det\left(\begin{smallmatrix} a & -b \\ b & a \end{smallmatrix}\right) = a^2 + b^2$.
The fact that the norm is multiplicative implies that units in $\mathbb{Z}[i]$ must have norm 1. The only elements with this property in $\mathbb{Z}[i]$ are $1, -1, i, -i$, which are obviously units.

[3] Warning: this is absolutely far from true in general: for example, the only quadratic imaginary fields, that is, fields of the form $\mathbb{Q}[\sqrt{d}]$, $d < 0$, whose ring of algebraic integers is a PID are those corresponding to $d = -163, -67, -43, -19, -11, -7, -3, -2, -1$. We also mention, *en passant*, that for $d = -19$ we obtain an example of a ring which is a PID but not a euclidean domain.

Let us determine explicitly the prime ideals in $\mathbb{Z}[i]$ (which are ideals generated by prime elements, as $\mathbb{Z}[i]$ is a PID).
There is only one prime of norm 2 up to associates, namely $1+i$.
Assume now that $\pi \in \mathbb{Z}[i]$ is prime with norm different from 2. Then $\pi \mid \pi\bar{\pi} = N(\pi) \in \mathbb{Z}$, hence, as π is prime, π divides one of the prime factors of $N(\pi)$, call it p. Hence, $N(\pi) \mid p^2$, so $N(\pi) = p$ or p^2.
If $p \equiv 3 \pmod 4$ then the equation $a^2+b^2 = p$ has no integer solutions, hence $N(\pi) = p^2$ and p is associated with π.
If $p \equiv 1 \pmod 4$, by Proposition 3.5 there exist $a, b \in \mathbb{Z}$ such that $a^2+b^2 = 1 \Rightarrow (a+bi)(a-bi) = p = \pi\bar{\pi}$. Now, $\pi' = a+bi$ is prime in $\mathbb{Z}[i]$. In fact, suppose $\pi' = \alpha\beta$. Then $p = N(\pi') = N(\alpha)N(\beta) \Rightarrow N(\alpha) = 1$ or $N(\beta) = 1$. So either α or β is a unit, hence π' is prime. It follows that $\pi' \mid \pi$ or $\pi' \mid \bar{\pi}$. Hence π is associated with π' or to $\bar{\pi}'$. We've obtained the following:

Proposition 1.5. *Up to associates, primes $\pi \in \mathbb{Z}[i]$ are of the form:*

1. $\pi = 1+i$
2. $\pi = p$, *with p prime in $\mathbb{Z}$, $p \equiv 3 \pmod 4$*
3. $\pi = a+ib$, *with $a^2+b^2 = p$, p prime, $p \equiv 1 \pmod 4$.*

Each prime in $\mathbb{Z}$ is associated with its opposite; anyway, each set of associated primes contains one and only one *positive* element.
In $\mathbb{Z}[i]$ each prime is associated with three other elements. We want to find a way to choose a "canonical" element among the set of associates of a fixed prime.

Definition 1.7. A prime $\pi \in \mathbb{Z}[i]$ is called *primary* if $\pi \equiv 1 \pmod{2+2i}$

An easy computation shows that $\pi = a+ib$ is primary if and only if $a \equiv 1 \pmod 4$, $b \equiv 0 \pmod 4$ or $a \equiv 3 \pmod 4$, $b \equiv 2 \pmod 4$.
More generally, we will say that a nonunit $\alpha \in \mathbb{Z}[i]$ is primary if $\alpha \equiv 1 \pmod{2+2i}$.

Proposition 1.6. *Let $\pi = a+bi \in \mathbb{Z}[i]$ prime, $(\pi) \neq (1+i)$. Then there exists a unique π' primary associated with π.*

Proof. As a and b cannot be both even, there is a unit ϵ such that $\epsilon\pi = a'+ib'$ with a' odd, b' even. Then either $\epsilon\pi$ or $-\epsilon\pi$ is primary.
If ϵ, ϵ' are units such that $\epsilon\pi$ and $\epsilon'\pi$ are primary, then $2(1+i) \mid (\epsilon-\epsilon')$ (as $(1+i) \nmid \pi$). This forces $\epsilon = \epsilon'$. ∎

1.4.2 The ring $\mathbb{Z}[\omega]$

Let $\omega = (-1+\sqrt{-3})/2$. Then $\mathbb{Z}[\omega] = \{a+b\omega, a, b \in \mathbb{Z}\}$ is the ring of algebraic integers of the field $\mathbb{Q}(\omega)$. It shares many properties with the ring $\mathbb{Z}[i]$, which can be proved in a very similar way. We will list the fundamental properties of $\mathbb{Z}[\omega]$ without proof; the interested reader can find them in [8], or adapt the previous proofs.

1. $\mathbb{Z}[\omega]$ is a euclidean domain;

2. if $x = a + b\omega \in \mathbb{Z}[\omega]$, $N(x) = x\bar{x} = a^2 - ab + b^2$;
3. the units in D are $\pm 1, \pm\omega, \pm\omega^2$;
4. up to associates, primes $\pi \in \mathbb{Z}[i]$ are of the form:
   (a) $\pi = 1 - \omega$
   (b) $\pi = p$, with p prime in $\mathbb{Z}$, $p \equiv 2 \pmod 3$
   (c) $\pi = a + \omega b$, with $a^2 - ab + b^2 = p$, p prime, $p \equiv 1 \pmod 3$.

Definition 1.8. A prime $\pi \in \mathbb{Z}[\omega]$ is called *primary* if $\pi \equiv 2 \pmod 3$

As in the case of $\mathbb{Z}[i]$, we find that primes in $\mathbb{Z}[\omega]$ which are not associated with $1 - \omega$ have a unique primary associate.

1.4.3 The Ideal Class Group of a Number Field

Proofs of the statements in this section can be found in [16].
Let K be a number field with ring of algebraic integers $\mathcal{O}_K$. If I, J are non zero ideals of $\mathcal{O}_K$, we say that I and J are equivalent, and we write $I \sim J$, if there exist two principal ideals $(\alpha), (\beta)$ such that $(\alpha)I = (\beta)J$. This is easily seen to be an equivalence relation. The set of equivalence classes of ideals of $\mathcal{O}_K$ is denoted by $Cl(K)$.
Let $\bar{I}$ denote the equivalence class of an ideal I. The composition law: $\overline{I} \cdot \overline{J} = \overline{IJ}$ is well defined, and turns $Cl(K)$ into an abelian group, called the *ideal class group* of K. The identity element of the group is the equivalence class of all principal ideals of $\mathcal{O}_K$.
Clearly, the ideal class group of K is trivial if and only if $\mathcal{O}_K$ is a PID. We can consider the size of the ideal class group as a measure of the extent to which $\mathcal{O}_K$ fails to be a PID.
The rings of algebraic integers of number fields enjoy the following fundamental property (which is not true for arbitrary Dedekind domains):

Theorem 1.2. *The ideal class group of a number field is finite.*

This important result can be proven using geometric ideas due to Minkowski.
We will also need the following more general result, whose proof needs more advanced tools:

Proposition 1.7. *Let $\mathcal{O}_K$ be the ring of algebraic integers of a number field K. Let M be a non zero ideal of $\mathcal{O}_K$, and $C_M = \{I : I \text{ ideal of } \mathcal{O}_K,\ (I, M) = (1)\}/\sim$ where $I \sim J$ if there exist two principal ideals $(\alpha), (\beta)$ with $\alpha \equiv 1 \pmod M$, $\beta \equiv 1 \pmod M$, such that $(\alpha)I = (\beta)J$. Then the composition law $\overline{I} \cdot \overline{J} = \overline{IJ}$ is well defined and turns C_M into a group; moreover, this group is finite.*

Observe that for $M = \mathcal{O}_K$ we obtain the finiteness of the class number.

Chapter 2

Algebraic curves

2.1 Affine curves

Let K be a field. The *affine space* of dimension n over K is the set:

$$\mathbb{A}^n(K) = \{(x_1, \dots, x_n) : \ x_i \in K \ \forall i = 1, \dots, n\}$$

In the sequel, we will be mainly interested in the case $n = 2$. In this case, the affine space is called *affine plane*.

Definition 2.1. Let $f(x, y) \in K[x, y]$. The set $C(K) = \{(x, y) \in \mathbb{A}^2(K) : f(x, y) = 0\}$ is called an *affine (plane) curve*.
$C(K)$ is said *geometrically irreducible* if f is an irreducible polynomial in $\bar{K}[x, y]$.
The degree of the polynomial $f(x, y)$ is called the *degree* of the curve.

Remark 2.1.

1. Let $C(\bar{K}) = \{(x, y) \in \mathbb{A}^2(\bar{K}) : f(x, y) = 0\}$, with $f(x, y) \in \bar{K}[x, y]$. Of course, there are many other polynomials whose zeros are exactly the points of $C(\bar{K})$: it is enough to consider $g(x, y) = \lambda f(x, y), \ \lambda \in \bar{K}^*$. We will say that $C(\bar{K})$ is defined over K if there exists a $\lambda \in \bar{K}^*$ such that $g(x, y) = \lambda f(x, y) \in K[x, y]$. In this case, we will denote: $C(K) = \{(x, y) \in \mathbb{A}^2(K) : g(x, y) = 0\}$.
   If C is defined over K, the Galois group $Gal(\bar{K}/K)$ acts on $C(\bar{K})$ in a natural way: for $\sigma \in Gal(\bar{K}/K), \ P = (a, b) \in C(\bar{K}), \ P^\sigma = (\sigma(a), \sigma(b))$. The points of $C(K)$ are exactly the fixed points of this action.
   More generally, let $P = (a, b) \in C(\bar{K})$. We say that P is *defined over* a field $L \supseteq K$ if $K(P) = K(a, b) \subset L$. Clearly P is defined over L if and only if it is fixed by the action of $Gal(\bar{K}/L)$.

2. We will often denote an affine curve simply by C, without mentioning explicitly the field K.

Definition 2.2. Let $f(x, y) \in K[x, y]$ and $C(K) = \{(x, y) \in \mathbb{A}^2(K) : f(x, y) = 0\}$ be an affine plane curve. The ring $K[C] = K[x, y]/(f)$ is the *affine coordinate ring* of C.
If C is a geometrically irreducible curve then its affine coordinate ring is an integral domain. Its field of fractions is called the *field of (rational) functions* of C, and is denoted by $K(C)$.

Remark 2.2. Since we will only deal with geometrically irreducible curves, from now on the word affine curve will indicate, unless the contrary is explicitly stated, a *geometrically irreducible curve*. The expression *irreducible curve* will also be used to refer to a geometrically irreducible curve. However, this terminology is not standard: usually a curve $C(K) = \{(x, y) \in \mathbb{A}^2(K) : f(x, y) = 0\}$ is called irreducible if $f(x, y)$ is irreducible in $K[x, y]$.

Given an affine curve C, our aim is to study its geometric properties looking at algebraic properties of the associated coordinate ring.
The following theorem is the fundamental example of the close correspondence between the structure of geometric objects and properties of algebraic structures naturally associated with them.

Theorem 2.1. *(Hilbert Nullstellensatz) Let $f(x, y) \in K[x, y]$ and $C(K) = \{(x, y) \in \mathbb{A}^2(K) : f(x, y) = 0\}$ be an affine plane curve with coordinate ring $K[C] = K[x, y]/(f)$.*

1. *Every maximal ideal of $K[C]$ is of the form $\psi_P = ker(ev_P)/(f)$ for some $P = (a, b) \in C(\bar{K})$, where $ev_P : K[x, y] \to \bar{K}$, $g(x, y) \mapsto g(a, b)$ is the* evaluation map *at the point P.*

2. *If $K = \bar{K}$ is algebraically closed, then the map $P \mapsto \psi_P$ is a bijection between $C(K)$ and $Max(K[C])$.*

Proof. See [15] (in which a proof of a more general result is given). □

This Theorem is the first "bridge" between algebra and geometry. It states that, over an algebraically closed field, points on the geometric object $C(K)$ correspond bijectively to maximal ideals in the associated coordinate ring $K[C]$.
If the field K is not algebraically closed, things are not so simple. For example, the ideal (x^2+1) in $\mathbb{R}[x]$ is maximal, but doesn't correspond to any point in $\mathbb{R}$. Anyway, the above result tells us that maximal ideals in $K[C]$ are kernels of evaluations at points on the curve defined over algebraic field extensions of K.
In general, evaluations at different points can correspond to the same maximal ideal in $K[C]$. In fact, let $\sigma \in Gal(\bar{K}/K)$, $P, Q \in C(\bar{K})$ such that $P = \sigma(Q)$. Then for each $g(x, y) \in K[x, y]$ we have: $ev_P(g) = \sigma \circ ev_Q(g) = 0 \Leftrightarrow ev_Q(g) = 0$. Hence $\psi_P = \psi_Q$.
Conversely, if $\psi_P = \psi_Q$ then $K(P) \cong K[C]/\psi_P = K[C]/\psi_Q \cong K(Q)$. The isomorphism between the two $K-$algebras $K(P)$ and $K(Q)$, which sends P to Q, can be extended to an isomorphism $\sigma \in Gal(\bar{K}/K)$ (because of the universal property of the algebraic closure). Thus P and Q are in the same orbit with respect to the action of $Gal(\bar{K}/K)$.
To sum up: maximal ideals in $K[C]$ correspond bijectively to orbits of points in $C(\bar{K})$ with respect to the action of $Gal(\bar{K}/K)$.
Observe also that, if $P \in C[\bar{K}]$, the orbit of P contains $[K(P) : K]$ different points, since the stabilizer of P is $Gal(\bar{K}/K(P))$, which has index $[K(P) : K]$ in $Gal(\bar{K}/K)$. (Section 1.2)

Example 2.1. Let $C(K)$ be a curve defined on $K = \mathbb{F}_p$. Then, if N_m denotes the number of points on $C(\bar{K})$ defined over $\mathbb{F}_{p^m}$ and $b_d = |\{M \in Max(K[C]) : [K[C]/M : \mathbb{F}_p] = d\}|$, we have:

$$N_m = \sum_{d|m} d b_d$$

In fact, $P = (a, b) \in C(\bar{K})$ is defined over $\mathbb{F}_{p^m}$ if and only if ψ_P is such that $d = [K[C]/\psi_P : \mathbb{F}_p] = [K(P) : \mathbb{F}_p] \mid m$. Moreover, each maximal ideal ψ_P corresponds to an orbit containing $[K(P) : \mathbb{F}_p] = [K[C]/\psi_P : \mathbb{F}_p] = d$ different points defined over $\mathbb{F}_{p^d} \subseteq \mathbb{F}_{p^m}$. Hence the total number of points on the curve defined on $\mathbb{F}_{p^m}$ is exactly $\sum_{d|m} d b_d$.

Definition 2.3. The *dimension* of an affine curve $C(K)$ is the Krull dimension of the ring $\bar{K}[C]$.

Recall that $K[x_1, \ldots, x_n]$ is a ring of dimension n (Example 1.2); in particular, $dim\ \bar{K}[x, y] = 2$. As a consequence we obtain the following

Lemma 2.1. *Affine curves have dimension 1.*

Proof. Recall that curve is always, for us, an irreducible curve. Its coordinate ring is $\bar{K}[C] = \bar{K}[x, y]/(f)$, where f is an irreducible polynomial. Let I be a non zero prime ideal in $\bar{K}[C]$. Then $I = \pi(J)$, where $J \in Spec(\bar{K}[x, y])$, $J \supset (f)$. Hence $0 \subset (f) \subset J$ is a chain of length 2 in $\bar{K}[x, y]$; it follows that $J \in Max(\bar{K}[x, y])$ and $\pi(J) = I \in Max(\bar{K}[C])$. □

We're now going to define the notion of smooth point and tangent line to a point of a curve. We will first begin with a geometric definition, borrowed from differential geometry; after that, according to our philosophy, we will look for a characterisation of smoothness in terms of the algebraic properties of the coordinate ring of the curve.

Definition 2.4. Let $f(x, y) \in K[x, y]$. A point $P = (a, b) \in C(\bar{K}) = \{(x, y) \in \mathbb{A}^2(\bar{K}),\ f(x, y) = 0\}$ is called a *singular point* if $\partial f/\partial y(P) = \partial f/\partial x(P) = 0$.
A point which is not singular is called *smooth*. A curve is smooth if all of its points in $\bar{K}$ are smooth.
The affine space $T_P = \{(x, y) \in \mathbb{A}^2(\bar{K}),\ \partial f/\partial y(P)(x - a) + \partial f/\partial y(P)(y - b) = 0\}$ is called *tangent line* to the curve at P.

It is well known that the geometric notion of smooth point on a curve is a *local* notion, which depends only on the structure of the curve in an arbitrarily small neighbourhood of the point. The ring-theoretic translation of the idea of "neighbourhood of a point" is that of *localisation* at a *prime ideal*.
Let $P = (a, b)$ be a point on a curve C. Let $\bar{K}[C]_P$ denote the localisation of $\bar{K}[C]$ at the maximal ideal $(x - a, y - b)$. This is called the *local ring* of C at P.

Proposition 2.1. *P is nonsingular if and only if $\bar{K}[C]_P$ is a discrete valuation ring.*

Proof. As $\bar{K}[C] = \bar{K}[x, y]/(f)$ is noetherian, the localisation $\bar{K}[C]_P$ is noetherian (by Proposition 1.3). Moreover we know that the localisation at a maximal ideal of a dimension 1 ring has dimension 1 (Corollary 1.1).
Let m_P be the only maximal ideal of $\bar{K}[C]_P$. By definition, $\bar{K}[C]_P$ is a discrete valuation ring if and only if m_P is principal. Hence, we need to show that $P = (a, b)$ is

smooth if and only if m_p is principal. We may assume, after a linear change of variables, that $P = (a, b) = (0, 0)$. Let us also assume, without loss of generality, that $\partial f/\partial y(0, 0) = \delta \neq 0$. Then the Taylor expansion of $f(x, y)$ at $(0, 0)$ can be written in the form:

$$\begin{aligned} f(x,y) &= \partial f/\partial x(0,0)x + \delta y + \text{(higher order terms)} = \\ &= \sum_{i=1}^{n} b_i x^i + y(\delta + g(x,y)) \end{aligned}$$

where $b_i \in \bar{K}$, $g(x, y) \in \bar{K}[x, y]$, $g(0, 0) = 0$. In $\bar{K}[C]_P$, the above equality becomes:

$$y(\delta + g(x,y)) = -\sum_{i=1}^{n} b_i x^i$$

(by an abuse of notation, we still denote by x, y the images in $\bar{K}[C]_P$ of $x, y \in \bar{K}[x, y]$). Now $g(0, 0) = 0 \Rightarrow g(x, y) \in m_p$. As δ is a unit, $g(x, y) + \delta \notin m_p \Rightarrow g(x, y) + \delta$ is a unit. Hence, in $\bar{K}[C]_P$, we find that y belongs to the ideal generated by x. Since x and y generate m_p, we conclude that m_p is principal.

