On the Use of Elliptic Curves in Public-Key Cryptography

Christoph Strobl

2020
September

Abstract

In this thesis, an introduction to public key cryptography over finite fields and elliptic curves is given. Elliptic curves are introduced using affine and projective spaces. The thesis also gives an introduction to algorithms that are able to break the discrete logarithm problem over finite fields and elliptic curves faster than an exhaustive search.

Uppsatsen behandlar asymmetrisk kryptering över ändliga kropper och elliptiska kurver. Elliptiska kurver införs med hjälp av affint och projektivt rum. Uppsaten också ger en introduction till algoritmer som kan knäcka diskret logaritm problemet över kropper och elliptiska kurver.

Keywords: elliptic curves, public key cryptography, index calculus, Pohlig-Hellman, Diffie-Hellman Problem, asymmetric cryptography

Danksagung

Ich danke meinem Betreuer Oscar Marmon zuallererst für den Vorschlag des Themas und dann für seine Betreuung während dieser Aufgabe. Es war nicht die einfachste Zeit und ich wohl auch nicht der einfachste Student.

Ich danke meinem Urgroßvater und meinem Großvater. Beide haben mir mein Erststudium in Deutschland und mein jetziges Zweitstudium in Schweden ermöglicht. Ich hatte das große Glück beide zu treffen und Erinnerungen an sie in meinem Gedächtnis vorhalten zu können.

Auch wenn ich durch das Studium der Mathematik viel an Erkentniss gewann, so vermag ich mit jedem Schreiten durch eine Bibliothek zu verspüren wie wenig es doch ist und wie viel es noch zu entdecken gibt. Ich hoffe mein Durst nach Erkenntnis und Wahrheit wird niemals gestillt werden und ich kann das Erlernte nutzen ohne Anderen zu schaden.

Contents

Notation

DHP	Diffie-Hellman Problem.
DLP	Discrete Logarithm Problem.
ECDHP	Elliptic Curve Diffie–Hellman Problem.
ECDLP	Elliptic Curve Discrete Logarithm Problem.
$\#E(\mathbf{F}_p)$	Number of rational points on the elliptic curve E over the field $\mathbf{F}_p$.
$E(\mathbf{K})$	Elliptic curve E over the field $\mathbf{K}$.
$\mathbf{F}_p^*$	Nonzero elements in the field $\mathbf{F}_p$.
$\phi(n)$	Euler Phi-Function – number of integers up to n which are coprime to n.

1 Introduction

A secret is a piece of information that someone wishes to keep entirely to themselves, or to be able to share with a limited group of people whom they trust. Furthering this idea is the notion of how long somone would like to keep this piece of information secret for. In the short term we can look at the example of buying a present for our partner, a piece of information that can be shared discretely with a number of individuals over a short period of time until it no longer needs to be kept secret. In the long term it could also be a piece of information that remains hidden for many years or even beyond death.

When keeping information secret for a long time, one has to take extended precautions. With rising computation power and available storage, this can become a real problem. For asymmetric cryptography large prime numbers or compositions of large prime numbers are needed. These primes are used as keys in asymmetric cryptography schemes and factoring the composition of two large primes breaks the system. A prime of 2048 bits is expected to be able to secure data until approximately 2022, 3072 bits until 2038, and 4096 bits until 2050 [FSK10]. But on the other hand, the rise in computation power, storage and transmission speed is not only on the attackers side. So where lies the problem? Take the German passport for example, access to biometric photos and fingerprints are managed by asymmetric encryption [Ben+08] and the transmission is done using contactless radio transmission. The problem when the key size is increasing lies in the rate of radio transmission. With increasing key size transmissions will take longer and longer in the future. Therefore switching to a different scheme, considered equally secure, with a smaller key size is advisable as not only the key size but also the size of the stored data is expected to increase in the future. Elliptic curve cryptography is considered to be as secure as choosing larger prime numbers as encryption keys with the added benefit of maintaining smaller key sizes.

History

Neal I. Koblitz and Victor S. Miller are the fathers of elliptic curve cryptography [Kob87] [Mil86a]. The first application of elliptic curves concerning cryptography was done by Hendrik Lenstra in 1984. He developed a method using elliptic curves to factor large integers into their prime factors [Len86].

In late 1984, Lenstra sent a copy of his algorithm to Koblitz shortly before Koblitz left for a study trip to the Soviet Union. During his trip Koblitz got the idea on using elliptic curve groups to construct a cryptosystem. Although fluent in Russian, he knew nobody to discuss his ideas with because cryptographic research was not done openly at universities in the Soviet Union at the time. Instead Koblitz wrote a letter to Andrew Odlyzko at Bell Labs, describing his ideas. The letters took a couple of weeks each way and so it wasn't until a month later when he got positive feedback from Odlyzko with the hint that Victor Miller was also working on this topic at the same time at IBM. Koblitz, during that time, had no notion of commercializing his idea and Miller was discouraged by the bureaucracy at IBM. Both ended up not filing patents for their research [Kob08].

The aim of this work is to introduce asymmetric cryptography using finite fields and elliptic curves over finite fields. The reader is expected to have taken a course on abstract algebra. To introduce elliptic curve cryptography from an mathematical standpoint, the discrete logarithm is introduced first. Building on the idea of the discrete logarithm, the Diffie-Hellman cryptosystem over finite fields and its challenges are introduced in the next section. The work continues with the introduction of elliptic curves from affine and projective curves. In the final section Diffie-Hellman over elliptic curves is presented, as well as an algorithm to break a wrong set up system.

The application of mathematical ideas to real world problems depends on numerous theorems as well as a broad toolkit of techniques, covering the entirety of information relating to elliptic curve cryptography is beyond the scope of this thesis. Therefore, only the theorems, lemmas and propositions of the utmost importance are presented here with proofs. It has been an arduous task determining what is most important to this thesis.

2 The Discrete Logarithm

The discrete logarithm (DL) can be viewed from an algebraic perspective or a number theoretic one. In Number Theory the DL is referred to as the index [Bur11], and algebraically we refer to the following definition [HPS14].

Definition 2.1. Let $\mathbf{G}$ be a group with $\star$ as its group operation. Then x is called a solution to the discrete logarithm if:

$$\underbrace{g \star g \star \cdots \star g}_{x \text{ times}} = h.$$

The following definition is the specific case when the group operation is multiplication.

Definition 2.2. Let g and h be elements of the multiplicative group $\mathbf{F}_p^*$, both non zero and g a primitive root in the group. Then x is called a solution to the discrete logarithm if

$$\begin{aligned} g^x &\equiv h \mod p \\ x &= \log_g(h). \end{aligned}$$

While the multiplicative version of this definition is used in cryptography, the additive counterpart is unusable for this purpose. The reason for this is that the DL can be calculated by using the Euclidean algorithm for finding the inverse to the primitive root in $\mathbf{F}_p$ [Wer13].

The next step is to look at basic properties of the multiplicative DL.

Theorem 2.3. *Let $a, b, g \in \mathbf{F}_p^*$, with g a primitive root. Then the following rules hold:*

1. $\log_g(ab) \equiv \log_g(a) + \log_g(b) \pmod{\phi(p)}$
2. $\log_g(a^k) \equiv k \log(a) \pmod{\phi(p)}$
3. $\log_g(1) \equiv 0 \pmod{\phi(p)}$

Proof. From the definition of the discrete logarithm, it follows that $g^{\log_g(a)} \equiv a \pmod{m}$ and $g^{\log_g(b)} \equiv b \pmod{p}$. When multiplying both congruences,

$$g^{\log_g(a)+\log_g(b)} \equiv ab \pmod{p}$$

is the result. But by definition of the exponential function: $g^{\log_g(ab)} \equiv ab \pmod{p}$. This leads to the conclusion that:

$$g^{\log_g(a)+\log_g(b)} \equiv g^{\log_g(ab)} \pmod{p}.$$

In case $\log_g(a) + \log_g(b)$ is bigger than $\phi(p)$, it still holds, because $a^i \equiv a^j \pmod{p}$ if and only if $i \equiv j \pmod{\phi(p)}$

Rule 2 can be proven by using the definition of the discrete logarithm again:

$$g^{\log_g(a^k)} \equiv a^k \pmod{p}.$$

Applying exponentiation results in:

$$(g^{\log_g(a)})^k = g^{k \log_g(a)} \equiv a^k \pmod{p}.$$

Which shows:

$$g^{\log_g(a^k)} = g^{k \log_g(a)} \pmod{p}.$$

From this it can be deduced that $\log_g(a^k) \equiv k \log(a) \pmod{\phi(p)}$.