Conversely, assume that $\bar{K}[C]_P = (z)$. As $\bar{K}[C]_P = (x, y)$ we have:

$$\begin{aligned} ux + vy &= z \quad \text{for } u, v \in m_p \\ x &= zs \quad \text{for } s \in m_p \\ y &= zr \quad \text{for } r \in m_p \end{aligned}$$

Since $\bar{K}[C]_P$ is a domain, $us + vr = 1$. Then either s or r is a unit in $\bar{K}[C]_P$. Assume that s is a unit. Since $rx - sy = 0$ in $\bar{K}[C]_P$, we can find polynomials $\tilde{r}, \tilde{s}, g \in \bar{K}[x, y]$ such that $f(x, y)g(x, y) = \tilde{r}(x, y)x - \tilde{s}(x, y)y$ and $\tilde{s}(x, y)$ has non-trivial constant term. By comparing the coefficients of the monomial y on the two sides we conclude that $\partial f/\partial y(0, 0) \neq 0$ □

What determines whether a point P on a curve C is smooth or singular is thus the local behaviour of the coordinate ring $\bar{K}[C]$ at the point P. By collecting local information at each point of the curve we obtain the following global characterisation of smooth curves in terms of their coordinate ring:

Proposition 2.2. *An affine curve $C(\bar{K})$ is smooth if and only if its coordinate ring $\bar{K}[C]$ is a Dedekind domain.*

Proof. The ring $\bar{K}[C]$ is noetherian and 1-dimensional (Lemma 2.1). Thus it is a Dedekind domain if and only if it is integrally closed. Proposition 1.2 states that $\bar{K}[C]$ is integrally closed if and only if $\bar{K}[C]_P$ is integrally closed for every $P \in C(\bar{K})$ (maximal ideals correspond to points on the curve, by Theorem 2.1). By Proposition 1.4 $\bar{K}[C]_P$ is integrally closed if and only it is a PID. We just proved that this is equivalent to the fact that P is a smooth point on C. Therefore, $\bar{K}[C]$ is a Dedekind domain if and only if all points of $C(\bar{K})$ are smooth, that is, if C is a smooth curve. □

Corollary 2.1. *Let K be a field, $C(K)$ an irreducible curve. If $\bar{K}[C]$ is a Dedekind domain, then $K[C]$ is also a Dedekind domain.*

Proof. $K[C]$ has dimension 1 and is noetherian. Let us show that every localisation of $K[C]$ at a maximal ideal is a PID. This will prove the theorem.
By the Theorem 2.1 we know that each maximal ideal M of $K[C]$ is of the form $M = ker(ev_P)/(f)$, $P = (a, b) \in C(\bar{K})$. Let $g(x)$ denote the minimal polynomial of a over K. Then there exists $h \in K[x, y]$ such that $ker(ev_P) = (g(x), h(x, y))$. In fact, let $I = g(x)K[x, y]$. Then:

$$K[x, y]/I \cong (K[x]/g(x))[y] \cong L[y]$$

where $L = K(a)$. The ideal M/I is maximal in $K[x, y]/I$, which is a PID; hence, M/I is generated by a $\pi(h(x, y))$, $h \in k[x, y]$, and $M = (g(x), h(x, y))$.
Now, since $f(x, y) \in ker(ev_P)$ there exist $\alpha(x, y)$, $\beta(x, y) \in K[x, y]$ such that

$$f(x, y) = \alpha(x, y)g(x) + \beta(x, y)h(x, y)$$

Therefore in $K[C]_P$ $\bar{g}$ divides $\bar{h}\bar{\beta}$ (where $\bar{g}$ denotes the class of g in $K[C]_P$). Since by hypothesis $\bar{K}[C]$ is a Dedekind domain, which is equivalent to the fact that the curve C is smooth, we may assume that $\partial f/\partial y(P) \neq 0$. We have

$$\partial f/\partial y(x, y) = \partial\alpha/\partial y(x, y)g(x) + \partial\beta/\partial y(x, y)h(x, y) + \beta(x, y)\partial h/\partial y(x, y)$$

and so: $\partial f/\partial y(P) = \beta(P)\partial h/\partial y(P)$. Hence $\beta(P) \neq 0 \Rightarrow \beta \notin ker(ev(P)) \Rightarrow \bar{\beta}$ is a unit in $K[C]_P$. As $\bar{g}$ divides $\bar{h}\bar{\beta}$, we deduce that $\bar{g}$ and $\bar{h}$ are associates, and so each of them generates $K[C]_P$, which is therefore principal. □

2.2 Projective curves

The *projective space* of dimension n over a field K is the set:

$$\mathbb{P}^n(K) = \{[x_0, \dots, x_n] : x_i \in K \ \forall i = 1, \dots, n\}$$

Recall that coordinates in the projective space are defined only up to a non zero multiplicative constant: $[x_0, \dots, x_n] = [y_0, \dots, y_n]$ if and only if there exists $\lambda \in K^*$ such that $x_i = \lambda y_i \ \forall i = 0, \dots, n$. The projective space of dimension 2 is called *projective plane.*
Algebraic curves in the projective plane are defined in the same way as in the affine case:

Definition 2.5. Let $f(x, y, z) \in K[x, y, z]$ be an homogeneous polynomial. The set $C(K) = \{[x, y, z] \in \mathbb{P}^2(K) : f(x, y, z) = 0\}$ is called a *projective (plane) curve.* We say that $C(K)$ is *(geometrically) irreducible* if f is an irreducible polynomial in $\bar{K}[x, y, z]$. The *degree* of the curve $C(K)$ is the degree of the polynomial f.

Remark 2.3. 1. Everything we said in Remark 2.1 can be repeated for projective curves. Terminology and notations introduced there will be also used for projective curves.

2. Note that the definition makes sense, as the polynomial is required to be homogeneous.

3. As in the affine case, we will always assume that projective curves are (geometrically) irreducible.

4. Let $H_0 = \{[x, y, z] \in \mathbb{P}^2(K),\ x = 0\}$.
   The bijection

$$\begin{aligned}\phi_0 : \mathbb{P}^2(K) \setminus H_0 &\to \mathbb{A}^2(K)\\ [x, y, z] &\mapsto \left(\frac{y}{x}, \frac{z}{x}\right)\end{aligned}$$

   whose inverse is the function $(a, b) \mapsto [1, a, b]$, identifies the affine plane with the projective plane deprived of a "line at infinity".
   If $C(K) = \{[x, y, z] \in \mathbb{P}^2(K) : f(x, y, z) = 0\}$ is a projective curve, then $C(K) \cap (\mathbb{P}^2(K) \setminus H_0)$ corresponds, with the above bijection, to the affine curve: $C^*(K) = \{[a, b] \in \mathbb{A}^2(K) : f^*(a, b) = 0\}$, where $f^*(a, b) = f(1, a, b)$ is the dehomogenization of f with respect to x.
   Conversely, to an affine curve $C^*(K) = \{[a, b] \in \mathbb{A}^2(K) : f^*(a, b) = 0\}$ we can associate its *projective closure* $C(K) = \{[x, y, z] \in \mathbb{P}^2(K) : f(x, y, z) = 0\}$, where $f(x, y, z) = x^{deg(f^*)} f^*(y/x, z/x)$ is the homogenization of f^*.
   The same process can be carried out with respect to any other variable, substituting $H_1 = \{[x, y, z] \in \mathbb{P}^2(K),\ y = 0\}$ or $H_2 = \{[x, y, z] \in \mathbb{P}^2(K),\ z = 0\}$ to H_0. The corresponding maps will be denoted by ϕ_1, ϕ_2.
   Roughly speaking, the projective plane allows us to look at our curve from a global point of view; on the contrary, we can always work in the affine setting when we are concerned with local properties of the curve. This motivates the following definitions (the doubtful reader can check that they're all good definitions):

Definition 2.6. Let C be a projective curve, H_i such that $C \nsubseteq H_i$. The *function field* of C, denoted by $K(C)$, is the function field of the affine curve $\phi_i(C \cap (\mathbb{P}^2 \setminus H_i))$.
The *dimension* of C is the dimension of $\phi_i(C \cap (\mathbb{P}^2 \setminus H_i))$.
Let $P \in C \cap (\mathbb{P}^2 \setminus H_i)$. The *local ring* of C at P, denoted by $K[C]_P$, is the local ring of $\phi_i(C \cap (\mathbb{P}^2 \setminus H_i))$ at $\phi_i(P)$.
We say that P is *smooth* if the point $\phi_i(P)$ on the affine curve $\phi_i(C \cap (\mathbb{P}^2 \setminus H_i))$ is a smooth point. A curve is smooth if all its points are smooth.

Remark 2.4. From now on, all curves will be supposed to be smooth.

Now, take a projective curve $C(\bar{K})$ defined over K. We want to state what it means for a point $P = [x, y, z]$ on C to be defined over a field $\bar{K} \supseteq L \supseteq K$. We cannot just say, as in the affine case, that $K(P) = K(x, y, z) \subseteq L$ as x, y, z are defined only up to a multiplicative constant $\lambda \in \bar{K}^*$.
Instead, we're going to exploit the action of the Galois group $Gal(\bar{K}/K)$ on the points of the curve, which we already studied in the affine case. The action is defined in the same way: for $\sigma \in Gal(\bar{K}\ K)$, $P = (x, y, z) \in C(\bar{K})$, $P^\sigma = (\sigma(x), \sigma(y), \sigma(z))$ (it's easy to verify that this is well defined).

Definition 2.7. Let $C(\bar{K})$ be a projective curve, $P \in C$. We say that P is *defined over a field* $\bar{K} \supseteq L \supseteq K$ if $\sigma(P) = P \ \forall \sigma \in Gal(\bar{K}/L)$, that is, if P is fixed by $Gal(\bar{K}/L)$.

One of the reasons why the projective plane is a good place to do algebraic geometry is that two irreducible curves $C(\bar{K})$ and $D(\bar{K})$ in the projective plane $\mathbb{P}^2(\bar{K})$ meet as many times as possible. This is the content of the following fundamental

Theorem 2.2. *(Bezout) Let $C(\bar{K})$ and $D(\bar{K})$ be two curves of degree m and n. Then C and D intersect in exactly mn points, counted with multiplicities.*

For a precise explanation of the concept of intersection multiplicity of two curves at a given point and an elementary proof of this result see [3].

2.2.1 Divisors and Riemann-Roch Theorem

Concepts and results presented in this section are of fundamental importance for the study of curves, and have a lot of extraordinary consequences (one of them is described in Chapter 4). Unfortunately, the proof of most statements requires some rather advanced machinery, and will not be given here. The standard reference for these topics is [6]. The same results can also be stated and proved within the (maybe) simpler theory of Riemann Surfaces (see [14]).
Anyway, the reader can just trust the main theorems in this section (expecially Riemann-Roch) and see them at work in the following chapters.

Definition 2.8. Let $C(\bar{K}) = \{[x, y, z] \in \mathbb{P}^2(\bar{K}) : p(x, y, z) = 0\}$ be a projective curve defined over K. The *divisor group* of C, denoted by $Div(C)$, is the free abelian group generated by the points of $C(\bar{K})$. A divisor on C is therefore a formal sum:

$$D = \sum_{P \in C(\bar{K})} n_P P$$

where the coefficients n_P are integers, and only finitely many of them are non zero. The *degree* of the divisor is by definition:

$$deg(D) = \sum_{P \in C(\bar{K})} n_P$$

If $D = \sum_{P \in C(\bar{K})} n_P P$, and $D' = \sum_{P \in C(\bar{K})} n'_P P$, we say that $D \leq D' \Leftrightarrow n_P \leq n'_P \ \forall P \in C(\bar{K})$.
A divisor $D \geq 0$ is called a *positive* (or effective) divisor.
The action of an element $\sigma \in Gal(\bar{K}/K)$ on a divisor $D = \sum_{P \in C(\bar{K})} n_P P$ is given by:

$$(\sigma, D) \mapsto D^\sigma = \sum_{P \in C(\bar{K})} n_P P^\sigma$$

As usual, a divisor D is said to be *defined over a field* $\bar{K} \supseteq L \supseteq K$ if it is fixed by the action of every $\sigma \in Gal(\bar{K}/L)$. The set of divisors defined over L is denoted by $Div_L(C)$.

Recall that if C is a smooth (affine or projective) curve and $P \in C$, the local ring $\bar{K}[C]_P$ is a discrete valuation ring with field of fractions $\bar{K}(C)$ (Proposition 2.1). Denote by v_P the standard valuation on $\bar{K}[C]_P$, extended to its field of fractions.
We say that $f \in \bar{K}(C)^*$ has a *zero* (respectively *pole*) at P if $v_P(f) > 0$ (respectively $v_P(f) < 0$).

Definition 2.9. Let $f \in \bar{K}(C)^*$; the divisor:

$$div(f) = \sum_{P \in C(\bar{K})} v_P(f)P$$

is called the divisor *associated with* f. Two divisors D and D' are called *equivalent* if $D - D' = div(f)$ for some $f \in \bar{K}(C)^*$. If D is equivalent to D', we write $D \sim D'$.

Remark 2.5. To show that $div(f)$ is actually a divisor, it is necessary to prove that a function $f \in \bar{K}(C)^*$ has only a finite number of zeros or poles on a curve C. This is a consequence of Bezout theorem (or it can be verified by more elementary means, with the aid of discriminants of polynomials). By the same means one can also show that if $f \neq 0$ then $deg(div(f)) = 0$. Moreover $div(f) = 0 \Leftrightarrow f \in \bar{K}^*$.

Notation 2.1. Let $D \in Div(C)$. We define:

$$L(D) = \{f \in \bar{K}(C)^* : div(f) + D \geq 0\} \cup \{0\}$$

If $D \in Div_K(C)$, let

$$L_K(D) = \{f \in K(C)^* : div(f) + D \geq 0\} \cup \{0\}$$

Then $L(D)$ (resp. $L_K(D)$) is a $\bar{K}$-vector space (resp. K-vector space), whose dimension is denoted by $l(D)$ (resp. $l_K(D)$).

Remark 2.6. It can be proved that $L(D)$ is always a finite dimensional vector space; moreover, if $deg(D) < 0$, then $L(D) = 0$. In fact, let $f \in L(D)$, $f \neq 0$; then:

$$0 = deg(div(f)) \geq deg(-D) = -deg(D) > 0$$

which is absurd.

The dimension of the space $L(D)$ (or $L_K(D)$) tells us, roughly speaking, how many functions we can find on a curve, with poles (zeros) in a fixed set of points not exceeding a prescribed order (having at least a prescribed order). It is extremely useful, in order to study the properties of a curve, to have precise information about the dimension of $L(D)$ or $L_K(D)$. This is the content of the following theorem, of crucial importance in algebraic geometry:

Theorem 2.3. *(Riemann-Roch) Let $C(\bar{K})$ be a (smooth) projective curve defined over K. There exists a divisor $K_C \in Div_K(C)$ and an integer $g \geq 0$, called the* genus *of the curve, such that:*

$$l(D) - l(K_C - D) = deg(D) + 1 - g$$
$$l_K(D) - l_K(K_C - D) = deg(D) + 1 - g$$

Proof. A proof via sheaf theory is given in [6]. A more elementary treatment can be found in [3]. □

Corollary 2.2. *With the same hypotheses as above, we have:*

1. $l(K_C) = g$
2. $deg(K_C) = 2g - 2$
3. *if* $deg(D) > 2g - 2$, *then* $l(D) = l_K(D) = deg(D) + 1 - g$

Proof.
1. Use the previous theorem with $D = 0$, and Remark 2.5, which implies that $l(0) = 1$.
2. Use the previous point and Riemann-Roch, with $D = K_C$.
3. From (2) we obtain $deg(K_C - D) < 0 \Rightarrow l(K_C - D) = 0$, from remark 2.6. Now use Riemann-Roch.

□

The last thing we need to know in order to be able to use Riemann-Roch theorem is how to compute the genus of a curve. For smooth projective curves, we have the following

Proposition 2.3. *If C is a smooth projective curve of degree d, then:*

$$g = (d - 1)(d - 2)/2$$

Chapter 3

Gauss and Jacobi sums

3.1 Multiplicative characters

Definition 3.1. Let G be a finite group. A *character* on G is a morphism of groups

$$\chi : G \to \mathbb{C}^*$$

The set of characters on a group G with the composition law $\chi\psi(g) = \chi(g)\psi(g)$ is a group, called the *dual group* of G and denoted by $\hat{G}$. Its identity element will be denoted by ϵ and will be called the *trivial character*.
Let $\mathbb{F}_p$ denote the field with p elements. A multiplicative character on $\mathbb{F}_p$ is a character defined on the multiplicative group $\mathbb{F}_p^*$ of $\mathbb{F}_p$.
We extend the domain of definition (as well as the codomain) of characters χ on $\mathbb{F}_p$ as follows: $\chi(0) = 0$ if $\chi \neq \epsilon$; $\epsilon(0) = 1$.

Remark 3.1. Properties: Let $\chi \in \hat{G}, g \in G$; let e be the identity of G. The following facts are easy to verify:

1. $\chi(e) = 1$
2. $\chi(g)^{|G|} = 1$
3. $\chi(g^{-1}) = \chi(g)^{-1} = \overline{\chi(g)}$

Recall that $\mathbb{F}_p^*$ is a cyclic group. Let a be a generator of $\mathbb{F}_p^*$. Then $\chi \in \hat{\mathbb{F}}_p^*$ is determined by its value $\chi(a) = \zeta_{p-1}^k$, where ζ_{p-1} is a fixed primitive $(p-1)$-th root of unity. For each $k \in \{0, \dots, p-1\}$, the formula

$$\chi(a^n) = \zeta_{p-1}^{kn}$$

defines a character of $\mathbb{F}_p$, and all characters have this form. it follows that $\hat{\mathbb{F}}_p^*$ is cyclic, and is isomorphic to $\mathbb{F}_p^*$.
Generators of $\hat{\mathbb{F}}_p^*$ correspond to the values of k which are coprime with $p-1$.

Proposition 3.1. *Let $\chi \in \hat{\mathbb{F}}_p^*, a \in \mathbb{F}_p^*$*

1. *if $\chi \neq \epsilon$, there exists $b \in \mathbb{F}_p^*$ such that $\chi(b) \neq 1$;*

2. *if $a \neq 1$, there exists $\psi \in \hat{\mathbb{F}}_p^*$ such that $\psi(a) \neq 1$;*

3. *if $\chi \neq \epsilon$, $\sum_{t \in \mathbb{F}_p} \chi(t) = 0$;*
   $\sum_{t \in \mathbb{F}_p} \epsilon(t) = p$;

4. *if $a \neq 1$, $\sum_{\chi \in \hat{\mathbb{F}}_p^*} \chi(a) = 0$.*

Proof. 1. This is clear.

2. Let b be a generator of $\mathbb{F}_p^*$, $a = b^k$. By hypothesis, $p-1$ doesn't divide k. Define $\psi(b^m) = \zeta_{p-1}^m$. Then ψ has the required properties.

3. Suppose $\chi \neq \epsilon$ (otherwise the claim is obvious). From (1), we can take $b : \chi(b) \neq 1$. Then we obtain:

$$\chi(b) \sum_{t \in \mathbb{F}_p} \chi(t) = \sum_{t \in \mathbb{F}_p} \chi(tb) = \sum_{t \in \mathbb{F}_p} \chi(t) \Rightarrow (\chi(b) - 1) \sum_{t \in \mathbb{F}_p} \chi(t) = 0$$

   from which the claim follows.

4. take $\psi \in \hat{\mathbb{F}}_p^*$ such that $\psi(a) \neq 1$ (point (2)), and use the same trick as in the previous point.

□

We're going to use characters in order to determine the number of solutions of equations in $\mathbb{F}_p$. In particular, characters turn out to be very useful in the study of equations of the form:

$$x^n = a$$

where $a \in \mathbb{F}_p^*$, $n \mid p-1$.
We make the following preliminary remark: fix a generator b of $\mathbb{F}_p^*$. Then $a = b^k$ is an $n-th$ power (equivalently, $x^n = a$ has a solution) if and only if $n \mid k$.

Proposition 3.2. *Let $a \in \mathbb{F}_p^*$, $n \mid p-1$;*

1. *if $x^n = a$ has a solution in $\mathbb{F}_p$ and χ is a character such that $\chi^n = \epsilon$, then $\chi(a) = 1$*

2. *if $x^n = a$ has no solutions in $\mathbb{F}_p$, then there is a character χ such that $\chi^n = \epsilon$ and $\chi(a) \neq 1$*

Proof. 1. fix a generator b of $\mathbb{F}_p^*$. From the above observation we have: $a = b^{kn} \Rightarrow \chi(a) = \chi(b^{kn}) = (\chi(b^k))^n = 1$

2. by he above remark $a = b^l$ and $n \nmid l$. Write $l = kn + m$, with $0 < m < n$. Define $\chi(b^t) = \zeta_{p-1}^{(p-1)t/n}$. Then χ is a character of order n, and $\chi(a) = \chi(b^{kn})\chi(b^m) = \zeta_{p-1}^{(p-1)m/n} \neq 1$.

□

Let $n \mid p-1$ and $a \in \mathbb{F}_p^*$. Notice that if the equation $x^n = a$ has a solution α, then it has n solutions. In fact, $\beta^n = a \Leftrightarrow (\alpha\beta^{-1})^n = \delta^n = 1$. Hence each solution of $x^n = a$ is of the form $\alpha\delta$, where δ is an element of order dividing n in $\mathbb{F}_p^*$. But $\mathbb{F}_p^*$ is cyclic, and it is known that if G is a cyclic group of finite order and $d \mid |G|$, G has d elements of order dividing d.
Since we know that $\hat{\mathbb{F}}_p^*$ is cyclic of order $p-1$, the same reasoning proves that there are n characters in $\hat{\mathbb{F}}_p^*$ of order dividing n. Collecting this information we obtain the following fundamental:

Proposition 3.3. *Let $a \in \mathbb{F}_p, n \mid p-1$. Then:*

$$N(x^n = a) = \sum_{\chi^n = \epsilon} \chi(a)$$

Proof. For $a = 0$, $x^n = 0$ has one solution ($x = 0$) and $\sum_{\chi^n=\epsilon} \chi(0) = \epsilon(0) = 1$ (this is why we defined $\chi(0) = 0$ if $\chi \neq \epsilon$ and $\epsilon(0) = 1$).
Suppose now that $a \neq 0$. Then, as we observed above, $x^n = a$ has either 0 or n solutions. In the latter case the previous proposition tells that, for every character χ such that $\chi^n = \epsilon$, $\chi(a) = 1$. It follows that $\sum_{\chi^n=\epsilon} \chi(a) = n$.
If a is not a $n-th$ power, from the same proposition we know that there is a character ψ such that $\psi^n = \epsilon$ and $\psi(a) \neq 1$. With a familiar trick, we obtain:

$$\psi(a) \sum_{\chi^n=\epsilon} \chi(a) = \sum_{\chi^n=\epsilon} \chi\psi(a) = \sum_{\chi^n=\epsilon} \chi(a) \Rightarrow (\psi(a) - 1) \sum_{\chi^n=\epsilon} \chi(a) = 0$$

(as characters of order $\leq n$ form a group). It follows that $\sum_{\chi^n=\epsilon} \chi(a) = 0$. □

3.2 Gauss sums

Definition 3.2. Let χ be a character on $\mathbb{F}_p$, $a \in \mathbb{F}_p$, $\zeta = e^{2\pi i/p}$.

$$g_a(\chi) = \sum_{t \in \mathbb{F}_p} \chi(t)\zeta^{at}$$

is called a *Gauss sum* on $\mathbb{F}_p$ belonging to the character χ.

Notice that, if $\zeta = e^{2\pi i/p}$ and $t \in \mathbb{F}_p$, ζ^t is well defined.
The following lemma follows easily from the definition:

Lemma 3.1. *If $a \neq 0$ and $\chi \neq \epsilon$ then $g_a(\chi) = \chi(a^{-1})g_1(\chi)$; if $a = 0$ and $\chi \neq \epsilon$ we have $g_0(\chi) = 0$. If $a \neq 0$ and $\chi = \epsilon$, $g_a(\epsilon) = 0$. If $a = 0$ and $\chi = \epsilon$, $g_0(\epsilon) = p$.*

From now on we shall denote $g_1(\chi)$ by $g(\chi)$.

Proposition 3.4. *If $\chi \neq \epsilon$, then $|g(\chi)| = \sqrt{p}$*

Proof. We're going to evaluate $A = \sum_a g_a(\chi)\overline{g_a(\chi)}$ in two different ways:
if $a \neq 0$, then by the previous proposition $g_a(\chi) = \chi(a^{-1})g(\chi)$, and $\overline{g_a(\chi)} = \chi(a)\overline{g(\chi)}$. Thus $g_a(\chi)\overline{g_a(\chi)} = |g(\chi)|^2$. Since $g_0(\chi) = 0$ we obtain $A = (p-1)|g(\chi)^2|$.
On the other hand, we also have:

$$g_a(\chi)\overline{g_a(\chi)} = \sum_{x,y} \chi(x)\overline{\chi(y)}\zeta^{ax-ay}$$

Observe that, for $t \in \mathbb{F}_p$ fixed, $\sum_{a\in\mathbb{F}_p} \zeta^{at} = p$ if $t \equiv 0 \pmod p$, otherwise this sum equals 0.
Hence we have:

$$A = \sum_{x,y} \chi(x)\overline{\chi(y)}\delta(x,y)p = (p-1)p$$

Comparing the two values of A, we get the desired result. $\square$

Remark 3.2. Let $\overline{\chi}$ denote the character whose value at a is $\overline{\chi(a)}$ (which coincides with χ^{-1}). Then we have:

$$\overline{g(\chi)} = \sum_t \overline{\chi(t)}\zeta^{-t} = \chi(-1)\sum_t \overline{\chi(-t)}\zeta^{-t} = \chi(-1)g(\bar{\chi})$$

Thus, the equality $|g(\chi)|^2 = p$ can be written as: $g(\chi)g(\bar{\chi}) = \chi(-1)p$

Remark 3.3. Let $\chi(a) = \left(\frac{a}{p}\right) = a^{(p-1)/2}$ be the quadratic residue character on $\mathbb{F}_p$. Evaluating the sum $\sum_a g_a g_{-a}$ in two ways, in the same way as in the previous proof, we find that $g^2 = (-1)^{(p-1)/2}p$. This is the starting point of an elegant proof of the law of quadratic reciprocity, whose conciseness and beauty make it worth describing.
Let p, q two different odd primes. Let ζ be a primitive $q-$th root of unity in $\bar{\mathbb{F}}_p$. Let χ be the quadratic residue character on $\mathbb{F}_q$ and $g = \sum_{t\in\mathbb{F}_q} \chi(t)\zeta^t \in \bar{\mathbb{F}}_p$ be the corresponding Gauss sum. Then, as $g^2 = (-1)^{\frac{q-1}{2}}q$, $\left(\frac{(-1)^{\frac{q-1}{2}}q}{p}\right) = 1$ if and only if $g \in \mathbb{F}_p \Leftrightarrow g^p = g$. Now

$$g^p = \sum_{t\in\mathbb{F}_q} \chi(t)^p\zeta^{pt} = \sum_{t\in\mathbb{F}_q} \chi(t)\zeta^{pt} = \chi(p)\sum_{t\in\mathbb{F}_q} \chi(pt)\zeta^{pt} = \chi(p)g$$

So $g^p = g \Leftrightarrow \chi(p) = \left(\frac{p}{q}\right) = 1$.
Hence we obtain:

$$\left(\frac{(-1)^{\frac{q-1}{2}}q}{p}\right) = 1 \Leftrightarrow \left(\frac{p}{q}\right) = 1$$

hence

$$(-1)^{\frac{p-1}{2}\frac{q-1}{2}}\left(\frac{q}{p}\right) = 1 \Leftrightarrow \left(\frac{p}{q}\right) = 1$$

which is the law of quadratic reciprocity.

Remark 3.4. As $g \in \mathbb{Q}[\zeta]$, we obtain that $\sqrt{\pm p} \in \mathbb{Q}[e^{2\pi i/p}]$. With a little extra work, it can be deduced from this fact that each quadratic extension of $\mathbb{Q}$ is contained in a cyclotomic extension. This is a very special case of the Kronecker-Weber theorem, which asserts that the same is true for every *abelian* extension of $\mathbb{Q}$ (i.e. a finite Galois extension of $\mathbb{Q}$ whose Galois group is abelian).[1]

3.3 Jacobi sums

Definition 3.3. Let χ and λ be characters on $\mathbb{F}_p$. The expression:

$$J(\chi, \lambda) = \sum_{a+b=1} \chi(a)\lambda(b)$$

is called a *Jacobi sum*.

The proof of the following elementary properties of Jacobi sums is left to the reader (see [8] if you need help):

Lemma 3.2. *Let χ, λ be nontrivial characters. Then*

1. $J(\chi, \chi) = p$
2. $J(\epsilon, \chi) = 0$
3. $J(\chi, \chi^{-1}) = -\chi(-1)$
4. *If $\chi\lambda \neq \epsilon$, then*
$$J(\chi, \lambda) = \frac{g(\chi)g(\lambda)}{g(\chi\lambda)}$$

Corollary 3.1. *If $\chi, \lambda, \chi\lambda$ are nontrivial, then $|J(\chi, \lambda)| = \sqrt{p}$. In particular, if χ, χ^2 are nontrivial, $|J(\chi, \chi)| = \sqrt{p}$*

This seemingly innocuous statement actually has many useful consequences, as we will soon discover. It allows us to obtain good estimates for the number of solutions of certain types of equations over finite fields.
The following proposition is another non trivial consequence of the previous corollary:

Proposition 3.5. *Let p be a prime number.*

1. $p \equiv 1 \pmod 4 \Leftrightarrow \exists\, a, b \in \mathbb{Z} : a^2 + b^2 = p$
2. $p \equiv 1 \pmod 3 \Leftrightarrow \exists\, a, b \in \mathbb{Z} : a^2 - ab + b^2 = p$

Proof. If $p \equiv 1 \pmod 4$ then $4 \mid (p-1) = |\mathbb{F}_p^*| \Rightarrow \hat{\mathbb{F}}_p^*$ has an element of order 4. Let us call it λ. Then λ takes its values in the set $\{1, -1, i, -i\}$, so $J(\lambda) \in \mathbb{Z}[i]$, and $|J(\lambda, \lambda)|^2 = |(a + ib)|^2 = a^2 + b^2 = p$, by Corollary 3.1.
If $p \equiv 1 \pmod 3$, take a character λ on $\mathbb{F}_p$ of order 3. Then $J(\lambda) \in \mathbb{Z}[\omega]$, and $p = |J(\lambda, \lambda)|^2 = |(a + \omega b)|^2 = a^2 - ab + b^2$.
The opposite implications are easy (look at the congruence class modulo 4 and 3). □

[1] Actually, the first correct proof of this theorem is due to Hilbert. To prove it one has to explore much deeper waters: it is one of the main results of the so called *class field theory*; see [16].

Proposition 3.6. *Let $p \equiv 1 \pmod n$ be a prime, χ a character of order n in $\mathbb{F}_p$. Then*

$$g(\chi)^n = \chi(-1)pJ(\chi,\chi)J(\chi,\chi^2)\dots J(\chi,\chi^{n-2})$$

In particular, if $n = 3$, $g(\chi)^3 = pJ(\chi,\chi)$.

Proof. Using Lemma 3.2 we obtain: $g(\chi)^2 = J(\chi,\chi)g(\chi^2)$. Multiplication by $g(\chi)$ and Lemma 3.2 again give $g(\chi)^3 = J(\chi,\chi)J(\chi,\chi^2)g(\chi^3)$. Continuing in this way, we get:

$$g(\chi)^{n-1} = J(\chi,\chi)J(\chi,\chi^2)\dots J(\chi,\chi^{n-2})g(\chi^{n-1})$$

Now, $\chi^{n-1} = \overline{\chi} \Rightarrow g(\chi^{n-1})g(\chi) = \chi(-1)p$ (Remark 3.2). Multiplication of the above equality by $g(\chi)$ gives the result. □

Proposition 3.7. *Let $p \equiv 1 \pmod 3$ be a prime, χ a character of order 3 on $\mathbb{F}_p$. Set $J(\chi,\chi) = a + b\omega$. Then $a \equiv -1 \pmod 3, b \equiv 0 \pmod 3$.*

Proof. We have the following congruences in the ring of algebraic integers $\mathbb{Z}[\omega]$:

$$g(\chi)^3 = (\sum_t \chi(t)\zeta^t)^3 \equiv \sum_t \chi(t)^3\zeta^{3t} \equiv \sum_{t\neq 0} \zeta^{3t} \equiv -1 \pmod 3$$

Thus:

$$g(\chi)^3 = pJ(\chi,\chi) \equiv a + b\omega \equiv -1 \pmod 3$$

Working with $\overline{\chi}$ instead of χ, and observing that $g(\overline{\chi}) = \overline{g(\chi)}$ (as χ is a cubic character, hence $\chi(-1)^3 = 1 \Rightarrow \chi(-1) = 1$) we get:

$$g(\overline{\chi})^3 = pJ(\overline{\chi},\overline{\chi}) \equiv a + b\overline{\omega} \equiv -1 \pmod 3$$

Subtraction yields $b(\omega - \overline{\omega}) = b\sqrt{-3} \equiv 0 \pmod 3 \Rightarrow 9 \mid -3b^2 \Rightarrow 3 \mid b$. Now, since $a + b\omega \equiv -1 \pmod 3$, we obtain $a \equiv -1 \pmod 3$. □

We can now prove a beautiful result due to Gauss; this is a first example of how to use Jacobi sums and their properties in order to obtain information about the number of solutions of a polynomial equation in $\mathbb{F}_p$.

Theorem 3.1. *(Gauss) Let $N(x^3+y^3 = 1)$ denote the number of solutions of the equation $x^3 + y^3 = 1$ in $\mathbb{F}_p$.*

1. *If $p \equiv 2 \pmod 3$, then $N(x^3 + y^3 = 1) = p$.*

2. *If $p \equiv 1 \pmod 3$, then there are integers A, B such that $4p = A^2 + 27B^2$; A is uniquely determined by the condition $A \equiv 1 \pmod 3$, and $N(x^3+y^3 = 1) = p-2+A$*

Proof. 1. As $p \equiv 2 \pmod 3$, $x \mapsto x^3$ is an automorphism of $\mathbb{F}_p^*$. Then each element in $\mathbb{F}_p$ is the cube of a unique element in $\mathbb{F}_p$, hence $N(x^3+y^3 = 1) = N(x+y = 1) = p$. Of course, we did not need Gauss to prove this.

2. This is going to be more interesting.
   Take a character χ on $\mathbb{F}_p$ of order 3. Then ϵ, χ, χ^2 are all the characters on $\mathbb{F}_p$ of order dividing 3 and, by Proposition 3.3, we have:

$$N(x^3+y^3=1) = \sum_{a+b=1} N(x^3=a)N(y^3=b) = \sum_{a+b=1}\left(\sum_{i=0}^{2}\chi^i(a)\sum_{j=0}^{2}\chi^j(b)\right) =$$
$$= \sum_{i,j=0}^{2} J(\chi^i,\chi^j) = J(\epsilon,\epsilon)+J(\epsilon,\chi)+J(\chi,\epsilon)+J(\chi,\chi^2)$$
$$+ J(\chi^2,\chi)+J(\chi,\chi)+J(\chi^2,\chi^2)$$

   Let $J(\chi,\chi)=a+b\omega$. By Lemma 3.2, we obtain:

$$\begin{aligned} N(x^3+y^3=1) =& p+0+0-\chi(-1)-\chi^2(-1)+J(\chi,\chi)+J(\bar\chi,\bar\chi) = \\ &= p-2+2Re(J(\chi,\chi)) = p-2+2Re(a+b\omega) = \\ &= p-2+(2a-b) = p-2+A \end{aligned}$$

   where $A=2a-b\equiv 1 \pmod 3$, $b\equiv 0 \pmod 3$ and $a^2-ab+b^2=p$ (Proposition 3.5 and Proposition 3.7), hence $4p=(2a-b)^2+3b^2=A^2+27B^2$, where $B=b/3$. It remains only to show that such an A is unique.
   Suppose that $4p=A'^2+27B'^2$, with $A'\equiv 1 \pmod 3$. Then, the equations:

$$\begin{aligned} 3B' =& b' \\ (2a'-b') =& A' \end{aligned}$$

   determine uniquely two integers a', b' such that $4p=(2a'-b')^2+3b'^2 \Rightarrow p = a'^2-a'b'+b'^2$. Moreover, $3|b'$ and $2a'-b'\equiv 1 \pmod 3$.
   In $\mathbb{Z}[\omega]$ we have $N(a+\omega b)=N(a'+\omega b')=p$, so $a+\omega b$ is associated with $a'+\omega b'$. This means that $a'+\omega b'=u(a+\omega b)$ where $u=\pm1,\pm\omega$ or $\pm\omega^2$. Examination of each case shows that only for $u=1$ we have $3|b'$ and $2a'-b'\equiv 1 \pmod 3$. Therefore $A'=A,\ B'=B$.

□

Example 3.1. $p=97$. Then $4p=388=19^2+27$. So $A=19, B=1$. Hence, the curve of affine equation $x^3+y^3=1$ has $97-2+19=114$ points in $\mathbb{A}^2(\mathbb{F}_{97})$. It would have been much harder to obtain the same conclusion by brute force.
Notice that this curve has 3 points at infinity, corresponding to the projective solutions of the equation

$$x^3+y^3=0$$

in $\mathbb{P}^1(\mathbb{F}_{97})$. Let $a=y/x$ (observe that in the above equation we must have $x\neq 0, y\neq 0$). Then we obtain the equivalent equation $a^3=-1$ in $\mathbb{F}_{97}^*$, which has exactly 3 solutions, since $3 \mid (97-1)$. Hence our curve has 117 points in $\mathbb{P}^2(\mathbb{F}_{97})$.

Remark 3.5. If we consider the projective closure of the affine curve $x^3+y^3=1$, namely $x^3+y^3-z^3=0$, it is easy to show, using the previous result, that the number of points on this curve defined on $\mathbb{F}_p$ is $p+1$ if $p\equiv 2 \pmod 3$, $p+1-A$ if $p\equiv 1 \pmod 3$ (it suffices to add to affine points the points at infinity of the curve, solutions of the equation $x^3 + y^3 = 0$; these can be determined as in the previous example). In both cases, we observe that the number of projective points on the curve $x^3 + y^3 - z^3 = 0$ satisfies the inequality: $|N(x^3+y^3-z^3=0)-(p+1)| \leq 2\sqrt{p}$. This is a special case of Hasse theorem: we will prove it for other smooth projective curves of degree 3 in Chapter 5.

We conclude this section with a technical result that will be used in the sequel:

Lemma 3.3. *Let p be an odd prime, ρ a character of order 2 and χ any non trivial character of $\mathbb{F}_p$. Then $J(\rho,\chi) = \chi(4)J(\chi,\chi)$.*

Proof.

$$\begin{aligned} J(\rho,\chi) &= \sum_{u+v=1} \rho(u)\chi(v) = \sum_{u+v=1} (1+\rho(u))\chi(v) = \\ &= \sum_{u+v=1} N(t^2=u)\chi(v) = \sum_t \chi(1-t^2) = \\ &= \chi(4)\sum_t \chi\Big(\frac{1-t}{2}\Big)\chi\Big(\frac{1+t}{2}\Big) = \chi(4)J(\chi,\chi) \end{aligned}$$

□

3.4 Cubic residue character

Let $\pi \in \mathbb{Z}[\omega]$, be a prime, with $N(\pi) \neq 3$. Equivalently, π is not associate with $1-\omega$. Then it's immediate to see that the residue classes of $1, \omega, \omega^2$ are distinct in $\mathbb{Z}[\omega]/\pi\mathbb{Z}[\omega]$. Thus $(\mathbb{Z}[\omega]/\pi\mathbb{Z}[\omega])^*$ contains a subgroup of order 3, and so $3 \mid |(\mathbb{Z}[\omega]/\pi\mathbb{Z}[\omega])^*| = N(\pi)-1$. Now, take $\alpha \in \mathbb{Z}[\omega]$. If π does not divide α, then $\alpha^{(N(\pi)-1)/3} \equiv 1, \omega$ or $\omega^2 \pmod{\pi}$. In fact, let $\alpha^{(N(\pi)-1)/3} \equiv A \pmod{\pi}$. Then A is a zero of the polynomial $(x-1)(x-\omega)(x-\omega^2) \in (\mathbb{Z}[\omega]/\pi\mathbb{Z}[\omega])[x]$, as $A^3 - 1 \equiv \alpha^{N(\pi)-1} \equiv \alpha^{|(\mathbb{Z}[\omega]/\pi\mathbb{Z}[\omega])^*|} \equiv 1 \pmod{\pi}$.
Therefore we can give the following definition:

Definition 3.4. If $N(\pi) \neq 3$, the cubic residue character of α modulo π is defined by:

1. $(\alpha/\pi)_3 = 0$ if $\pi \mid \alpha$;
2. $(\alpha/\pi)_3 \equiv \alpha^{(N(\pi)-1)/3} \pmod{\pi}$, with $(\alpha/\pi)_3 \in \{1, \omega, \omega^2\}$, if $(\pi, \alpha) = 1$.