As an immediate consequence of the definition of exponents $a^1 = a$, therefore, dividing both by a leaves $a^0 = 1$ and rule 3 holds. □

3 Diffie-Hellman Key Exchange

A classical problem in cryptography is the exchange of a shared secret over an insecure channel. With the Diffie-Hellman key exchange, the problem can be avoided by creating a shared secret over an insecure channel with two or more participants. Following to classical cryptography notation, Alice and Bob are two parties who wish to communicate secretly, while Eve would like to intercept their communication and steal their secrets. In the beginning, Alice and Bob agree on a prime p and a primitive root g in $\mathbf{F}_p^*$ using the insecure channel. Eve will take notice of that and therefore also know which g they will use. As a next step, Alice chooses an secret integer a. Bob does likewise choose an integer b and keeps it secret. In the next step, Alice and Bob will use their secret integers to compute two values A and B

$$A \equiv g^a \pmod p$$

and

$$B \equiv g^b \pmod p.$$

If Alice and Bob now exchange the values A and B, Eve will also record those two values. Alice and Bob now raise the received number to the power of their secret integer. This then gives:

$$B' = B^a \equiv g^{ab} \pmod p$$

and

$$A' = A^b \equiv g^{ab} \pmod p.$$

Both have now a shared secret which they created together over an insecure channel. Their shared secret is never transmitted, both keep it to themselves. For more than two participants the algorithm stays essentially the same but involves a more complicated exchange of intermediate calculation results. Eve ends up with A, B and g and is not able to reconstruct a or b in a simple way from the observed values, which she needs to get g^{ab}. The interesting question is now: How could Eve reconstruct a or b?
This is called the Diffie-Hellman problem.

Definition 3.1 (Diffie-Hellman-Problem)**.** Let p be a large prime and g a primitive root in $\mathbf{F}_p^*$. Then the Diffie-Hellman Problem (DHP) is the problem of constructing the value $g^{ab} \pmod p$ from the values $g^a \pmod p$ and $g^b \pmod p$.

The DHP is not harder to solve than the DLP introduced in section 2. When the attacker can solve the DLP, the secret exponents of Bob and Alice can be obtained which makes it easy to calculate their shared secret g^{ab}.

3.1 Implementation Issues

There are several possible issues that can decrease the difficulty of solving the DHP and therefore weaken the encryption. Besides programming issues, there are also mathematical traps. Kohno, Ferguson and Schneier [FSK10] give insight into these traps. A first mistake that can be done is choosing an arbitrary prime p for $\mathbf{F}_p^*$. The next theorem states the reason for this.

Theorem 3.2. *When a cyclic group G has order n, then G has a unique subgroup of order k if k is a divisor of n.*

Proof. This is a well-known theorem and a proof can be found in most books on abstract algebra. □

Choose a random element g^* from $\mathbf{F}_p^*$ to be the generator of the group. g^* then generates a subgroup which divides to order of $\mathbf{F}_p^*$. The order of $\mathbf{F}_p^*$ is $p-1$ and therefore an even number. It can be the case that for a bad choice of a generator, g^* only generates a small subgroup $\mathbf{F}_p^*$. This can enable, depending on the subgroup's size, Eve to try all possible values in a reasonable amount of time. To avoid this problem, one can use a special sort of primes:

Definition 3.3 (Safe prime)**.** A prime number of the form $p = 2q + 1$ where q is also prime is called a safe prime or Sophie Germain prime.

By Theorem 3.2, if p is a safe prime, the number of multiplicative subgroups of $\mathbf{F}_p^*$ reduces to four, namely:

1. The trivial subgroup containing only the neutral element 1.
2. A subgroup of size two, containing 1 and $p - 1$.
3. A subgroup of size q.
4. The full group of size $2q$ itself.