In this section, will also denote $(\alpha/\pi)_3$ by $\chi_\pi(\alpha)$. The following properties of $\chi_\pi(\alpha)$ are clear:

Lemma 3.4. *1. $\chi_\pi(\alpha\beta) = \chi_\pi(\alpha)\chi_\pi(\beta)$*

2. if $\alpha \equiv \beta \pmod{\pi}$, then $\chi_\pi(\alpha) = \chi_\pi(\beta)$

As a consequence of the previous Lemma, χ_π gives rise to a character defined on the group $(\mathbb{Z}[\omega]/\pi\mathbb{Z}[\omega])^*$, which we will still denote with the same symbol.
Let $\pi \in \mathbb{Z}[\omega], N(\pi) = p \equiv 1 \pmod{3}$, p prime in $\mathbb{Z}$. Then π is prime in $\mathbb{Z}[\omega]$ and $(\mathbb{Z}[\omega]/\pi\mathbb{Z}[\omega])$ is a field of cardinality $N(\pi) = p$, thus it may be identified with $\mathbb{F}_p$. Thus, χ_π can be seen as a character of order 3 on $\mathbb{F}_p$. We can determine explicitly the value $J(\chi_\pi, \chi_\pi)$ if π is primary:

Proposition 3.8. *If $\pi \in \mathbb{Z}[\omega]$ is primary and $N(\pi) = p$, then:*

$$J(\chi_\pi, \chi_\pi) = \pi$$

Proof. By corollary 3.1 $J(\chi_\pi, \chi_\pi) = \pi'$ is prime in $\mathbb{Z}[\omega]$ (as its norm is p). By Proposition 3.7 we know that π' is primary. Thus, we need only to show that $\pi \mid \pi'$, as this will imply that π and π' are associates and both primary, hence equal. Now, we have by definition:

$$J(\chi_\pi, \chi_\pi) = \sum_x \chi_\pi(x)\chi_\pi(1-x) \equiv \sum_x x^{(p-1)/3}(1-x)^{(p-1)/3} \pmod{\pi}$$

As the polynomial $x^{(p-1)/3}(1-x)^{(p-1)/3}$ is of degree $< (p-1)$, we conclude applying the following proposition to each monomial in $x^{(p-1)/3}(1-x)^{(p-1)/3}$. □

Proposition 3.9. *1. If $(p-1) \mid n$, $\sum_{x\in\mathbb{F}_p} x^n = p-1$;*

2. If $(p-1) \nmid n$, $\sum_{x\in\mathbb{F}_p} x^n = 0$.

Proof. (1) is clear. Let us prove (2): as $(p-1) \nmid n$, there exists $y \in \mathbb{F}_p^*$ such that $y^n \neq 1$. It suffices to take a generator of the multiplicative group $\mathbb{F}_p^*$. Hence we have, as usual:

$$\sum_{x\in\mathbb{F}_p} x^n = \sum_{x\in\mathbb{F}_p} (xy)^n = y^n \sum_{x\in\mathbb{F}_p} x^n \Rightarrow (y^n - 1) \sum_{x\in\mathbb{F}_p} x^n = 0 \Rightarrow \sum_{x\in\mathbb{F}_p} x^n = 0$$

□

As a consequence of this simple fact, we obtain the following important:

Theorem 3.2. *(Chevalley-Warning) Let $f(x,y,z) \in \mathbb{F}_p[x,y,z]$ be a homogeneous polynomial of degree 2. For each prime p there is always at least a non zero solution of the equation $f(x,y,z) = 0$ in $\mathbb{A}^3(\mathbb{F}_p)$.*
Equivalently, there is always at least a point on the projective curve $C(\mathbb{F}_p) = \{[x,y,z] \in \mathbb{P}^2(\mathbb{F}_p) : f(x,y,z) = 0\}$.

Proof. Fix a prime p. Let N be the number of solutions of the equation $f(x,y,z) = 0$ with $x,y,z \in \mathbb{F}_p$. We're going to show that $N \equiv 0 \pmod{p}$. This will imply the result. The key observation is that $f(x,y,z) \neq 0$ if and only if $f(x,y,z)^{p-1} \equiv 1 \pmod{p}$. Hence

$$N \equiv \sum_{x,y,z\in\mathbb{F}_p} 1 - f(x,y,z)^{p-1} \pmod{p}$$

Now notice that the each monomial $m(x,y,z) = x^i y^j z^k$ in the polynomial $1 - f(x,y,z)^{p-1}$ has degree at most $2(p-1)$. Hence one of x,y,z, say x, appears in m with an exponent which is less than $p-1$. By the above Proposition: $\sum_{x,y,z\in\mathbb{F}_p} m(x,y,z) = \left(\sum_{y,z} y^j z^k\right) \sum_x x^i \equiv 0 \pmod{p}$. The result follows. □

Remark 3.6. A careful examination of the above proof shows that the Theorem is true for any homogeneous polynomial whose degree is strictly inferior to the number of variables. Moreover, a very similar proof works for an arbitrary finite field $\mathbb{F}_{p^m}$.

Remark 3.7. Cubic residue characters allow us to state in a very simple way the law of cubic reciprocity:

Theorem 3.3. *If $\lambda, \pi \in \mathbb{Z}[\omega]$ are primary, then $(\lambda/\pi)_3 = (\pi/\lambda)_3$.*

See [8] for a proof of this result.

3.5 Biquadratic residue character

In this section we're going to define a biquadratic (or quartic) residue character, in the same way as we defined the cubic residue character in the previous section. We will work in the ring $\mathbb{Z}[i]$ instead of the ring $\mathbb{Z}[\omega]$.
Let π be a prime in $\mathbb{Z}[i]$, with $N(\pi) \neq 2 \Leftrightarrow (\pi) \neq (1+i)$. Then the residues classes of $1, -1, i, -i$ are distinct in $\mathbb{Z}[i]/\pi\mathbb{Z}[i]$, so $(\mathbb{Z}[i]/\pi\mathbb{Z}[i])^*$ contains a subgroup of order 4, hence $4 \mid |(\mathbb{Z}[i]/\pi\mathbb{Z}[i])^*| = N(\pi) - 1$.
Let $\alpha \in \mathbb{Z}[i]$, and suppose that π doesn't divide α. Then in $(\mathbb{Z}[i]/\pi\mathbb{Z}[i])^*$ we have $\alpha^{N(\pi)-1} = \alpha^{|(\mathbb{Z}[i]/\pi\mathbb{Z}[i])^*|} \equiv 1 \pmod{\pi} \Rightarrow \alpha^{(N(\pi)-1)/4} \equiv 1, -1, i$ or $-i \pmod{\pi}$.

Definition 3.5. If $N(\pi) \neq 2$, the *biquadratic residue character* of α modulo π is defined by:

1. $(\alpha/\pi)_4 = 0$ if $\pi \mid \alpha$;
2. $(\alpha/\pi)_4 \equiv \alpha^{(N(\pi)-1)/4} \pmod{\pi}$, with $(\alpha/\pi)_4 \in \{1, -1, i, -i\}$, if $(\pi, \alpha) = 1$

In this section we will also denote $(\alpha/\pi)_4$ by $\chi_\pi(\alpha)$.

Lemma 3.5. *1.* $\chi_\pi(\alpha\beta) = \chi_\pi(\alpha)\chi_\pi(\beta)$

2. if $\alpha \equiv \beta \pmod{\pi}$*, then* $\chi_\pi(\alpha) = \chi_\pi(\beta)$

It follows from these properties that χ_π induces a character on the group $(\mathbb{Z}[i]/\pi\mathbb{Z}[i])^*$, which we will still denote with the same symbol.
Suppose that π is prime in $\mathbb{Z}[i]$ and $N(\pi) = p \equiv 1 \pmod{4}$. Then $|\mathbb{Z}[i]/\pi\mathbb{Z}[i]| = N(\pi) = p$, so $\mathbb{Z}[i]/i\mathbb{Z}[i]$ is isomorphic to $\mathbb{F}_p$ and χ_π can be seen as a character of order 4 on $\mathbb{F}_p$. Suppose that π is primary. We would like to determine, as in the previous section, the value of $J(\chi_\pi, \chi_\pi)$. We need the following Proposition:

Lemma 3.6. $-\chi_\pi(-1)J(\chi_\pi, \chi_\pi)$ *is primary.*

Proof.
$$J(\chi_\pi, \chi_\pi) = 2\sum_{t=2}^{(p-1)/2} \chi_\pi(t)\chi_\pi(1-t) + \chi_\pi\left(\frac{p+1}{2}\right)^2$$
Now, any unit in $\mathbb{Z}[i]$ is congruent to $1 \pmod{1+i}$. Moreover, $(2+2i) \mid 4 \mid (p-1) \Rightarrow p \equiv 1 \pmod{2+2i}$. Finally, $\chi_\pi(\frac{p+1}{2})^2 = \chi_\pi(2^{-1})^2 = \chi_\pi(2)^{-2} = \chi_\pi(2)^2 = \chi_\pi(-i(1+i)^2)^2 = \chi_\pi(-i)^2 = \chi_\pi(-1)$. Thus:
$$J(\chi_\pi, \chi_\pi) \equiv 2\left(\frac{p-3}{2}\right) + \chi_\pi(-1) \equiv -2 + \chi_\pi(-1) \pmod{2+2i}$$
Thus:
$$-\chi_\pi(-1)J(\chi_\pi, \chi_\pi) \equiv 2\chi_\pi(-1) - 1 \equiv 1 \pmod{2+2i}$$
since $\chi_\pi(-1) = \pm 1$. □

Proposition 3.10. *If* $\pi \in \mathbb{Z}[i]$ *is primary, then:*
$$-\chi_\pi(-1)J(\chi_\pi, \chi_\pi) = \pi$$

Proof. We know that $N(J(\chi,\chi)) = p$ (Corollary 3.1), hence $J(\chi,\chi)$ is prime in $\mathbb{Z}[i]$, and $-\chi_\pi(-1)J(\chi_\pi,\chi_\pi) = \pi'$ is also prime. Moreover, we know by the above Lemma that π' is primary. Thus, it's enough to show that $\pi \mid \pi' \Leftrightarrow \pi \mid J(\chi_\pi,\chi_\pi)$. By definition

$$J(\chi_\pi,\chi_\pi) \equiv \sum_{t=1}^{p-1} t^{(p-1)/4}(1-t)^{(p-1)/4} \pmod{\pi}$$

Now, the polynomial $\sum_{t=1}^{p-1} t^{(p-1)/4}(1-t)^{(p-1)/4}$ has degree $< p-1$, hence we conclude by Proposition 3.9. □

Remark 3.8. 1. We can also define the quartic residue character $(a/p)_4$ with respect to a prime $p \equiv 3 \pmod{4}$, which is also prime in $\mathbb{Z}[i]$. This is a character on $\mathbb{Z}[i]/p\mathbb{Z}[i] \cong \mathbb{F}_{p^2}$. For $a \in \mathbb{Z}$ such that $(a,p) = 1$ we obtain: $(a/p)_4 \equiv a^{(N(p)-1)/4} \equiv a^{(p^2-1)/4} \equiv a^{(p-1)(p+1)/4} \equiv 1 \pmod{p}$ as $a^{p-1} \equiv 1 \pmod{p}$.

2. Even more generally, we can define, for $a \in \mathbb{Z}[i]$ such that $(1+i) \nmid a$, the biquadratic residue symbol modulo a: $(\alpha/a)_4 = \prod_i (\alpha/\lambda_i)_4$ where $\alpha = \prod_i \lambda_i$ and each λ_i is prime in $\mathbb{Z}[i]$.
The law of biquadratic reciprocity describes the relation between $(\alpha/a)_4$ and $(a/\alpha)_4$. We will need the following special case of this result:

Theorem 3.4. *If* $a \equiv 1 \pmod{4}$ *and* α *is primary , with* $(\alpha, a) = 1$*, then*

$$(\alpha/a)_4 = (a/\alpha)_4$$

We will also need the following "supplement":

Lemma 3.7. *If* $\alpha = a + bi$ *is primary, then*

$$\left(\frac{1+i}{\alpha}\right)_4 = i^{(a-b-b^2-1)/4}$$

For a proof, see [8].

Chapter 4

The Zeta function

In this chapter we're going to define and study some of the fundamental properties of the zeta function of a curve. This function, defined in terms of a power series, provides information about the number of points on a curve over the fields $\mathbb{F}_{p^m}$, where p is a fixed prime and m varies among all natural numbers.
First of all, let us fix some notation. Let $K = \mathbb{F}_p$, $f(x, y) \in \mathbb{F}_p[x, y]$. Let $C(K) = \{(x, y) \in \mathbb{A}^2(K) : f(x, y) = 0\}$ be an irreducible smooth curve. Let $N_m = |\{(x, y) \in C(\bar{K}) : (x, y)$ is defined over $\mathbb{F}_{p^m}\}|$.

Definition 4.1. The *zeta function* of the affine curve $C(K)$ is the (formal) series given by:

$$Z_C(t) = exp\left(\sum_{m=1}^{\infty} \frac{N_m t^m}{m}\right)$$

We will denote the zeta function simply by $Z(t)$ when the corresponding curve is clear from the context.
The reader may wonder why we defined the zeta function in such a (seemingly) strange way. Hopefully, the reason should become clear in a few pages.
Let us start by giving a simple example:

Example 4.1. Let $f(x, y) = x$, so that $C(\mathbb{F}_p) = \mathbb{A}^1(\mathbb{F}_p)$. Then, $N_m = p^m$. An easy computation gives:

$$Z_C(t) = exp\left(\sum_{m=1}^{\infty} \frac{p^m t^m}{m}\right) = exp(-log(1 - pt)) = (1 - pt)^{-1}$$

At least in this case, the zeta function turns out to be much simpler than one could expect from the definition.

Another reasonable question is the following: is there any relation between the zeta function we just defined and the well known Riemann zeta function? Before facing this problem, let us recall briefly some of the basic properties of the Riemann zeta function (for a proof, see [20]):

Proposition 4.1. *The series*

$$\sum_{n=1}^{\infty} \frac{1}{n^s}$$

converges for all $s \in \mathbb{C}$ such that Re $s > 1$, and defines an holomorphic function $\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} = \prod_{p\ prime} (1 - p^{-s})^{-1}$ on this half plane, called the Riemann zeta function, with the following properties:

1. *(analytic continuation) $\zeta(s)$ can be analytically continued to a meromorphic function (still denoted by $\zeta(s)$) on the whole plane, with a simple pole at $s = 1$ and simple zeros at the points $-2, -4, \ldots, -2n, \ldots$, called the* trivial zeros *of $\zeta(s)$;*

2. *(functional equation) Let $\xi(s) = \pi^{-s/2}\Gamma(s/2)\zeta(s)$, where $\Gamma(s)$ denotes Euler's gamma function. Then $\xi(s)$ satisfies the following functional equation:*

$$\xi(s) = \xi(1 - s)$$

3. *(Riemann hypothesis) The non trivial zeros of $\zeta(s)$ all lie in the strip $0 <$ Re $s < 1$. Riemann hypothesis conjectures that they all lie on the line Re $s = 1/2$.*

If K is any number field, we can give a definition of the zeta function associated to the Dedekind domain $\mathcal{O}_K$ which generalizes the Riemann zeta function (corresponding to the case $K = \mathbb{Q}$):

$$Z_K(s) = \sum_I \frac{1}{N(I)^s}$$

where the sum is over all nonzero ideals in $\mathcal{O}_K$. It follows from unique factorization of ideals in $\mathcal{O}_K$ and the fact that the norm is multiplicative that the zeta function can be written in the following form (called an Euler product):

$$Z_K(s) = \prod_{M \in Max(\mathcal{O}_K)} (1 - N(M)^{-s})^{-1}$$

Observe now that the last two formulas make sense for arbitrary Dedekind domains D with finite quotients. Coordinate rings $K[C]$ of affine curves defined over $K = \mathbb{F}_p$ always have this property: in fact, Corollary 2.1 states that $K[C]$ is a Dedekind domain, and Theorem 2.1 tells that if $M \in Max(K[C])$, then $M = \psi_P$, where $P = (a, b) \in \bar{K}$. Hence $(K[C]/M) \cong K(P) = \mathbb{F}_p(P)$, and this is a finite field.
We will now prove that the zeta function of the Dedekind domain $\mathbb{F}_p[C]$, as defined above, coincides with the zeta function of the affine curve defined at the beginning of the paragraph (up to a change of variable).
As in Example 2.1, let $b_d = |\{M \in Max(K[C]) : [K[C]/M : \mathbb{F}_p]\}| = d$. Observe that if $[K[C]/M : \mathbb{F}_p]\}| = d$ then $N(M) = p^d$. Hence we obtain:

$$Z_{K[C]}(s) = \prod_{M \in Max(K[C]} (1 - N(M)^{-s})^{-1} = \prod_{d \in \mathbb{N}} \left(1 - \frac{1}{p^{sd}}\right)^{-b_d}$$

For $t = q^{-s}$, we have (denoting $Z_{K[C]}(t)$ by $Z(t)$):

$$
\begin{aligned}
Z(t) = \prod_{d\in\mathbb{N}}(1-t^d)^{-b_d} \Rightarrow log(Z(t)) &= -\sum_{d\in\mathbb{N}} b_d log(1-t^d) = \\
&= \sum_{d\in\mathbb{N}} b_d \left(\sum_{i=1}^{\infty} \frac{t^{di}}{i}\right) = \\
&= \sum_{n=1}^{\infty} \left(\sum_{d|n} db_d\right) \frac{t^n}{n} = \sum_{n=1}^{\infty} N_n \frac{t^n}{n}
\end{aligned}
$$

since by Example 2.1 we have $\sum_{d|n} db_d = N_n$. So we have: $log(Z(t)) = \sum_{n=1}^{\infty} N_n \frac{t^n}{n}$. Exponentiating both sides we get the desired equality; this also explains why the exponential is used in the definition of the zeta function.
Now we are going to define the zeta function of a projective plane curve. The definition is analogous to the affine case Anyway, as the projective world is better than the affine one, we will see that the zeta function of a projective plane curve enjoys extraordinary properties.
Let $K = \mathbb{F}_p$, $f(x, y, z) \in \mathbb{F}_p[x, y, z]$ homogeneous, $C(K) = \{[x, y, z] \in \mathbb{P}^2(K) : f(x, y, z) = 0\}$ be an irreducible smooth curve.
Let $N_m = |\{P \in C(\bar{K}) : P \text{ is defined over } \mathbb{F}_{p^m}\}|$.

Definition 4.2. The *zeta function* of the projective curve $C(K)$ is the (formal) series given by:

$$Z_C(t) = exp\left(\sum_{m=1}^{\infty} \frac{N_m t^m}{m}\right)$$

Example 4.2. Let $f(x, y, z) = x$, so that $C = \mathbb{P}^1(\mathbb{F}_p)$. Then $N_m = p^m + 1$. Hence we obtain:

$$Z_{\mathbb{P}^1(\mathbb{F}_p)}(t) = exp\left(\sum_{m=1}^{\infty} \frac{(p^m+1)t^m}{m}\right) = \frac{1}{(1-pt)(1-t)}$$

This example allows us to determine with no extra work the zeta function of all projective smooth curves of degree 2. Let C be such a curve.
By Chevalley-Warning Theorem and Remark 3.6 C has always at least one point on each $\mathbb{F}_{p^m}$. Once we have a point P, all the other points on the curve defined over $\mathbb{F}_{p^m}$ are obtained intersecting the curve with a line $ax + by + cz = 0$ passing through P, with $a, b, c \in \mathbb{F}_{p^m}$. Clearly different lines through P correspond to different points on C.
Hence $C(\mathbb{F}_{p^m}) \cong \mathbb{P}^1(\mathbb{F}_{p^m})$, so $N_m = p^m + 1$.
Therefore, the zeta function of an arbitrary smooth projective plane curve of degree 2 is:

$$Z(t) = \frac{1}{(1-pt)(1-t)}$$

Our aim now is to study the fundamental properties of the zeta function of a smooth projective curve. In particular, we want to find out whether this function enjoys similar properties to those of the Riemann zeta function.
First of all, it's easy to see that the zeta function, which we defined just as a formal power

series, can actually be thought as a holomorphic function defined on a small enough disk in the plane centred in the origin. The problem arises of determining whether there is an analytic continuation of this function to a meromorphic function defined on the whole plane. The above example shows that this is true for all curves of degree ≤ 2. In fact, in this case we've seen that much more is true, namely, the zeta function is a *rational* function.
Let us see what happens in the general case.

4.0.1 Rationality of the zeta function

Let C be a projective curve defined over $K = \mathbb{F}_p$.
First of all, we need a clever way to count the number of points on C which are defined over $\mathbb{F}_{p^m}$. Recall that $P \in C(\bar{K})$ is defined over $\mathbb{F}_{p^m}$ if and only if $P^\sigma = P \ \forall \sigma \in Gal(\bar{\mathbb{F}}_p/\mathbb{F}_{p^m}) \Leftrightarrow \phi^m(P) = P$, where $\phi(x) = x^p$ is the Frobenius automorphism.
Let $deg(P)$ denote the least value of d such that $\phi^d(P) = P$. This is called the *degree* of the point P. If P has degree d it is defined over $\mathbb{F}_{p^d}$, and the points $P, \phi(P), \ldots \phi^{d-1}(P)$ are distinct points on C defined over $\mathbb{F}_{p^d}$.

Definition 4.3. Let $P \in C(\bar{K})$ be a point of degree d. Then the divisor:

$$\mathcal{P} = P + \phi(P) + \phi^2(P) + \ldots + \phi^{d-1}P$$

is called a *prime divisor*.

Remark 4.1.
1. Observe that $\mathcal{P}^\phi = \mathcal{P}$, so $\mathcal{P}$ is a divisor of degree d defined over $\mathbb{F}_p$.
2. If $D \geq 0$ is a divisor defined over $\mathbb{F}_p$ in which a point P of degree d appears with a non zero coefficient, then the points $\phi^i(P),\ i = 1, \ldots, d-1$ must have the same coefficient. This allows to show that each divisor $D \geq 0$ defined over $\mathbb{F}_p$ can be written uniquely in the form: $D = i_1\mathcal{P}_1 + \ldots + i_s\mathcal{P}_s$, with $\mathcal{P}_1, \ldots, \mathcal{P}_s$ prime divisors.
3. Let a_d denote the number of prime divisors of degree d. Let $P \in C(\bar{K})$. Then P is defined over $\mathbb{F}_{p^m}$ if and only if its degree d divides m. Moreover, we can divide points of fixed degree d into disjoint sets of d elements, according to the prime divisor in which they appear (with non zero coefficient). Hence we deduce the following equality:

$$N_m = \sum_{d|m} da_d$$

Let us transform a bit our zeta function exploiting the above remarks. We have:

$$\begin{aligned}
\frac{d}{dt}(log(Z(t))) =& \frac{1}{t}\sum_{m=1}^{\infty} N_m t^m = \\
=& \frac{1}{t}\sum_{m=1}^{\infty}\left(\sum_{d|m} da_d\right)t^m = \frac{1}{t}\sum_{m=1}^{\infty} ma_m\left(\sum_{i=1}^{\infty} t^{mi}\right) = \\
=& \frac{1}{t}\sum_{m=1}^{\infty}\frac{ma_m t^m}{1-t^m} = \frac{d}{dt}\left(log\prod_{m=1}^{\infty}\left(\frac{1}{1-t^m}\right)^{a_m}\right)
\end{aligned}$$

Now, observe that

$$\prod_{m=1}^{\infty}\left(\frac{1}{1-t^m}\right)^{a_m} = \prod_{\mathcal{P} \text{ prime divisor}} \left(\frac{1}{1-t^{deg(\mathcal{P})}}\right)$$

Remark (2) above implies that the last expression equals

$$\sum_{D \in Div_K(C),\, D \geq 0} t^{deg(D)} = \sum_{m=0}^{\infty} A_m t^m$$

where $A_m = |\{D \in Div_K(C), D \geq 0, deg(D) = m\}|$ (it's easy to see that A_m is finite).
Summing up, we have obtained:

$$\frac{d}{dt}(log(Z(t))) = \frac{d}{dt}\left(log\big(\sum_{m=0}^{\infty} A_m t^m\big)\right)$$

From this it easily follows that $Z(t) = \sum_{m=0}^{\infty} A_m t^m$, since $log(Z(t))$ and $log\big(\sum_{m=0}^{\infty} A_m t^m\big)$ have the same derivative and the same value for $t = 0$.
So, we're left with the task of investigating the integers A_m in the above sum. This is where Riemann-Roch is going to help us.

Notation 4.1. 1. Let $\delta\mathbb{Z}$ be the subgroup of $\mathbb{Z}$ image of the map $deg : Div_K(C) \to \mathbb{Z}$.

2. Fix a divisor $D_0 \in Div_K(C)$ of degree δ.

3. Choose $\nu \in \mathbb{N}$ such that $(\nu - 1)\delta < g \leq \nu\delta$ where g is the genus of the curve.

4. Let $\{D_1, \dots, D_h\}$ be a maximal set of positive non equivalent divisors in $Div_K(C)$ of degree $\nu\delta$.

5. Choose a canonical divisor $K \in Div_K(C)$ and let $\mu \in \mathbb{N}$ such that $\mu\delta = 2g - 2 = deg(K)$.

Now, let $D \in Div_K(C)$ of degree $\nu\delta$. By Riemann-Roch:

$$l_K(D) \geq deg(D) + 1 - g \geq 1$$

hence there exists $f \in K(C)^*$ such that $div(f) + D \geq 0$. Then $div(f) + D$ must be equivalent to one of the D_i. Hence $D \sim D_i$ for some i, $1 \leq i \leq h$. Moreover such a D_i is unique, as $D_i \sim D$ and $D_j \sim D$ implies $D_i \sim D_j$, absurd. Hence, each divisor of degree $\nu\delta$ is equivalent to one and only one of the D_i.
Now, if $D \in Div_K(C)$ has degree $n\delta$, then $D - (n-\nu)D_0$ has degree $\nu\delta$, hence there is a unique i such that D is equivalent to $(n-\nu)D_0 + D_i$.
Finally, it follows easily from the definition that the number of positive divisors in $Div_K(C)$ which are equivalent to D is $\frac{p^{l_K(D)-1}}{p-1}$.
To simplify the notation, let us denote $l_K(D)$ simply by $l(D)$. The above observations allow us to write the zeta function in the form:

$$Z(t) = \sum_{n=0}^{\infty}\left(\sum_{i=1}^{h} \frac{p^{l(D_i+(n-\nu)D_0)} - 1}{p-1}\right) t^{n\delta}$$

Now, denote

$$\begin{aligned} Z_1(t) =& \sum_{n=0}^{\mu} \left(\sum_{i=1}^{h} \frac{p^{l(D_i+(n-\nu)D_0)}}{p-1} \right) t^{n\delta} \\ Z_2(t) =& Z(t) - Z_1(t) \end{aligned}$$

Then an easy computation gives:

$$Z_2(t) = \sum_{i=1}^{h} \sum_{n=\mu+1}^{\infty} \frac{p^{l(D_i+(n-\nu)D_0)}}{p-1} t^{n\delta} - \frac{h}{p-1} \sum_{n=0}^{\infty} t^{n\delta}$$

The key fact at this point is that for $n > \mu$ we have $deg(D_i+(n-\nu)D_0) = n\delta > \mu\delta = 2g-2$, hence $l(D_i + (n-\nu)D_0) = deg(D_i + (n-\nu)D_0) + 1 - g = n\delta + 1 - g$ by Corollary 2.2. Then

$$\begin{aligned} Z_2(t) =& \sum_{i=1}^{h} \sum_{n=\mu+1}^{\infty} \frac{p^{n\delta+1-g}}{p-1} t^{n\delta} - \frac{h}{p-1} \sum_{n=0}^{\infty} t^{n\delta} = \\ =& \sum_{i=1}^{h} \sum_{n=\mu+1}^{\infty} \frac{p^{1-g}}{p-1} (pt)^{n\delta} - \frac{h}{p-1} \sum_{n=0}^{\infty} t^{n\delta} = \\ =& \frac{hp^{1-g}}{p-1} \sum_{n=\mu+1}^{\infty} (pt)^{n\delta} - \frac{h}{p-1} \sum_{n=0}^{\infty} t^{n\delta} = \\ =& \frac{h}{p-1} \Big(\frac{p^{1-g}(pt)^{(\mu+1)\delta}}{1-(pt)^{\delta}} - \frac{1}{1-t^{\delta}} \Big) \end{aligned}$$

Let us stop and think about what we've done up to now: we had our zeta function, written in the form

$$Z(t) = \sum_{n=0}^{\infty} \left(\sum_{i=1}^{h} \frac{p^{l(D_i+(n-\nu)D_0)} - 1}{p-1} \right) t^{n\delta}$$

We decomposed it into two pieces, Z_1 and Z_2. The first one is a polynomial in t fo degree $\mu\delta = 2g-2$. On the other hand,

$$Z_2(t) = \sum_{i=1}^{h} \sum_{n=\mu+1}^{\infty} \frac{p^{l(D_i+(n-\nu)D_0)}}{p-1} t^{n\delta} - \frac{h}{p-1} \sum_{n=0}^{\infty} t^{n\delta}$$

is *a priori* an infinite series. However, the fact that $n > \mu$ in the sum allows us to apply Riemann-Roch, and we discover that we're actually dealing with a geometric series, whose sum we can compute explicitly, and turns out to be a rational function of t.
Therefore, our computation lead us to show that the zeta function of a smooth projective curve is a *rational function*. Also note that everything we made for a curve $C = C(\mathbb{F}_p)$ in order to obtain this result works in exactly the same way for a curve $C = C(\mathbb{F}_{p^m})$ (just replace p with $q = p^m$).

Lemma 4.1. *The zeta function satisfies the identity:*

$$Z_{C(\mathbb{F}_{p^d})}(t^d) = \prod_{\epsilon^d=1} Z_{C(\mathbb{F}_p)}(\epsilon t)$$

Proof. The right hand side equals:

$$exp\left(\sum_{m=1}^{\infty} N_m \frac{t^m}{m} (\sum_{\epsilon^d=1} \epsilon^m)\right)$$

Now use the fact that $\sum_{\epsilon^d=1} \epsilon^m = 0$ if $d \nmid m$, d if $d \mid m$. □

Theorem 4.1. *The function $Z(t)$ may be written as*

$$Z(t) = \frac{P(t)}{(1-t)(1-pt)}$$

with $P(t) \in \mathbb{Z}[t]$ of degree $2g$.

Proof. The above computations show that

$$\begin{aligned} Z(t) =& Z_1(t) + Z_2(t) = \sum_{n=0}^{\mu} \Big(\sum_{i=1}^{h} \frac{p^{l(D_i+(n-\nu)D_0)}}{p-1} \Big) t^{n\delta} + \frac{h}{p-1} \Big(\frac{p^{1-g}(pt)^{(\mu+1)\delta}}{1-(pt)^{\delta}} - \frac{1}{1-t^{\delta}} \Big) = \\ =& \frac{P(t)}{(1-t^{\delta})(1-(pt)^{\delta})} \end{aligned} \tag{4.1}$$

for some $P(t) \in \mathbb{Z}[t]$. Moreover, Z_1 is a polynomial in t, whereas Z_2 has a simple pole at those t for which $t^{\delta} = 1$. Hence $Z_{C(\mathbb{F}_{p^{\delta}})}(t^{\delta})$ has a simple pole at each t such that $t^{\delta} = 1$. Now use Lemma 4.1 with $d = \delta$ to get:

$$Z_{C(\mathbb{F}_{p^{\delta}})}(t^{\delta}) = \prod_{\epsilon^{\delta}=1} \frac{P(\epsilon t)}{(1-(\epsilon t)^{\delta})(1-(\epsilon tp)^{\delta})} = \frac{\prod_{\epsilon^{\delta}=1} P(\epsilon t)}{(1-t^{\delta})^{\delta}(1-(tp)^{\delta})^{\delta}}$$

hence $Z_{C(\mathbb{F}_{p^{\delta}})}(t^{\delta})$ has a pole of order δ at each t such that $t^{\delta} = 1$, so $\delta = 1$. Formula 4.1 with $\delta = 1$ becomes:

$$Z(t) = Z_1(t) + \frac{h}{p-1} \Big(\frac{p^{1-g}(pt)^{(\mu+1)}}{1-pt} - \frac{1}{1-t} \Big)$$

where Z_1 is a polynomial of degree $\mu = 2g-2$ in t. Hence

$$Z(t) = \frac{P(t)}{(1-t)(1-pt)}$$

with $deg(P(t)) = 2g$ □

As clearly $Z(0) = 1$, we obtain the following

Corollary 4.1. $Z(t) = \frac{\prod_{i=1}^{2g}(1-\alpha_i t)}{(1-t)(1-pt)}$ *where $\alpha_1, \ldots, \alpha_{2g} \in \mathbb{C}$ are the inverses of the zeros of $Z(t)$.*

Corollary 4.2. $N_m = 1 + p^m - \sum_{i=1}^{2g} \alpha_i^m$

Proof. By the previous Corollary, we find:

$$\frac{tZ'(t)}{Z(t)} = \sum_{m=1}^{\infty} \Big(1 + p^m - \sum_{i=1}^{2g} \alpha_i^m\Big) t^m$$

On the other hand

$$Z(t) = exp\Big(\sum_{M=1}^{\infty} \frac{N_m t^m}{m}\Big) \Rightarrow \frac{tZ'(t)}{Z(t)} = \sum_{m=1}^{\infty} N_m t^m$$

Comparing coefficients we get the desired result. $\square$

The fact that the zeta function is rational gives us information about the rate of growth of the number of points on a curve defined over $\mathbb{F}_{p^m}$ for growing m. Note that this number is $p^m + 1 = |\mathbb{P}^1(\mathbb{F}_{p^m})|$ plus an "error term", which depends on the zeros of the zeta function. Thus if we can obtain good bounds for the size of these zeros, we will have an hopefully sharp estimate of the "error term".
In the case $g = 1$ (corresponding to curves of degree 3 by Proposition 2.3) the formula for the zeta function turns out to be very simple: recall that

$$Z(t) = Z_1(t) + Z_2(t) = \sum_{n=0}^{\mu} \Big(\sum_{i=1}^{h} \frac{p^{l(D_i + (n-\nu)D_0)}}{p-1}\Big) t^n + \frac{h}{p-1}\Big(\frac{p^{1-g}(pt)^{(\mu+1)}}{1-pt} - \frac{1}{1-t}\Big)$$

where $\mu = 2g - 2$. So $\mu = 0$ if $g = 1$. Hence the above formula becomes:

$$Z(t) = \Big(\sum_{i=1}^{h} \frac{p^{l(D_i - \nu D_0)}}{p-1}\Big) t^0 + \frac{h}{p-1}\Big(\frac{pt}{1-pt} - \frac{1}{1-t}\Big) = A + \frac{h}{p-1}\Big(\frac{pt}{1-pt} - \frac{1}{1-t}\Big)$$

where $A = \Big(\sum_{i=1}^{h} \frac{p^{l(D_i - \nu D_0)}}{p-1}\Big)$ is a constant not depending on t. Hence we have:

$$\begin{aligned} Z(\frac{1}{pt}) =& A + \frac{h}{p-1}\Big(\frac{(p(\frac{1}{pt}))}{1-p(\frac{1}{pt})} - \frac{1}{1-\frac{1}{pt}}\Big) = \\ =& A + \frac{h}{p-1}\Big(\frac{\frac{1}{t}}{1-\frac{1}{t}} - \frac{pt}{pt-1}\Big) = \\ =& A + \frac{h}{p-1}\Big(\frac{1}{t-1} - \frac{pt}{pt-1}\Big) = Z(t) \end{aligned}$$

Moreover, we know that

$$Z(t) = \frac{P(t)}{(1-t)(1-pt)} = \frac{1 - a_p t + b_p t^2}{(1-t)(1-pt)}$$

The condition $Z(t) = Z(\frac{1}{pt})$ implies:

$$\frac{1 - a_p t + b_p t^2}{(1-t)(1-pt)} = \frac{1 - a_p \frac{1}{pt} + b_p \frac{1}{(pt)^2}}{(1-\frac{1}{pt})(1-p\frac{1}{pt})} =$$
$$= \frac{1}{p}\frac{p^2t^2 - a_p pt + b_p}{(pt-1)(t-1)} = \frac{pt^2 - a_p t + \frac{b_p}{p}}{(pt-1)(t-1)}$$

The last equality holds if and only if $b_p = p$. Hence for curves of degree 3 the zeta function has the form:

$$Z(t) = \frac{1 - a_p t + pt^2}{(1-t)(1-pt)} \tag{4.2}$$

Now, it can be proven that $|a_p| \leq 2\sqrt{p}$, which implies that $1 - a_p t + pt^2$ has two complex conjugate zeros; therefore we obtain:

$$Z(t) = \frac{(1-\pi t)(1-\bar{\pi} t)}{(1-t)(1-pt)}$$

where $\pi, \bar{\pi}$ are the inverses of the zeros of $Z(t)$. Clearly, we also have $\pi\bar{\pi} = p \Rightarrow |\pi| = \sqrt{p}$. A direct proof of these facts for two particular families of elliptic curves (which are curves of degree 3) will be given in Chapter 5.
All the results we obtained in this paragraph are special cases of the following amazing Theorem:

Theorem 4.2. *Let $C(\mathbb{F}_p)$ be a smooth projective curve of genus g. Then the zeta function $Z(t)$ of C satisfies the following properties:*

1. *(analytic continuation) $Z(t)$ is a rational function, and can be written in the form:*

$$Z(t) = \frac{\prod_{i=1}^{2g}(1-\alpha_i t)}{(1-t)(1-pt)}$$

2. *(functional equation) $Z(\frac{1}{pt}) = (pT^2)^{1-g} Z(t)$*

3. *(Riemann hypothesis) $\alpha_1, \ldots, \alpha_{2g}$ satisfy $|\alpha_i| = \sqrt{p}$*

As a corollary, we obtain the following:

Theorem 4.3. *(Hasse) The number of points on a projective smooth curve $C(\bar{\mathbb{F}}_p)$ of genus g defined over $\mathbb{F}_{p^m}$ satisfy the inequality:*

$$|N_m - p^m - 1| \leq 2g\sqrt{p^m}$$

Proof. The theorem follows immediately from Corollary 4.2 and from the Riemann hypothesis. □

Corollary 4.3. *Let C be a projective curve of degree 3 defined over $\mathbb{F}_p$. Then for each $m \in \mathbb{N}$ there is at least a point on C defined over $\mathbb{F}_{p^m}$.*

Proof. As in this case $g = 1$ we obtain $|N_m - p^m - 1| \le 2\sqrt{p^m} \Rightarrow N_m = p^m + 1 + \epsilon$, with $|\epsilon| \le 2\sqrt{p^m} < p^m + 1$. The conclusion follows. □

This result is the analogue, for curves of degree 3, of Chevalley-Warning theorem (which asserts the same thing for curves of degree 2). Anyway, note that in this case the proof requires much more advanced techniques.
Moreover, the above statement is false for projective curves of degree greater than 3. For example, it's easy to check that the projective curve with equation $x^4 + y^4 + z^4 = 0$ has no points in $\mathbb{P}^2(\mathbb{F}_5)$.

Example 4.3. As an application of the previous results, we will now describe a counterexample to the local-global principle (Theorem 1.1) for a curve of degree 3. Precisely, we will show that the projective curve $C(\mathbb{Q}) = \{[x, y, z] \in \mathbb{P}^2(\mathbb{Q}) : 3x^3 + 4y^3 + 5z^3 = 0\}$ has a point in $\mathbb{P}^2(\mathbb{Q}_p)$ for each prime p and a point in $\mathbb{P}^2(\mathbb{R})$. The proof of the fact that it has no points in $\mathbb{P}^2(\mathbb{Q})$ is much more difficult, and will not be given here.
An easy computation shows that the reduced curve $C_p(\mathbb{F}_p) = \{[x, y, z] \in \mathbb{P}^2(\mathbb{F}_p) : \bar{3}x^3 + \bar{4}y^3 + \bar{5}z^3 = 0\}$ is smooth for $p \ne 2, 3$ or 5. Thus for such a p Corollary 4.3 tells us that there is at least a nontrivial solution of the equation $3x^3 + 4y^3 + 5z^3 \equiv 0 \pmod{p}$. By Hensel Lemma [1] this solution lifts to a nontrivial solution in $\mathbb{Z}_p$.
Let $p = 2$. Then the reduced curve $C_2(\mathbb{F}_2) = \{[x, y, z] \in \mathbb{P}^2(\mathbb{F}_2) : x^3 + z^3 = 0\}$ contains the point $[1, 0, 1]$, which lifts to a point on $C(\mathbb{Q}_2)$ thanks to Hensel Lemma.
For the same reason, the point $[1, 2, 0] \in C_5(\mathbb{F}_5) = \{[x, y, z] \in \mathbb{P}^2(\mathbb{F}_2) : 3x^3 + 4y^3 = 0\}$ lifts to a point on $C(\mathbb{Q}_5)$, and the point $[0, 1, 4] \in C_3(\mathbb{F}_3) = \{[x, y, z] \in \mathbb{P}^2(\mathbb{F}_3) : 4y^3 + 5y^3 = 0\}$ lifts to a point on $C(\mathbb{Q}_3)$.
Finally, $[0, \sqrt[3]{5/4}, -1] \in C(\mathbb{R})$.