The preferred subgroup to use is the one of size q. This can be seen from mathematical properties of the subgroups. While it is easy to rule out the use of the first two subgroups by trial and error, a more advanced idea has to be used to distinguish between the groups of size q and $2q$.

Definition 3.4 (Legendre symbol)**.** Let p be a prime bigger than 2, and $\gcd(a, p) = 1$, if there exists an a such that $a \equiv x^2 \pmod{p}$, then a is said to be a quadratic residue or nonresidue if that is not the case. The Legendre symbol is defined in the following way:

$$\left(\frac{a}{p}\right) = \begin{cases} 1 & \text{if } a \text{ is a quadratic residue modulo } p \text{ and } a \not\equiv 0 \pmod{p}, \\ -1 & \text{if } a \text{ is a quadratic nonresidues modulo } p, \\ 0 & \text{if } a \equiv 0 \pmod{p}. \end{cases}$$

Lemma 3.5. *The subgroup of order q of the group $\mathbf{F}_p^*$ with $p = 2q + 1$ consists only of quadratic residues* $\pmod p$.

Proof. The quadratic residues are congruent (in some order) to the even powers of the groups generator $g^{2k} \pmod p$.

Closure

If $a = g^{2l}$ and $b = g^{2m}$ are quadratic residues $\pmod p$, so is their product $ab = g^{2(k+l)}$ also a quadratic residue $\pmod p$.

Associativity

For $a = g^{2k}, b = g^{2l}$ and $c = g^{2m}$ it holds that $(a \cdot b) \cdot c = (g^{2k} g^{2l}) \cdot g^{2m} = g^{2(k+l)} \cdot g^{2m} = g^{2(k+l+m)} = g^{2k} \cdot g^{2(l+m)} = g^{2k} \cdot (g^{2l} \cdot g^{2m}) = a \cdot (b \cdot c)$. So the quadratic residues are associative.

Identity element

The element $g^0 = 1$ is the identity element in the group.

Inverse element

For an arbitrary element g^{2n} there exists and element g^{-2n} such that $g^{2n} \cdot g^{-2n} = g^0 = 1$. Where $-2n \equiv m \pmod{\phi(p)}$ for some $m \in 1, ..., \phi(p)$.

The quadratic residues obey the group axioms, to finish the proof it has to be shown that their group order is in fact p.

From Number Theory, it is known that a number can either be a quadratic residue or nonresidue, not both. It is also well known that exactly half of the numbers between 1 and $p - 1$ are nonresidues $\pmod p$. Therefore the subgroup of order q is formed by the quadratic residues $\pmod p$. □

The subgroup of order $2q$ is not used because the use of elements which are quadratic residues and nonresidues reveals information about the triple (g^x, g^y, g^z) where it can help to determine if $g^x \cdot g^y = g^z$ or not. This problem is related to the *Decisional Diffie–Hellman assumption* [FSK10].

3.2 Index Calculus

Besides being aware of issues that weaken the security of a cryptographic system, one should also take into account algorithms which are designed to exploit the nature of the problem effectively. The Baby-Step-Giant-Step or the Pohlig-Hellman algorithm can solve the DLP but run in *exponential* time.

A very powerful method for the DLP is the so-called *index calculus* algorithm. The first ideas for it came from Wester and Miller and got published 1968. In the late 1970s and early 1980s Adelman, Merkle and Pollard invented an algorithm independently from each other. That algorithm can give a solution in subexponential time but only works on $\mathbf{F}_p$ [HPS14][Ngu11].

The main idea of index calculus is that one builds a basis of known logarithm values and tries to construct the unknown value from this basis. A similar approach is also used when factoring composite numbers of big primes with the number field or quadratic sieve.[1]

Smooth Numbers

Definition 3.6. A number is called B-smooth if it can be factored into prime factors less than or equal to B.

The first step is to define which value B to choose for the so-called factor base. Choosing B is a trade-off between efficiency of the algorithm and the likelihood to find numbers that are B-smooth.

Definition 3.7. Let $\psi(x, B)$ denote the function that counts the numbers smaller or equal to x which are B-smooth.