Let us now explain why point (2) of Theorem 4.2 is called Riemann hypothesis: let $Z(t) = \frac{\prod_{i=1}^{2g}(1-\alpha_i t)}{(1-t)(1-pt)}$. By the change of variables $t = p^{-s}$ we obtain:

$$Z(s) = \frac{\prod_{i=1}^{2g}(1 - \alpha_i p^{-s})}{(1 - p^{-s})(1 - p^{1-s})}$$

The zeros of $Z(s)$ are: $\beta_i = a_i + ib_i$ such that $p^{\beta_i} = \alpha_i$. Thus $|\alpha_i| = p^{a_i} = p^{Re\,\beta_i}$. Hence property (2) in Theorem 4.2 is equivalent to the fact that the zeros of $Z(s)$ have real part $\frac{1}{2}$.
Remark 4.2. Theorem 4.2 was first proved by André Weil in the 1940*s*. Weil also conjectured that the same results were true for arbitrary algebraic varieties. We're not going to explain here what exactly are algebraic varieties. And in fact, the formulation of this concept in an appropriate language was one of the first problems of algebraic geometers trying to prove Weil conjectures. This gave birth to many of the most important ideas in modern algebraic geometry (for example the concept of scheme), which allowed to prove the rationality of the zeta function for an arbitrary algebraic variety (Dwork, 1960) and the fact that this function satisfies a certain functional equation. The hardest statement to prove was Riemann hypothesis for arbitrary algebraic varieties, which was proved by

[1] See [17, Theorem 1, pag. 15].

Deligne in 1974.
Some material about Weil conjectures can be found in [6, Appendix C].

In the following Chapter we're going to restrict our attention to a particular type of smooth projective curves of degree 3, called elliptic curves. We're going to compute explicitly the zeta function of some families of elliptic curves using the techniques developed so far. This will allow us to verify directly the validity if the Riemann hypothesis for these curves.
After that, we will introduce a new type of function associated to a curve, namely its L-function.

Chapter 5

Elliptic curves

We will now focus on a particular class of projective smooth curves, called *elliptic curves*. Observe that Example 4.2 completely solves the problem of determining the zeta function of a smooth curve of degree $n = 1$ and 2. The following natural case to study is that corresponding to $n = 3$.
Let K be a field.

Definition 5.1. An *elliptic curve* is a smooth projective curve $E = C(K)$ of degree 3 defined over K, together with a point $P \in C(K)$.

Remark 5.1. It follows from Proposition 2.3 that elliptic curves are curves of genus one.

Let $K = \mathbb{Q}$. In this case, it can be proven (see [19]) that with a change of coordinates we can always transform our curve in the form:

$$E = C(\mathbb{Q}) = \{[x_0, x_1, x_2] \in \mathbb{P}^2(\mathbb{Q}) : x_0x_2^2 = x_1^3 - Ax_0^2x_1 + Bx_0^3\}$$

with $A, B \in \mathbb{Q}$. This is called the *Weierstrass form* of the curve. The affine equation of the curve obtained by setting $x = x_1/x_0, y = x_2/x_0$ is:

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

Note that there is only one point at infinity, namely $[0, 0, 1]$.
The transformation $(x, y) \mapsto (c^2x, c^3y)$ transforms the equation into $y^2 = x^3 - c^4Ax + c^6B$. Thus, we may assume that $A, B \in \mathbb{Z}$. We will always suppose this from now on.
The number $\Delta = 16(4A^3 - 27B^2)$ is called the *discriminant* of the curve. An easy calculation shows that E is a smooth curve if and only if its discriminant is not zero.
Reducing A, B modulo a prime p we obtain a curve $E_p = \{(x, y) \in \mathbb{A}^2(\mathbb{F}_p) : y^2 = x^3 - \bar{A}x + \bar{B}\}$. This is a smooth affine curve if and only if $p \nmid \Delta$. If this is the case the projective completion of E_p, with equation $x_0x_2^2 = x_1^3 - \bar{A}x_0^2x_1 + \bar{B}x_0^3$, is also smooth. We will still denote it by E_p.
Primes p for which the reduced curve E_p is smooth are called primes of *good reduction* for E.
Let p be a prime of good reduction for E. By Equation 4.2 the zeta function of E_p, called the *local zeta function* of E at p, is the rational function

$$Z_{E_p}(t) = \frac{P(t)}{(1-t)(1-pt)} = \frac{pt^2 - a_pt + 1}{(1-t)(1-pt)} = \frac{(1-\pi t)(1-\bar{\pi}t)}{(1-t)(1-pt)}$$

Moreover, by Corollary 4.2 we have:

$$N_{p^m} = p^m + 1 - \pi^m - \bar{\pi}^m$$

where N_{p^m} denotes the number of points on E_p defined over $\mathbb{F}_{p^m}$. In particular $N_p = p + 1 - a_p$. Thus, if one calculates N_p this also determines a_p and so π and $\bar{\pi}$, which are the inverses of the roots of the polynomial $pt^2 - a_p t + 1$. Hence the number of points defined over $\mathbb{F}_p$ of the elliptic curve E_p uniquely determines its zeta function and the number of points on E_p defined on each $\mathbb{F}_{p^m}$.

Example 5.1. In Example 3.1 we computed the number of points on the curve E of projective equation $f(x, y, z) = x_0^3 + x_1^3 - x_2^3 = 0$ in $\mathbb{P}^2(\mathbb{F}_{97})$, and found out that they are $N_{97} = 117$.
It is easy to see that this curve is smooth, as $\partial f/\partial x_0(x_0, x_1, x_2) = \partial f/\partial x_1(x_0, x_1, x_2) = \partial f/\partial x_2(x_0, x_1, x_2) = 0 \Leftrightarrow x_0 = x_1 = x_2 = 0$. Thus E is an elliptic curve, and we have: $a_p = 98 - 117 = -19$. From this we obtain $\pi = (-19 + 3\sqrt{-3})/2$. Hence, for example, $N_{9409} = N_{97^2} = 9409 + 1 - \pi^2 - \bar{\pi}^2 = 9243$.

Now let us compute explicitly the zeta functions of two particular classes of elliptic curves.

5.0.2 The curve $y^2 = x^3 + D$

Let $D \in \mathbb{Z}$, $D \neq 0$. Then the curve $E = E(\mathbb{Q})$ defined by the equation $x_0x_2^2 - x_1^3 - Dx_0^3 = 0$ is smooth. The discriminant of E is $\Delta = -2^4 3^3 D^2$, so we will only consider primes different from 2 and 3 and not dividing D, which are primes of good reduction for E.
The affine equation of E_p is $y^2 = x^3 + \bar{D}$. There is only one point at infinity, hence the number of $\mathbb{F}_p$-points on E_p is $N_p = 1 + N(y^2 = x^3 + \bar{D})$ (we will forget the bar in what follows).
If $p \equiv 2 \pmod 3$ then $x \mapsto x^3$ is an automorphism of $\mathbb{F}_p^*$, hence $N(y^2 = x^3 + D) = N(y^2 = x + D) = p$. The last equality follows from Example 4.2 and the observation that $y^2 = x + D$ is the equation of an affine smooth curve of degree 2 with only one point at infinity. Thus, for $p \equiv 2 \pmod 3$, $N_p = p + 1$.
Let $p \equiv 1 \pmod 3$ and let χ be a character of order 3 on $\mathbb{F}_p$, ρ a character of order 2. Then:

$$\begin{aligned} N(y^2 = x^3 + D) = \sum_{u+v=D} N(y^2 = u)N(x^3 = -v) = \sum_{u+v=D} (1 + \rho(u))(1 + \chi(-v) + \chi^2(-v)) = \\ = p + \sum_{u+v=D} \rho(u)\chi(v) + \sum_{u+v=D} \rho(u)\chi^2(v) \end{aligned}$$

because of Proposition 3.1 and the fact that $\chi(-1) = 1$.
Setting $u = Du'$, $v = Dv'$ we find:

$$N_p = p + 1 + \rho\chi(D)J(\rho, \chi) + \overline{\rho\chi(D)J(\rho, \chi)}$$

By Lemma 3.3 we obtain

$$N_p = p + 1 + \rho\chi(4D)J(\chi, \chi) + \overline{\rho\chi(4D)J(\chi, \chi)}$$

Let us now specify ρ and χ. Since $p \equiv 1 \pmod 3$, $p = \pi\bar{\pi}$ in $\mathbb{Z}[\omega]$, and we can take $\pi, \bar{\pi}$ primary. Let $(a/\pi)_6$ be the sixth power residue symbol (defined in the same way as the cubic and biquadratic residue symbol) and take $\rho(a) = (a/\pi)_6^3$ and $\chi(a) = (a/\pi)_6^2 = (a/\pi)_3$. Then by Proposition 3.8 we obtain:

$$N_p = p + 1 + \overline{\left(\frac{4D}{\pi}\right)_6^5}\pi + \left(\frac{\overline{4D}}{\pi}\right)_6^5\bar{\pi} = p + 1 + \overline{\left(\frac{4D}{\pi}\right)_6}\pi + \left(\frac{4D}{\pi}\right)_6\bar{\pi}$$

Let us collect the results we obtained:

Theorem 5.1. *Let $p \neq 2$ or 3 and $p \nmid D$. The number of projective points on the elliptic curve $x_0x_2^2 - x_1^3 - Dx_0^3 = 0$ which are defined over $\mathbb{F}_p$ is:*

1. *If $p \equiv 2 \pmod 3$, $N_p = p + 1$;*
2. *If $p \equiv 1 \pmod 3$, let $p = \pi\bar{\pi}$ in $\mathbb{Z}[\omega]$, with π primary. Then*

$$N_p = p + 1 + \overline{\left(\frac{4D}{\pi}\right)_6}\pi + \left(\frac{4D}{\pi}\right)_6\bar{\pi}$$

Thanks to this result we're able to determine the local zeta function of E for all primes p for which the reduced curve E_p is smooth.
Recall that we have:

$$Z_{E_p}(t) = \frac{1 - a_pt + pt^2}{(1-t)(1-pt)}$$

and $a_p = p + 1 - N_p$. If $p \equiv 2 \pmod 3$ the above formula implies that $a_p = 0$. If $p \equiv 1 \pmod 3$ then:

$$|a_p| = \left| \overline{\left(\frac{4D}{\pi}\right)_6}\pi + \left(\frac{4D}{\pi}\right)_6\bar{\pi} \right| \leq 2|\pi| = 2\sqrt{p}$$

Thus in both cases $|a_p| \leq 2\sqrt{p}$. As we already pointed out, this implies that the polynomial $1 - a_pt + pt^2$ has two complex conjugate zeros, hence the local zeta function of E at a prime p of good reduction has the form:

$$Z_{E_p}(t) = \frac{(1-\pi t)(1-\bar{\pi}t)}{(1-t)(1-pt)}$$

with $|\pi| = \sqrt{p}$. Hence the Riemann hypothesis for elliptic curves of affine equation $y^2 = x^3 + D$ is proved.

Example 5.2. Consider the elliptic curve E of affine equation $y^2 = x^3 + 1$. Let us determine the number of points of its reduction modulo 19. We have $19 = (5+3\omega)(5+3\omega^2)$ in $\mathbb{Z}[\omega]$. Hence:

$$N_{19} = 20 + \overline{\left(\frac{4}{5+3\omega}\right)_6}(5+3\omega) + \left(\frac{4}{5+3\omega}\right)_6(5+3\omega^2)$$

It remains to calculate

$$\left(\frac{4}{5+3\omega}\right)_6 = \left(\frac{2}{5+3\omega}\right)_3$$

We have

$$2^{(19-1)/3} = 2^6 = 64 \equiv 64 - 19 \times 3 \equiv 7 \pmod{(5+3\omega)}$$

Direct calculation shows that $(5+3\omega) \mid (7-\omega^2)$, hence $(4/(5+3\omega))_6 = \omega^2$. Hence, we obtain:

$$N_{19} = 20 + \omega(5+3\omega) + \omega^2(5+3\omega^2) = 20 + 8(\omega+\omega^2) = 20 + 8(2Re\ \omega) = 20 - 8 = 12$$

The local zeta function of E at $p = 19$ is therefore:

$$Z_{E_{19}}(t) = \frac{19t^2 - a_{19}t + 1}{(1-t)(1-pt)}$$

where $a_{19} = 19 + 1 - N_{19} = 20 - 12 = 8$. So

$$Z_{E_{19}}(t) = \frac{19t^2 - 8t + 1}{(1-t)(1-pt)} = \frac{(1-\pi t)(1-\bar{\pi}t)}{(1-t)(1-pt)}$$

where $\pi = (4+\sqrt{-3})$. Now we can determine with no extra work the number of points on our curve defined over an arbitrary finite extension $\mathbb{F}_{19^m}$ of $\mathbb{F}_{19}$. For example: $N_{130321} = N_{19^4} = 19^4 + 1 - \pi^4 - \bar{\pi}^4 = 130368$.

5.0.3 The curve $y^2 = x^3 - Dx$

Consider the elliptic curve E defined by $x_0x_2^2 - x_1^3 + Dx_1x_0^2 = 0$, where $D \in \mathbb{Z}$, $D \neq 0$. In affine coordinates E has equation $y^2 = x^3 - Dx$, and only one point at infinity. The discriminant of E is $\Delta = 2^6D^3$, hence we will only consider odd primes not dividing D, which are of good reduction for E.
Let p be such a prime. Let us determine the number N_p of points on the reduced curve E_p defined over $\mathbb{F}_p$.
Notice that the polynomial $p(x) = x^3 - Dx$ is odd.
If $p \equiv 3 \pmod 4$ then -1 is not a square in $\mathbb{F}_p$. $p(x)$ has 1 or 3 zeros, depending on whether D is or not a square in $\mathbb{F}_p$.
Take $x \in \mathbb{F}_p$ such that $p(x) \neq 0$. Then either $p(x)$ is a square or $p(-x) = -p(x)$ is a square in $\mathbb{F}_p$. In both cases, there are exactly two points on E_p with first coordinate equal to x or $-x$. So, if D is a square we have $2\frac{p-3}{2} + 3 = p$ affine points on E_p. If D is not a square, we find $2\frac{p-1}{2} + 1$ affine points on E_p. In both cases we have $N_p = p + 1$.
Let $p \equiv 1 \pmod 4$. We want to use again the technique of Jacobi sums in order to count points on E_p. It should be clear from the examples we dealt with so far that this method works for "diagonal" equations such as $x^3 + y^3 - 1 = 0$. Let us transform our equation in this form. If C denotes the curve of equation $u^2 = v^4 + 4D$, it's easy to see that the transformation

$$T(u,v) = \big((u+v^2)/2, v(u+v^2)/2\big)$$

maps bijectively C to $E \setminus (0,0)$ (the inverse map being $(x,y) \mapsto (2x - y^2/x^2, y/x)$). Therefore, we have $N(y^2 = x^3 - Dx) - 1 = N(u^2 = v^4 + 4D)$.

Let λ be a character of order 4 on $\mathbb{F}_p$, $\rho = \lambda^2$. Then we find:

$$\begin{aligned} N(u^2 = v^4 + 4D) = & \sum_{r+s=4D} N(u^2 = r)N(v^4 = -s) = \\ & = \sum_{r+s=4D} (1+\rho(r))(1+\lambda(-s)+\lambda^2(-s)+\lambda^3(-s)) = \\ & = p + J(\rho,\rho) + \rho\lambda(-4D)J(\rho,\lambda) + \rho\lambda^3(-4D)J(\rho,\lambda^3) = \\ & = p - 1 + \overline{\lambda(-4D)}J(\rho,\lambda) + \lambda(-4D)\overline{J(\rho,\lambda)} = \\ & = p - 1 + \overline{\lambda(D)}\lambda(-1)J(\lambda,\lambda) + \lambda(D)\overline{\lambda(-1)J(\lambda,\lambda)} \end{aligned}$$

since by Lemma 3.3 we have $J(\rho,\lambda) = \lambda(4)J(\lambda,\lambda)$.
Now, since $\pi \equiv 1 \pmod 4$ we can write $p = \pi\bar{\pi}$ in $\mathbb{Z}[i]$, where π is primary. Choosing $\lambda(a) = (a/\pi)_4$ (the biquadratic residue on $\mathbb{Z}[i]/\pi\mathbb{Z}[i]$) we have by Proposition 3.10 $-\lambda(-1)J(\lambda,\lambda) = \pi$. Substituting this in the above equation and collecting all the information we've obtained we arrive at:

Theorem 5.2. *Let $p \neq 2$ and $p \nmid D$. The number of projective points on the elliptic curve $x_0x_2^2 - x_1^3 + Dx_1x_0^2 = 0$ defined over $\mathbb{F}_p$ is:*

1. *If $p \equiv 3 \pmod 4$, $N_p = p + 1$;*
2. *If $p \equiv 1 \pmod 4$, let $p = \pi\bar{\pi}$ in $\mathbb{Z}[i]$, with π primary. Then*

$$N_p = p + 1 - \overline{\left(\frac{D}{\pi}\right)_4}\pi - \left(\frac{D}{\pi}\right)_4\bar{\pi}$$

Again, this theorem allows us to determine the local zeta function of E for all p for which the reduced curve E_p is smooth. We obtain, as in the previous section, that:

$$Z_{E_p}(t) = \frac{(1-\pi t)(1-\bar{\pi}t)}{(1-t)(1-pt)}$$

with $|\pi| = \sqrt{p}$. This proves the Riemann hypothesis for elliptic curves of affine equation $y^2 = x^3 - Dx$.

Example 5.3. Let us determine the number of points of the reduction E_5 of the elliptic curve E whose affine equation is $y^2 = x^3 - 4x$.
We have: $5 = (2i-1)(-2i-1)$, and $2i - 1$ is primary. Hence:

$$N_5 = 5 + 1 - \overline{\left(\frac{4}{-1+2i}\right)_4}(-1+2i) - \left(\frac{4}{-1+2i}\right)_4(-1-2i)$$

Moreover: $(4/(-1+2i))_4 \equiv 4^{\frac{N(-1+2i)-1}{4}} \equiv 4 \equiv -1 \pmod{-1+2i}$.
Hence $N_5 = 6 + (-1+2i) + (-1-2i) = 4$.
In this simple case, one can verify directly that the points on E_5 which are defined over $\mathbb{F}_5$ are: $(0,0), (2,0), (-2,0), \infty$.
The local zeta function of E at $p = 5$ is:

$$Z_{E_5}(t) = \frac{5t^2 - 2t + 1}{(1-t)(1-5t)} = \frac{(1-(1+2i)t)(1-(1-2i)t)}{(1-t)(1-5t)}$$

Hence we have for example: $N_{625} = 625 + 1 + (1+2i)^4 + (1-2i)^4 = 612$.