In order to investigate the complexity of the index calculus method, it is necessary to understand how smooth numbers are distributed. In other terms, how many B-smooth numbers are there in the interval from 1 to x.

Silverman and Hoffstein [HPS14] give a result by Canfield, Erdős and Pomerance:

Theorem 3.8. *For a fixed $0 < \epsilon < \frac{1}{2}$, let x and B increase such that the following inequality is always satisfied:*

$$\ln(x)^{\epsilon} < \ln(B) < \ln(x)^{1-\epsilon}.$$

Then it holds that:

$$\psi(x, B) = x \cdot u^{-u(1+o(1))}$$

with $u = \frac{\ln(x)}{\ln(B)}$.

Proof. For a proof of this theorem and further information about smooth numbers, the reader is advised to consult [Gra08] and [Pom08]. □

Time Complexity

The time complexity of an algorithm or a function is the approximated cost of computations it would take to solve the task. This cost can be measured in group operations or bit operations. The difference is that when bit operations are done, the operation costs also depends on the length of the input, while group operations are independent of input size.

[1] See [HPS14],[FSK10] and especially [Pom08]

Operation	Expected complexity for cryptography
Addition of and m and an n-bit-integer	$\mathcal{O}(\max\{\log n, \log m\})$
Multiplication of m-bit integers – $M(m)$	$\mathcal{O}(m^2)$
Multiplication $\pmod n$	$\mathcal{O}(M(\log(n)))$
Inversion $\pmod n$	$\mathcal{O}(\log^2(n))$
Computation $g^m \pmod n$	$\mathcal{O}(\log(m)M(\log(n)))$

Table 1: Taken from [Gal12]

Definition 3.9 (Big-Oh-Notation)**.** In this work $f(n) \in \mathcal{O}(g(n))$ will denote the following:

If there exist constants c and N such that for all n $\geqslant$ N it holds that $|f(n)| \leqslant c|g(n)|$, then this can be written as:

$$\limsup_{n\to\infty} \frac{|f(n)|}{|g(n)|} < \infty.$$

Definition 3.10 (Little-Oh-Notation)**.** In this work $f(n) \in o(g(n))$ will denote the following:

$$\limsup_{n\to\infty} \frac{|f(n)|}{|g(n)|} = 0.$$

The cost of operations usually varies, depending on the implement algorithm. For example the cost of integer multiplication can vary between $\mathcal{O}(n^2)$ for classical *schoolbook* multiplication to as low as $\mathcal{O}(n \log n \log\log n)$ for a multiplication using a *Fast Fourier Transformation* [Sut19]. Therefore the abbreviation $M(m)$ will be used when multiplication is part of an algorithm.

If needed the cost stated in table 1 will be assumed for the different operations when time complexity is derived in this work.

Another notation that is commonly used when analyzing algorithms is the so-called L-Notation.

Definition 3.11. Let the L-notation be denoted as follows:

$$L_n[\alpha, c] = e^{(c+o(1))(\ln n)^{\alpha}(\ln\ln n)^{1-\alpha}}$$

with c a positive constant and $0 \leqslant \alpha \leqslant 1$.

Depending on the value of α the complexity described in this notation is referred to in three different terms. For the case that $\alpha = 0$ the algorithm is said of being of polynomial time. If $\alpha = 1$ it is of exponential and for $0 < \alpha < 1$ of sub-exponential time.

Algorithm 3.12 (Index Calculus)**.** *Let α denote a primitive root in $\mathbf{F}_p^*$ and β the element whose discrete logarithm is to be computed. The problem to solve is then to find the value of $log_\alpha(\beta)$.*

1. *Decide on a smoothness bound B and and compute the factor base $F_B := \{p_1, p_2, ..., p_b\}$ where b denotes the total of number that are B-smooth.*

2. *Generate a random power of α^r and attempt to factor α^r over the factor base. If successful save the relation $r_i = \alpha^{e_i}\beta^{-1} = \prod p_j^{ei,j}$. Repeat this step until $b+1$ relations are found.*

3. *The resulting equations from step 2 form a system of linear equations when the logarithm with respect to α is taken on both sides. Solve this system of equations.*

4. *Return the solution $log_\alpha(\beta)$ if found, otherwise go to step 2 again.*