Remark 5.2. Take an elliptic curve E defined over $\mathbb{Q}$ with affine equation $y^2 = x^3 + D$ or $y^2 = x^3 - Dx$. Let p be a prime of good reduction for E. Let N_p be the number of points on the reduced curve E_p which are defined on $\mathbb{F}_p$. In the previous two paragraphs we verified directly that $N_p = p + 1 + \epsilon_p$, with $|\epsilon_p| \leq 2\sqrt{p}$ (Hasse Theorem). More precisely, when p varies among all primes of good reduction for E, Theorem 5.1 and Theorem 5.2 show that, roughly speaking, $\epsilon_p = 0$ for half of the primes p. Thus the "error term" ϵ_p is not uniformly distributed in the interval $[-2\sqrt{p}, 2\sqrt{p}]$.
The problem arises of determining the distribution of the error terms ϵ_p when p varies among all primes of good reduction for an arbitrary elliptic curve E defined over $\mathbb{Q}$. Let us consider, instead of ϵ_p, the normalised error $\sigma_p = \epsilon_p / 2\sqrt{p}$. As $|\sigma_p| \leq 1$, for each p there exist a unique $\theta_p \in [0, \pi]$ such that $\cos\theta_p = \sigma_p$. It has been conjectured that, for "most" elliptic curves, the values of θ_p are "uniformly distributed" in the interval $[0, \pi]$. To formulate exactly the conjecture, we need to clarify what it means for a sequence of real numbers to be "uniformly distributed". This is the content of the following definition:

Definition 5.2. Let $\{x_n\}_{n\in\mathbb{N}}$ be a sequence of points in the interval $[a, b] \subseteq \mathbb{R}$. Let μ be a measure on $[a, b]$. We say that $\{x_n\}$ is a *μ-equidistributed* sequence if, for every interval $I = [c, d] \subseteq [a, b]$,

$$\lim_{n\to\infty} \frac{|\{k \leq n \in \mathbb{N} : x_k \in I\}|}{n} = \frac{d-c}{b-a}$$

Conjecture 5.1. *(Sato-Tate) Let E be an elliptic curve defined over $\mathbb{Q}$, without complex multiplication.*[1] *Then θ_p is a μ-equidistributed sequence with respect to the probability measure $\mu = \frac{2}{\pi}\sin^2\theta d\theta$ on the interval $[0, \pi]$.*

This is actually no more a conjecture: in fact, it has recently been proven by L. Clozel, M. Harris, N. Shepherd-Barron, R. Taylor.

In the next section we're going to collect local information in order to build the global zeta function of a given elliptic curve, and the closely related L-function. The study of this new object, as we will soon see, originates problems whose importance in Number Theory could hardly be overestimated.

5.1 The L-function

Let $E = \{[x_0, x_1, x_2] \in \mathbb{P}^2(\mathbb{Q}) : x_0 x_2^2 = x_1^3 - A x_0^2 x_1 + B x_0^3\}$ be an elliptic curve. If p is a prime of good reduction for E, we have defined the local zeta function of E at p as the zeta function of the reduced curve E_p:

$$Z_{E_p}(t) = \frac{pt^2 - a_p t + 1}{(1-t)(1-pt)}$$

[1]See Section 5.2.3 for an explanation of what is complex multiplication. In the same paragraph we also notice that elliptic curves of the form $y^2 = x^3 + D$ and $y^2 = x^3 - Dx$ *do* have complex multiplication. As for these curves the claim in the conjecture is clearly false, we see that the hypothesis of not having complex multiplication cannot be dropped.

By the change of variable $t = q^{-s}$ we obtain:

$$Z_{E_p}(s) = \frac{p^{1-2s} - a_p p^{-s} + 1}{(1-p^{-s})(1-p^{1-s})}$$

If p is *not* a prime of good reduction, that is, if $p \mid \Delta$, we define the local zeta function of E at p as[2]:

$$Z_{E_p}(s) = \frac{1}{(1-p^{-s})(1-p^{1-s})}$$

In this way, for each prime p we have defined the local zeta function of E at p. We define the global zeta function of the elliptic curve E as:

$$Z_E(s) = \prod_{p \in Max(\mathbb{Z})} Z_{E_p}(s)$$

Now recall that the Riemann zeta function can be written as: $\zeta(s) = \prod_{p \text{ prime}} (1-p^{-s})^{-1}$ (Proposition 4.1).
Let

$$L_E(s) = \prod_{p \nmid \Delta} (1 - a_p p^{-s} + p^{1-2s})^{-1}$$

Then we obtain the equality:

$$Z_E(s) = \zeta(s)\zeta(1-s)L_E(s)^{-1}$$

The function $L_E(s)$ we just defined is called the L-function of the elliptic curve E. We will denote it simply by $L(s)$ when this will not be too dangerous.

Lemma 5.1. *The product defining the L-function $L(s)$ converges in the half plane $Re\ s > 3/2$, and $L(s)$ is a holomorphic function on this half plane.*

Proof. Observe that, by Hasse theorem, for each $p \nmid \Delta$ the zeros of $1 - a_p p^{-s} + p^{1-2s}$ have real part $1/2$, hence $(1 - a_p p^{-s} + p^{1-2s})^{-1}$ is holomorphic and nonzero on the half plane $Re\ s > 3/2$. Recall that a product of nonzero holomorphic functions f_n is said to converge if the sum $\sum_n log(f_n)$ converges uniformly on compact sets.
We need to check that $-\sum_{p \nmid \Delta} log(1 - a_p p^{-s} + p^{1-2s})$ converges for $Re\ s > 3/2$.
By Hasse theorem we have $1 - a_p p^{-s} + p^{1-2s} = (1 - \pi p^{-s})(1 - \bar{\pi} p^{-s})$ with $|\pi| = \sqrt{p}$. Hence

$$\sum_{p \nmid \Delta} (log(1 - \pi p^{-s}) + log(1 - \bar{\pi} p^{-s}))$$

converges uniformly on compact sets K such that $|\pi p^{-s}| = p^{1/2 - Re\ s} \leq p^{-1-\delta}\ \forall s \in K$, for any fixed $\delta > 0$. It follows that our product converges for all s such that $1/2 - Re\ s < -1 \Leftrightarrow Re\ s > 3/2$. □

[2] Actually, this is not the "official" definition. To state it correctly, it would be necessary to investigate more in depth what happens when we reduce an elliptic curve at a prime which is not of good reduction. For simplicity, we shall not do this here and we will use this simplified form of the definition, which will be enough for us.

Of course, we're not satisfied with this result: we would like to know if it is possible to analytically continue $L(s)$ to a holomorphic function defined on the whole complex plane. Notice that by Proposition 4.1 $\zeta(s)$ can be analytically continued to a meromorphic function on $\mathbb{C}$. As $Z_E(s) = \zeta(s)\zeta(1-s)L_E(s)^{-1}$, our problem is strictly related to the problem of determining if the global zeta function $Z_E(s)$ of an elliptic curve E can be analytically continued to a meromorphic function defined on the whole complex plane. Anyway it turns out that it's easier to work with the L-function, which enjoys some remarkable properties.
In order to explain the importance of the L-function, and to show how its analytic properties are (or could be) linked to fundamental arithmetic properties of the corresponding curve, we're going to briefly discuss one of the most important open problems involving L-functions of elliptic curves.

5.1.1 The Birch and Swynnerton-Dyer conjecture

Let E be an elliptic curve defined over an arbitrary field K, with a distinguished point $O \in E(K)$. Poincaré first discovered that the points of $E(K)$ can be given an abelian group structure, defined as follows.
Let $P, Q \in E(K)$. By Bezout theorem the line through P and Q intersects $E(\bar{K})$ in a third point $R \in \mathbb{P}^2(\bar{K})$; denote this point by $R = P * Q$. It's easy to verify that actually if $P, Q \in E(K)$ then also $P * Q \in E(K)$. Now, for each pair of points $P, Q \in E(K)$ define $P + Q = O * (P * Q)$. Then $(E(K), +)$ is an abelian group.

Remark 5.3. The careful reader will have noticed that we've been quite imprecise in the definition of the group law given above: for example, if $P = Q$ the line through P and Q is by no means unique. In this case we have to consider the tangent line to E at P. The detailed description of the group law, as well as the proof that this law actually enjoys the properties of a group law, can be found for example in [3] or [18]. A very readable description of the group law from the complex analytic point of view is given in [9].

The natural problem at this point is to determine the structure of the group of points of an elliptic curve $E(K)$ for a fixed field K. The following remarkable theorem, proved by Mordell (1923) for $\mathbb{Q}$ and later generalised by Weil, answers this question for number fields.

Theorem 5.3. *(Mordell-Weil) Let $E(K)$ be an elliptic curve defined over a number field K. Then the group $(E(K), +)$ is a finitely generated abelian group.*

Proof. For a proof of the general case see [18]. An elementary proof in a special case, containing all the main ideas that are used in the general proof, can be found in [19]. □

The above theorem and the structure theorem for finitely generated abelian groups imply that, as a group, we have: $E(K) = T \oplus \mathbb{Z}^r$, where T denotes the torsion part of the group. The integer r is called the *rank* of the elliptic curve $E(K)$.
Here is where our L-function plays an unexpected, yet crucial role: suppose for a moment that it is true that the L-function of any elliptic curve E defined over $\mathbb{Q}$ can be analytically continued to a holomorphic function on $\mathbb{C}$. Then it makes sense to speak of the behaviour

of the (extended) L-function at the point $s = 1$. Based on an extensive empirical work with curves of the form $y^2 = x^3 - Dx$ and $y^2 = x^3 + D$ (for which we will see that our supposition is actually true) Birch and Swinnerton-Dyer stated the following

Conjecture 5.2. *(Birch, Swinnerton-Dyer) Let E be an elliptic curve defined over $\mathbb{Q}$. Then the rank of E is equal to the order of vanishing of $L(E, s)$ at $s = 1$.*

The above conjecture, if true, should imply the existence of an intimate relation between analytic properties of the L-function of an elliptic curve and arithmetic properties of that curve. [3]
There is actually a more precise form of this conjecture, but we will not discuss it here. The interested reader can refer to [7].
In a recent paper Manjul Bhargava (Fields Medal 2014), Christopher Skinner and Wei Zhang have shown that for "most" elliptic curves the Birch and Swinnerton-Dyer conjecture is true (See `http://arxiv.org/abs/1407.1826`).
We will now return to our original problem of determining if there exists an analytic continuation of the L-function of a given elliptic curve. The key idea is to look at the L-function from a completely different perspective.

5.2 Hecke L-functions

Hecke L-functions are an important family of functions associated with the so called *Hecke characters*, which are particular characters defined over number fields.
In order to motivate the introduction of this new concept, we shall quickly review the definition of the more classical Dirichlet L-functions. [4]

5.2.1 Dirichlet L-functions

Let ψ be a character on the abelian group $(\mathbb{Z}/m\mathbb{Z})^*$. Extend ψ letting $\psi(a) = 0$ $\forall a \in (\mathbb{Z}/m\mathbb{Z}) \setminus (\mathbb{Z}/m\mathbb{Z})^*$. Let $\pi : \mathbb{Z} \to \mathbb{Z}/m\mathbb{Z}$ denote the canonical projection. The composition $\chi = \psi \circ \pi : \mathbb{Z} \to \mathbb{C}$ has the following properties:

1. $\chi(n + m) = \chi(n)$ $\forall n \in \mathbb{Z}$
2. $\chi(n) \neq 0$ if and only if $(n, m) = 1$
3. $\chi(kn) = \chi(k)\chi(n)$ $\forall k, n \in \mathbb{Z}$

A function $\chi : \mathbb{Z} \to \mathbb{C}$ satisfying the three properties above is called a *Dirichlet character* modulo m.

[3] The prominent position of this conjecture in modern mathematical research is also proved by the fact that it has been chosen as one of the seven (maybe we should say six, after Perelman) Millenium Prize Problems.

[4] These functions were introduced by Dirichlet who used them to give the first proof of the fact that each arithmetic progression $a + nb$, with $(a, b) = 1$ contains infinitely many prime numbers. See [8] for an account of the proof.

To a given Dirichlet character χ we associate the corresponding *L-function* defined as follows:

$$L_\chi(s) = \sum_{n=1}^{\infty} \frac{\chi(n)}{n^s}$$

As $|\chi(s)| = 1$, we immediately see that $L_\chi(s)$ is an analytic function on the half plane $Re\ s > 1$. Moreover, as $\chi(kn) = \chi(k)\chi(n)$ we obtain the product formula:

$$L_\chi(s) = \prod_{p \text{ prime}} (1 - \chi(p)p^{-s})^{-1}$$

As for the zeta function, it can be shown that Dirichlet L-functions have a meromorphic continuation to $\mathbb{C}$, which satisfies a certain functional equation (see [16]).
Moreover, just like the zeta function, Dirichlet L-functions can be generalized to arbitrary number fields. In what follows, we will study one possible generalisation to *CM* fields, due to Hecke.

5.2.2 Hecke algebraic characters and L-functions

Definition 5.3. Let K be a number field. A morphism $\sigma : K \to \mathbb{C}$ is called *real* if $\sigma(K) \subseteq \mathbb{R}$, otherwise it is called *complex*.
K is called *totally real* if every morphism of K into $\mathbb{C}$ is real; it is called *totally complex* if every morphism of K into $\mathbb{C}$ is complex.
A *CM field* is a totally complex quadratic extension of a totally real subfield K_0.

Example 5.4.

1. If $d \in \mathbb{Z}, d > 0$ is squarefree then $\mathbb{Q}(\sqrt{-d})$ is a CM field. In particular, $\mathbb{Q}(i)$ and $\mathbb{Q}(\omega)$ are CM fields.

2. If $\zeta_m = e^{2\pi i/m}$, $\mathbb{Q}(\zeta_m)$ is a CM field, with totally real subfield $\mathbb{Q}(\zeta_m + \zeta_m^{-1})$.

Remark 5.4. Let K be a Galois extension of $\mathbb{Q}$ which is also a CM field. Let $j \in Gal(K/\mathbb{Q})$ denote the restriction to K of complex conjugation. Let L be the fixed field of j. Then clearly $L \supseteq K_0$, and $[K : L] = 2 = [K : K_0]$. Hence $L = K_0$.
Moreover, j clearly commutes with every $\sigma \in Gal(K_0/\mathbb{Q})$, so it commutes with each element in $Gal(K/\mathbb{Q})$. In this paragraph, we will always assume these hypotheses.

Notation 5.1. Let $G = Gal(K/\mathbb{Q})$. We will denote by $\mathbb{Z}[G]$ the *group ring* of G, whose elements are formal finite sums $\sum_{\sigma\in G} n_\sigma\sigma$ with $n_\sigma \in \mathbb{Z}$. This is a ring with addition and multiplication defined in the only reasonable way. It acts on K as follows: for $\theta = \sum_\sigma n_\sigma\sigma \in \mathbb{Z}[G], \alpha \in K$, $\alpha^\theta = \prod_\sigma \sigma(\alpha)^{n_\sigma}$.

Definition 5.4. Let K be a CM field which is also a Galois extension of $\mathbb{Q}$. Let $\mathcal{O}$ be the ring of algebraic integers of K and M a fixed ideal of $\mathcal{O}$.
An *algebraic Hecke character* modulo M is a function $\chi : \{I : I \subseteq \mathcal{O} \text{ ideal}\} \to \mathbb{C}$ that satisfies the following properties:

1. $\chi(\mathcal{O}) = 1$

2. $\chi(A) \neq 0$ if and only if $(A, M) = (1)$

3. $\chi(AB) = \chi(A)\chi(B)$

4. There is an element $\theta = \sum_\sigma n_\sigma \sigma \in \mathbb{Z}[G]$ such that if $\alpha \in \mathcal{O}, \alpha \equiv 1 \pmod{M}$, then $\chi((\alpha)) = \alpha^\theta$

5. There is an integer $m > 0$, called the *weight* of χ, such that $n_\sigma + n_{j\sigma} = m \ \forall \sigma \in G$.

Remark 5.5.

1. Properties (2),(3) of Hecke characters defined above are analogous to properties (2),(3) of Dirichlet characters.

2. It follows from condition (5) above that $(1+j)\sigma = mN$, where $N = \sum \sigma$ is the norm element in $\mathbb{Z}[G]$. The name is motivated by the fact that, for $\alpha \in K$, $\alpha^N = N(\alpha)$ where $N(\alpha)$ is the norm as defined in Section 1.4.

Proposition 5.1. *Let χ be an algebraic Hecke character of weight m. If $(A, M) = 1$, $|\chi(A)| = N(A)^{m/2}$.*

Proof. With the notations of Proposition 1.7, let $h = |C_M|$. Then h is finite by Proposition 1.7, hence there exist $\alpha, \beta \in \mathcal{O}$, $\alpha, \beta \equiv 1 \pmod{M}$ such that $(\alpha)A^h = (\beta)$. Applying χ we get:

$$\alpha^\theta \chi(A)^h = \beta^\theta$$

Taking complex conjugates and multiplying we obtain, recalling that $(1 + j)\sigma = mN$:

$$N(\alpha)^m |\chi(A)|^{2h} = N(\beta)^m$$

On the other hand, we have $(\alpha)A^h = (\beta) \Rightarrow N(\alpha)N(A)^h = N(\beta)$. Comparing the last two equations we obtain:

$$|\chi(A)|^{2h} = N(A)^{mh} \Rightarrow |\chi(A)| = N(A)^{m/2}$$

□

Remark 5.6. The equality $\alpha^\theta \chi(A)^h = \beta^\theta$ shows that the values $\chi(A)$ are algebraic numbers (roots of elements of K). This explains why we call χ an "algebraic" Hecke character.

Now that we've defined a character, we can attack an L-function to it. The definition is the obvious generalisation of the one we gave for Dirichlet L-functions:

Definition 5.5. Let χ be an algebraic Hecke character on a CM field K with ring of algebraic integers $\mathcal{O}$. The *Hecke L-function* associated to χ is:

$$L_\chi(s) = \sum_A \chi(A) N(A)^{-s} = \prod_P (1 - \chi(P)N(P)^{-s})^{-1}$$

where the sum is over all non zero ideals of $\mathcal{O}$, the product over all ideals $P \in Max(\mathcal{O})$ and equality of the two expressions follows from unique factorisation of ideals in $\mathcal{O}$ and the fact that $\chi(AB) = \chi(A)\chi(B)$.

Proposition 5.2. *The product $L_\chi(s) = \prod_P (1 - \chi(P)N(P)^{-s})^{-1}$ converges absolutely in the half plane $Re\ s > 1 + m/2$, uniformly for $Re\ s \geq 1 + m/2 + \delta$ for any $\delta > 0$.*

Proof. The product converges absolutely if and only if the series $\sum_P |log(1-\chi(P)N(P)^{-s})|$ converges. This is equivalent to the convergence of the series $\sum_P |\chi(P)N(P)^{-s}| =$
$= \sum_P N(P)^{m/2}N(P)^{-u} = \sum_P N(P)^{(m/2)-u}$, where $u = Re\ s$.
Now $N(P) = |\mathcal{O}/P| \geq p \Rightarrow N(P)^{(m/2)-u} \leq p^{-(u-m/2)}$ where p is the prime such that $P \cap \mathbb{Z} = (p)$. Since we know by Section 1.4 that each prime p as at most $[K:\mathbb{Q}]$ primes in $\mathcal{O}$ above it, we obtain:

$$\sum_P N(P)^{(m/2)-u} \leq [K:\mathbb{Q}] \sum_p p^{-(u-m/2)}$$

and the last sum converges for $Re\ s = u > 1 + m/2$. □

One of the reasons why Hecke L-functions are so important is that they have the property that we would like to be enjoyed by L-functions of elliptic curves:

Theorem 5.4. *Let χ be an algebraic Hecke character and $L_\chi(s)$ the corresponding L-function. If $\chi(A)$ is not equal to 1 for some ideal A, then $L_\chi(s)$ can be analytically continued to a holomorphic function defined on the whole $\mathbb{C}$.*

Proof. The original proof of this (difficult) Theorem is due to Hecke. A different proof has been given by John Tate in his thesis, which is reproduced in [1]. □

With this result in mind, let us return to our problem: we have an elliptic curve E with L-function $L_E(s)$. We want to show that $L_E(s)$ has an analytic continuation to the whole complex plane. We have just discovered that another, seemingly unrelated family of L-functions enjoys this property. There's only one thing we can try to do: show that $L_E(s)$ can be realised as the L-function associated with a certain Hecke character χ defined over the ring of algebraic integers of a well chosen CM field K. Let us try to do this with our two favourite families of elliptic curves.