The following example is taken from [HPS14]:

Example 1. Let $p = 18443$, 37 be a primitive root $\pmod p$ and the discrete logarithm problem to be solved is

$$37^x \equiv 211 \pmod{18443}$$

Setting $B = 5$, the factor base consists of the elements 2,3 and 5. The first step is to compute 37^x for random values of x and check if the result is 5-smooth. If so, the value and its factorization are saved for the next step of the algorithm. For example:

$$g^{12708} \equiv 2^3 \cdot 3^4 \cdot 5 \pmod{18443}, \qquad g^{11311} \equiv 2^3 \cdot 5^2 \pmod{18443}$$
$$g^{15400} \equiv 2^3 \cdot 3^3 \cdot 5 \pmod{18443}, \qquad g^{2731} \equiv 2^3 \cdot 3 \cdot 5^4 \pmod{18443}$$

This can be turned into a linear equation with three unknowns:

$$12708 = 3 \cdot \log_{37}(2) + 4 \cdot \log_{37}(3) + \log_{37}(5)$$

Doing the same for all the before given examples gives rise to a system of linear equations:

$$\begin{aligned}
3x_2 + 4x_3 + \ \ x_5 &= 12708 \pmod{18442} \\
3x_2 + \qquad + 2x_5 &= 11311 \pmod{18442} \\
3x_2 + 3x_3 + \ \ x_5 &= 15400 \pmod{18442} \\
3x_2 + \ \ x_3 + 4x_5 &= 2731 \pmod{18442}
\end{aligned}$$

This gives the two solutions

$$(x_2, x_3, x_5) \equiv (1, 0, 1) \pmod 2$$
$$(x_2, x_3, x_5) \equiv (5733, 6529, 6277) \pmod{9221}$$

Combining those gives:

$$(x_2, x_3, x_5) \equiv (5733, 15750, 6277) \pmod{18442}$$

The next step is computing $211 \cdot 37^{-k} \pmod{18443}$ for random values of k until a B-smooth value is obtained.

$$211 \cdot 37^{-9549} \equiv 2^5 \cdot 3^2 \cdot 5^2 \pmod{18443}$$

Taking the previously calculated values for the discrete logarithms of 2,3 and 5 leads to:

$$\begin{aligned} \log_{37}(211) &= 9549 + 5\log_{37}(2) + 2\log_{37}(3) + 2\log_{37}(5) \\ &= 9549 + 5 \cdot 5733 + 2 \cdot 15750 + 2 \cdot 6277 \equiv 8500 \pmod{18442} \end{aligned}$$

And in fact, $37^{8500} \equiv 211 \pmod{18442}$.

The proof of the next theorem follows closely the argumentation of [Sut19]. The simplest approach for factoring over the factor base, trial-division, is used. There are faster algorithms for factoring, it does not change the fact that the algorithm is subexponential but simplifies the proof.

Theorem 3.13. *The index calculus algorithm for solving the discrete logarithm problem over finite fields is of subexponential complexity.*

Proof. At first, the second step of the algorithm will be examined, the reason for this will become clearer later. The second step takes approximately:

$$(b + 1) \cdot u^u \cdot b \cdot \mathrm{M}(\log N)$$

where $u = \frac{\log N}{\log B}$.

- $b + 1$: the number of equation needed for the linear algebra step
- u^u: the number of random exponents expected to try to obtain an B-smooth integer m within $[1, N]$
- b: number of trial divisions to test if the number m is B-smooth and then factor it
- $\mathrm{M}(\log N)$: the time for each trial division

The first assumption to do is that $b \approx b + 1$. So the equation simplifies to:

$$b^2 \cdot u^u \cdot \mathrm{M}(\log N).$$

The number of primes $b = \pi(B)$ up to B, can roughly be approximated with $\pi(B) \approx \frac{B}{\log B}$ by the *prime number theorem*, leading to:

$$\left(\frac{B}{\log B}\right)^2 \cdot u^u \cdot \mathrm{M}(\log N)$$

In the end, $\frac{B}{\log B}$ is a large number, replacing it with just B will help to analyse the algorithm better.

$$B^2 \cdot u^u \cdot \mathrm{M}(\log N)$$

The factor $\mathrm{M}(\log N)$ will be crossed out. The reason will be discussed near the end of the proof. The fact that $u = \frac{\log N}{\log B}$ makes it possible to rewrite B in terms of N such that:

$$B^2 u^u = N^{2/u} u^u.$$

Taking the logarithm on both sides leads to:

$$f(u) = \log(N^{2/u} u^u) = \frac{2}{u}\log(N) + u \log u.$$

Examining the derivative $f'(u)$ helps to minimize the function:

$$f'(u) = -\frac{2}{u^2}\log N + \frac{2}{uN} + \log u + 1 = 0.$$

Neglecting the terms 1 because it is not relevant enough for asymptotic behaviour when N is big, one simplifies the derivative to:

$$\tilde{f}'(u) = -\frac{2}{u^2}\log N + \log u + \mathcal{O}(1) = 0.$$

Which leads to the following approximation:

$$\begin{aligned} u^2 \log u &\approx 2\log N \\ u^2 &\approx \frac{2\log N}{\log(\log N - \log B)} \end{aligned}$$

noticing that $\log B \in o(\log N)$ simplifies then to:

$$u^2 \approx \frac{2\log N}{\log\log N}.$$

Using

$$u = 2\sqrt{\log N / \log\log N}$$

in the equation gives:

$$u^2 \log u = \frac{4 \log N}{\log \log N} \cdot \left(\log 2 + \frac{1}{2}(\log \log N - \log \log \log N) \right) = 2 \log N + o(\log N).$$

This value of u shows that one should use the following smoothness bound:

$$\begin{aligned} B &= N^{1/u} = \exp(\frac{1}{U} \log N) \\ &= \exp(\frac{1}{2}\sqrt{\log N \log \log N}) \\ &= L[\frac{1}{2}, \frac{1}{2}]. \end{aligned}$$

This also gives $u^u = \exp(u \log u)$ which equals $L_N = L[\frac{1}{2}, 1]$ in L-Notation. The factor $M(\log N)$ can be ignored because multiplying by a polynomial in $\log N$ does not change the asymptotic level of precision of this time complexity analysis. Putting those results together gives the expected running time of step two:

$$B^2 u^u = L_N[1/2, 1/2]^2 L_N[1/2, 1] = L_N[1/2, 2].$$

The step of solving the linear algebra part of the problem depends on the size of the smoothness bound only. Using Gaussian elimination to solve the system of equations, this step can be bounded by $\mathcal{O}(b^3)$. b was before approximated by B, so this leads to $\mathcal{O}(B^3)$. This translates to $L_N[1/2, 3/2]$ and, as it is added, is dominated by the previous step of generating the linear relations. □

4 Elliptic Curves

The main goal of this section is to define elliptic curves formally and investigate the group structure they contain. To achieve this, affine curves and projective curves are introduced. This is necessary to formally introduce the *point at infinity $\mathcal{O}$* used as the neutral group element in cryptography. For this formal introduction Werner's book [Wer13] is closely followed, but some parts are also taken from [Eng12]. The rest about elliptic curves cryptography is based on the literature [ST15], [Bla+99]. In contrast to [Lan78] who writes:

> It is possible to write endlessly on elliptic curves. (This is not a threat.)

this section will have a finite number of pages.

4.1 Affine Curves

Definition 4.1. Let f be a polynomial function of two variables with coefficients in some field $\mathbf{F}$:

$$f(x,y) = \sum_{\mu_1,\mu_2 \geqslant 0} \alpha_{\mu_1,\mu_2} x^{\mu_1} y^{\mu_2} \qquad \alpha_{\mu_1,\mu_2} \in \mathbf{F}$$

with only finitely many α_{μ_1,μ_2} nonzero and $f \neq 0$. The set of zeros of f in $\mathbf{F} \times \mathbf{F}$ shall be then called $C_f(\mathbf{F})$:

$$C_f(\mathbf{F}) = \{(a,b) \in \mathbf{F} \times \mathbf{F} \mid f(a,b) = 0\}.$$

This will be called an affine plane curve. The notation can be simplified by just writing $C(\mathbf{F})$ when it is clear which polynomial function f is meant.