The L-function of $y^2 = x^3 - Dx$

Let E be the curve whose affine equation is $y^2 = x^3 - Dx$, $D \in \mathbb{Z}$, $D \neq 0$. Since $\Delta = 2^6 D^3$ we have:

$$L_E(s) = \prod_{p \nmid 2D} (1 - a_p p^{-s} + p^{1-2s})^{-1}$$

where $a_p = p + 1 - N_p$ and N_p has been determined in Section 5.0.3. Recall that in that Section we worked with the ring $\mathbb{Z}[i]$ in order to count points on E. If we have a chance to find an Hecke character whose associated L-function coincides with that of the curve, there is no better place to look for it than the ring $\mathbb{Z}[i]$.
Precisely, we are going to construct an algebraic Hecke character χ on $\mathbb{Z}[i]$ with respect to the modulus $(8D)$. To define it, it is clearly sufficient to specify the value of χ on prime ideals (P) of $\mathbb{Z}[i]$.

1. If $P \mid 2D$, $\chi(P) = 0$.
2. If $P \nmid 2D$ and $N(P) = p \equiv 1 \pmod 4$, let $(P) = (\pi)$, with $\pi \equiv 1 \pmod{2+2i}$. Then $\chi(P) = \overline{(D/p)_4}\pi$. Let $\mathcal{P}_1$ denote the set of prime ideals (P) of this type.

3. If $P \nmid 2D$ and $N(P) = p^2$ then $p \equiv 3 \pmod 4$ and $(P) = (p)$. Then $\chi(P) = -p$. Let $\mathcal{P}_2$ denote the set of prime ideals (P) of this type.

Thanks to Remark 3.8 points (2) and (3) can be joined, and we can define $\chi(P)$ uniformly for primes (P) such that $P \nmid 2D$ as: $\chi(P) = \overline{(D/\pi)_4}\pi$ where $(P) = (\pi)$ and $\pi \equiv 1 \pmod{2+2i}$.
Now suppose for a moment that we knew that χ is an algebraic Hecke character. Then we would obtain:

$$L_E(s) = \prod_{p \nmid 2D, p\equiv 3 \pmod 4} (1 - a_p p^{-s} + p^{1-2s})^{-1} \prod_{p \nmid 2D, p \equiv 1 \pmod 4} (1 - a_p p^{-s} + p^{1-2s})^{-1}$$

By Theorem 5.2 we have: $N_p = p+1 \Rightarrow a_p = 0$ if $p \equiv 3 \pmod 4$; $a_p = \overline{(D/\pi)_4}\pi + (D/\pi)_4\bar{\pi}$ if $p \equiv 1 \pmod 4$, where $p = \pi\bar{\pi}, \pi \equiv 1 \pmod{2+2i}$ Hence we obtain:

$$\begin{aligned} L_E(s) = & \prod_{p \nmid 2D,\, p\equiv 3 \pmod 4} (1 + p^{1-2s})^{-1} \prod_{p \nmid 2D,\, p \equiv 1 \pmod 4} (1 - \overline{(D/\pi)}_4 \pi p^{-s})^{-1}(1 - (D/\pi)_4\bar{\pi}p^{-s})^{-1} = \\ & = \prod_{(P)\in\mathcal{P}_2} (1 - \chi(P)N(P)^{-s})^{-1} \prod_{(P)\in\mathcal{P}_1} (1 - \chi(P)N(P)^{-s})^{-1} = \\ & = L_\chi(s) \end{aligned}$$

This is exactly what we wanted. It remains only to show that χ is an algebraic Hecke character.
As $\chi(P) = \overline{(D/\pi)_4}\pi$ for all ideals (P) with $P \nmid 2D$, for any ideal A which is coprime with $(2D)$ we have $\chi(A) = \overline{(D/\alpha)_4}\alpha$, where α is the unique generator of A such that $\alpha \equiv 1 \pmod{2+2i}$ (existence and uniqueness of such an α are checked as in Proposition 1.6).
Thus, to prove that χ is an algebraic Hecke character (of weight one) with respect to the modulus $(8D)$ it will be enough to check that

$$\alpha \equiv 1 \pmod{8D} \Rightarrow (D/\alpha)_4 = 1$$

This is the typical situation in which reciprocity laws can help us. Let us distinguish three cases:

1. $D \equiv 1 \pmod 4$ Then by Proposition 3.4 we have $(D/\alpha)_4 = (\alpha/D)_4 = 1$, since $\alpha \equiv 1 \pmod D$

2. $D \equiv 3 \pmod 4$ In this case, we have: $(D/\alpha)_4 = (i^2\alpha/D)_4 = 1$ as $(i/\alpha)_4 = 1$ for $\alpha \equiv 1 \pmod 8$ (easy computation). Hence, by Proposition 3.4, $(D/\alpha)_4 = (-D/\alpha)_4 = (\alpha/D)_4 = 1$.

3. D is even. Write $D = 2^t D_0$, with D_0 odd. Then by the previous points we have $(D_0/\alpha)_4 = 1$. It remains to show that $(2/\alpha)_4 = 1$.
   Since $\alpha = a + bi \equiv 1 \pmod{8D}$ and D is even ve have $\alpha \equiv 1 \pmod{16}$, hence $a - 1 \equiv 0 \pmod{16}$ and $b \equiv 0 \pmod{16}$. Lemma 3.7 implies that $((1+i)/\alpha)_4 = 1$. Thus:
$$1 = \left(\frac{1+i}{\alpha}\right)_4^2 = \left(\frac{2i}{\alpha}\right)_4 = \left(\frac{2}{\alpha}\right)_4$$

Hence we have proved the following

Theorem 5.5. *Let $D \in \mathbb{Z}, D \neq 0$. The L-function of the elliptic curve E of affine equation $y^2 = x^3 - Dx$ coincides with the Hecke L-function of the Hecke character modulo $M = (8D)$ defined on $\mathbb{Z}[i]$ by: $\chi(A) = \overline{(D/\alpha)_4}\alpha$ for $A = (\alpha)$ such that $(A, M) = 1$ and α is primary. Hence the L-function of E can be analytically continued to the whole complex plane.*

The L-function of $y^2 = x^3 + D$

We are now going to study the L-function of the elliptic curve E whose affine equation is $y^2 = x^3 + D$, $D \in \mathbb{Z}$, $D \neq 0$. The ideas involved are the same as in the previous paragraph. We will therefore describe the main steps without going in depth into the details.
Since in this case $\Delta = -2^4 3^3 D^2$ we have:

$$L_E(s) = \prod_{p \nmid 6D} (1 - a_p p^{-s} + p^{1-2s})^{-1}$$

where $a_p = p + 1 - N_p$ and N_p has been determined in section 5.0.2. By Theorem 5.1 we have: $a_p = 0$ if $p \nmid 6D$, $p \equiv 2 \pmod 3$; $a_p = -\overline{(D/\pi)_6}\pi - (D/\pi)_6\bar{\pi}$ if $p \equiv 1 \pmod 3$, $p \nmid 6D$ and $p = \pi\bar{\pi}, \pi \equiv 2 \pmod 3$. Hence:

$$L_E(s) = \prod_{p \nmid 6D, p \equiv 2 \pmod 3} (1 + p^{1-2s})^{-1} \prod_{p \nmid 6D, p \equiv 1 \pmod 3} (1 + \overline{(D/\pi)}_6 \pi p^{-s})^{-1} (1 + (D/\pi)_6 \bar{\pi} p^{-s})^{-1}$$

As in the previous paragraph, we will construct an algebraic Hecke character whose L-function coincides with the L-function of our curve; of course, in this case we will work with the ring $\mathbb{Z}[\omega]$.
We will construct an Hecke character χ of weight 1 on $\mathbb{Z}[\omega]$ with modulus $M = (12D)$. The definition is the following:

1. If $P \mid 6D$, $\chi(P) = 0$.
2. If $P \nmid 6D$ and $N(P) = p \equiv 1 \pmod 3$, let $(P) = (\pi)$, $\pi \equiv 2 \pmod 3, \pi\bar{\pi} = p$. Then $\chi(P) = -\overline{(4D/\pi)_6}\pi$.
3. If $P \nmid 6D$ and $N(P) = p^2$ then $p \equiv 2 \pmod 3$ and $(P) = (p)$. Then $\chi(P) = -p$.

As usual, we have used the symbol $(4D/\pi)_6$ to indicate $(4D)^{(N(\pi)-1)/6}$.
Notice that if p is an odd prime and $p \equiv 2 \pmod 3$ we have $(4D/p)_6 = (4D)^{(N(p)-1)/6} = ((4D)^{p-1})^{(p+1)/6} \equiv 1 \pmod p$, so $(4D/p)_6 = 1$.
Hence we can define, for (P) such that $(P, M) = 1$, $\chi(P) = -\overline{(4D/\pi)_6}\pi$ where π is such that $(P) = (\pi)$ and $\pi \equiv 2 \pmod 3$. As a consequence we have, for any ideal A such that $(A, M) = 1$, $\chi(A) = \overline{(4D/\alpha)_6}\alpha$, where α is the generator of A such that $\alpha \equiv 1 \pmod 3$.
Now, if $N(P) = p \equiv 1 \pmod 3$, $(P) = (\pi), \pi \equiv 2 \pmod 3$ we have $\chi(P) = -\overline{(4D/p)_6}\pi$.
Hence

$$(1 - \chi(P)N(P)^{-s})(1 - \chi(\bar{P})N(\bar{P})^{-s}) = (1 + \overline{(4D/p)_6}\pi p^{-s})(1 + (4D/p)_6 \bar{\pi} p^{-s})$$

If $N(P) = p^2$, $p \equiv 2 \pmod 3$ and $(P) = (p)$ then $\chi(P) = -p$. Hence

$$(1 - \chi(P)N(P)^{-s}) = 1 + p^{1-2s}$$

Therefore we obtain:

$$L_\chi(s) = \prod_{P \in Max(\mathbb{Z}[\omega])} (1 - \chi(P)N(P)^{-s})^{-1} = L_E(s)$$

It remains to show that χ is an algebraic Hecke character. As $\chi(A) = \overline{(4D/\alpha)_6}\alpha$ for any ideal $A = (\alpha)$ such that $(A, M) = 1$, it suffices to show that $\alpha \equiv 1 \pmod{12D} \Rightarrow (4D/\alpha)_6 = 1$. This can be shown using cubic reciprocity. Details can be found in [8].
Summing up, we obtain the following:

Theorem 5.6. *Let $D \in \mathbb{Z}$, $D \neq 0$. The L-function of the elliptic curve E of affine equation $y^2 = x^3 + D$ coincides with the Hecke L-function of the Hecke character modulo $M = (12D)$ defined on $\mathbb{Z}[\omega]$ by: $\chi(A) = \overline{(4D/\alpha)_6}\alpha$ for $A = (\alpha)$ such that $(A, M) = 1$ and $\alpha \equiv 1 \pmod 3$. Hence the L-function of E can be analytically continued to the whole complex plane.*

5.2.3 Why did things work?

Let us try to understand what we've been doing in the last two paragraphs: we were given an elliptic curve with affine equation $y^2 = x^3 + D$ or $y^2 = x^3 - Dx$. We were able to write down the L-function of such an elliptic curve explicitly, and we guessed an algebraic Hecke character defined on the ring of integers of a (reasonable) CM field whose L-function turned out to be exactly equal to the L-function of our elliptic curve.
There are (at least) two natural question which arise at this point:

1. Is there a peculiar property of the elliptic curves we considered that allows to express their L-function as the L-function of an algebraic Hecke character? (notice that, *a priori*, these two kinds of L-functions have nothing in common except the name).

2. Can this method be generalised to other elliptic curves, hence showing that their L-function can be analytically continued to $\mathbb{C}$?

Of course, these questions are closely related.
In this section, we will try to answer them. The key fact, as we shall see, is that elliptic curves of the form $y^2 = x^3 + D$ or $y^2 = x^3 - Dx$ have *complex multiplication*. We will explain what this means, and state the main results which connect the theory of algebraic Hecke characters with that of elliptic curves with complex multiplication. The advanced nature of these topics is far beyond the level of this exposition. Hence we will give no proofs; our aim is only to present concisely some fundamental results, leaving to the interested reader the opportunity to study more in depth these fascinating topics. Some useful material can be found in [7], [16].
Take an elliptic curve E defined over $\mathbb{Q}$, with affine equation: $y^2 = x^3 - Ax + B$, $A, B \in \mathbb{Z}$. Up to now, we've always studied E from the algebraic point of view. Anyway, we can

also look at $E(\mathbb{C})$, the set of complex projective points on E, which is a subset of $\mathbb{P}^2(\mathbb{C})$. This is easily seen to be a Riemann surface. It is of course very reasonable (and in fact it is true) that our algebraic definition of the genus of a curve coincides with the geometric definition of the genus of a Riemann surface (roughly speaking, its number of holes). Hence an elliptic curve is a Riemann surface of genus one, which is a torus.[5] It is well known that a torus can be realised as a quotient space of $\mathbb{C}$ with respect to a lattice L (a discrete subgroup of the plane not contained in a line). Up to an homothety, we can suppose that this lattice is generated by the points $(1, 0)$ and $\tau = (a, b)$ with $b > 0$.
Now consider the ring $\mathcal{O} = \{z \in \mathbb{C} : zL \subseteq L\}$. Clearly, $\mathcal{O} \supseteq \mathbb{Z}$.
It can happen that this inclusion is proper, which means, intuitively, that the lattice L has some nontrivial symmetries. If this is the case, we say that the elliptic curve $\mathbb{C}/L$ has complex multiplication. It is not hard to show that when this happens the ring $\mathcal{O}$ is a subring of the ring of algebraic integers of an imaginary quadratic field $\mathbb{Q}(\tau)$ which generates $\mathbb{Q}(\tau)$ over $\mathbb{Q}$ (such an $\mathcal{O}$ is called an *order* in $\mathbb{Q}(\tau)$).
The algebraic formulation of this concept is the following: to an elliptic curve E we can associate the corresponding *endomorphism ring*, containing those morphisms of the algebraic curve E to itself with also respect the group structure. It's easy to see that for each integer m group multiplication by m endomorphism is an endomorphism of E. These endomorphisms correspond to the trivial symmetries of the lattice in the complex plane. Algebraically speaking, the elliptic curve E has complex multiplication if its endomorphism ring is strictly bigger than $\mathbb{Z}$.
It turns out that having complex multiplication is the key property allowing to realise the L-function of a given elliptic curve as the L-function of an Hecke character:

Theorem 5.7. *(Deuring) Let E be an elliptic curve defined over $\mathbb{Q}$ having complex multiplication. Then there exists an Hecke character χ on a number field K such that:*

$$L_E(s) = L_\chi(s)$$

In particular, $L_E(s)$ has an analytic continuation to a holomorphic function on $\mathbb{C}$.

Proof. See [7] □

Let us return to our old friends, the elliptic curves of affine equation $y^2 = x^3 + D$ or $y^2 = x^3 - Dx$. The corresponding lattices in $\mathbb{C}$ are represented in Figure 5.1. Precisely, a curve of the form $y^2 = x^3 + D$ corresponds to a multiple of the lattice $\mathbb{Z}[\omega]$, while a curve of the form $y^2 = x^3 - Dx$ corresponds to a multiple of the lattice $\mathbb{Z}[i]$ (see [9]). In

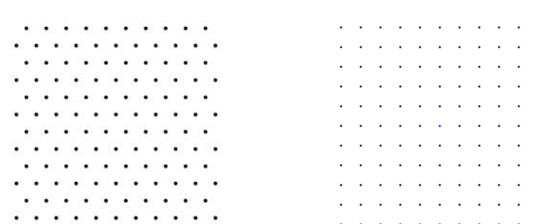


Figure 5.1

[5]This can also be shown directly with the aid of the Weierstrass $\wp$ function; see [9] and [20].

the first case, it's easy to see that $\{z \in \mathbb{C} : zL \subseteq L\} = \mathbb{Z}[\omega]$. In the second case, this set equals $\mathbb{Z}[i]$. Therefore our two families of elliptic curves have complex multiplication. That's the scientific reason why our computations were successful.
We eventually answered the questions at the beginning of the paragraph. Anything else?

Conclusion

Deuring's Theorem 5.7 allows to prove that the L-function of an elliptic curve defined over $\mathbb{Q}$ having complex multiplication can be analytically continued to the whole complex plane, by showing that this function can actually be realized as the Hecke L-function of a certain Hecke character. The converse of this result is also true: let E be an elliptic curve defined over $\mathbb{Q}$ such that there exists an Hecke character χ verifying $L_E(s) = L_\chi(s)$. Then E has complex multiplication. Unfortunately, most elliptic curves *do not* have complex multiplication. What can we say about the existence of an analytic continuation of the L-function of an *arbitrary* elliptic curve defined over $\mathbb{Q}$? Hasse and Weil proposed the following conjecture:

Conjecture 5.3. *(Hasse-Weil) The L-function of every elliptic curve defined over $\mathbb{Q}$ extends to an analytic function defined on the whole $\mathbb{C}$.*

A proof of this conjecture can be obtained as a consequence of another celebrated conjecture due to Taniyama, Shimura and Weil regarding the theory of *modular* elliptic curves.[6] The Taniyama-Shimura-Weil conjecture was first proved by Andrew Wiles in a special case (which was enough to deduce Fermat's last theorem) and then in its general form by Christophe Breuil, Brian Conrad, Fred Diamond, and Richard Taylor in 2001. Their result is now known as the *modularity theorem*, and is the tip of an enormous iceberg known as the Langlands program, a vast web of conjectures first proposed by Robert Langlands in 1967.[7] Very roughly speaking, these conjectures state that all the L-functions which arise in number theory (of which the L-functions of elliptic curves are a special case) can always be realised as L-functions of the so called *automorphic representations*. The problem is similar in nature to the one we faced in Section 5.2.2: in order to study the L-functions of certain number-theoretic objects (such as elliptic curves) we try to show that these L-functions can actually be obtained within a completely different context.

The amount of machinery needed just to state correctly Langlands's conjectures is very large, and most of these conjectures have not yet been proved. Much hard work still has to be done, but the promised rewards for such an effort are extraordinary. In this thesis we just explored the shore of that mysterious and fascinating sea which is modern number theory. We hope we encouraged the reader to sail towards further, unexplored waters.

[6] See [7], [9] and [12] for details.

[7] See [4] for a very good introduction to the Langlands program.

Bibliography

[1] Cassels G. W., Frohlich A., *Algebraic Number Theory*, London Mathematical Society, 2010

[2] Dummit D., Foote R., *Abstract Algebra*, Wiley, 2003

[3] Fulton W., *Algebraic Curves: An Introduction to Algebraic Geometry*, W. A. Benjamin, 1969

[4] Gelbart S., *An elementary introduction to the Langlands program*, Bulletin of the American Mathematical Society, 1984

[5] Hansen S. H., *Rational Points on Curves over Finite Fields*, Lecture Notes Series, 64. Aarhus Universitet, 1995

[6] Hartshorne R., *Algebraic Geometry*, Springer, 1997

[7] Husemoller D, *Elliptic Curves*, Springer, 2004

[8] Ireland K, Rosen M., *A Classical Introduction to Modern Number Theory*, Springer, 1998

[9] Kobliz N., *Introduction to Elliptic Curves and Modular Forms*, Springer, 1993

[10] Lang S., *Algebra*, Springer, 2005

[11] Lorenzini D., *An Invitation to Arithmetic Geometry*, American Mathematical Society, 1996

[12] Milne J. S., *Elliptic Curves*, BookSurge Publishing, 2006

[13] Milne J. S., *Fields and Galois Theory*, `http://www.jmilne.org/math/CourseNotes/ft.html`

[14] Miranda R., *Algebraic Curves and Riemann Surfaces*, American Mathematical Society, 1995

[15] Nekovar J., *Algebra 2*, Lecture Notes, ENS 2013-2014 `http://www.math.jussieu.fr/~nekovar/co/ens/`

[16] Neukirch J., *Algebraic Number Theory*, Springer, 1999

[17] Serre J-P., *A Course in Arithmetic*, Springer, 1996

[18] Silverman J. N., *The Arithmetic of Elliptic Curves*, Springer, 2009

[19] Silverman J. N., Tate J., *Rational Points on Elliptic Curves*, Springer, 1994

[20] Stein E. M., *Complex Analysis*, Princeton University Press, 2001