Definition 4.2. The space $\mathbf{F} \times \mathbf{F}$ may also be denoted as $\mathbf{A}^2(\mathbf{F})$.

$$\mathbf{A}^2(\mathbf{F}) = \{(a,b) \mid a,b \in \mathbf{F}\}$$

and called the *two dimensional affine space.*

Example 2. Let $\mathbf{F} = \mathbf{F}_5$ and $f(x,y) = y^2 - x^3 - x - 1$. Then the curve $C_f(\mathbf{F})$ consists of the set of solutions of the equation $y^2 = x^3 + x + 1$. Checking which elements of $\mathbf{F}_5$ satisfy the equation, gives the points:

$$C_f(\mathbf{F}_5) = \{(0,1),(0,4),(2,1),(2,4),(3,1),(3,4),(4,2),(4,3)\}$$

One can also consider a field $\mathbf{E}$ which contains $\mathbf{F}$. A special case would be $\bar{\mathbf{F}}$, its closure and see that $C_f(\mathbf{F})$ is a sub set of $C_f(\bar{\mathbf{F}})$.

Definition 4.3. An affine curve $C_f(\mathbf{F})$ is said to be *singular* in a point $(a, b) \in C_f(\mathbf{F})$ if f itself and both partial derivatives equal to zero in (a, b).

$$f(a,b) = \frac{\partial f}{\partial x}(a,b) = \frac{\partial f}{\partial y}(a,b) = 0$$

Definition 4.4. An affine curve $C_f(\mathbf{F})$ is said to be *non-singular* if the curve $C_f(\bar{\mathbf{F}})$ isn't singular in any point $(a, b) \in \mathbf{A}^2(\bar{\mathbf{F}})$.

So it may be that $C_f(\mathbf{F})$ is singular although it does not contain a singular point at all. An example can be given when one considers a curve of the reals $\mathbf{R}$ and its closure $\mathbf{C}$, the complex numbers. Werner [Wer13] gives the following example:

Example 3. Let $f(x, y) = y^2 - x^4 - 2x^2 - 1$ and then its partial derivatives are

$$\frac{\partial f}{\partial x} = -4x(x^2 + 1) \quad \text{and} \quad \frac{\partial f}{\partial y} = 2y$$

Then $f, \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}$ have no zeros over $\mathbf{R}$ in common. But over the closure of $\mathbf{R}$ there exist the two points $(i, 0)$ and $(-i, 0)$ which are in $C_f(\mathbf{C})$ so that the curve $C_f(\mathbf{R})$ is singular.

4.2 Projective Curves

To define the point $\mathcal{O}$, it is necessary to examine the curves in a space other than the affine space.

Considering the same curve $C_f(\mathbf{F})$ as before from example 2:

$$f(x, y) = y^2 - x^3 - x - 1$$

The solutions $(a, b) \in \mathbf{A}^2(\mathbf{F})$ follow the rule $b^2 = a^3 + a + 1$. Choosing an arbitrary number $c \neq 0 \in \mathbf{F}$, one can define $a' = ac$ and $b' = bc$. When then replacing a and b by a', b' the equation rewrites to:

$$\left(\frac{b'}{c}\right)^2 = \left(\frac{a'}{c}\right)^3 + \frac{a'}{c} + 1.$$

Multiplying with c^3 gives the equation $b'^2c = a'^2 + a'c^2 + c^3$ which makes $(a', b', c) \in \mathbf{F} \times \mathbf{F} \times \mathbf{F}$ a solution of an equation in three variables:

$$Y^2Z = X^3 + XZ^2 + Z^3.$$

The reason for doing this is that the extended equation has more solutions than the previous one in only two variables. Assuming that $(a, b, c) \in \mathbf{F} \times \mathbf{F} \times \mathbf{F}$, the solutions are of the form:

$$b^2c = a^3 + ac^2 + c^3.$$

This leads to two different cases:

First, the c which extends the equation is chosen to be zero, which leads to the case that $a^3 = 0$, which indicates that $a = 0$ and b can be chosen arbitrarily. This is not a solution to the affine form of the equation. This case will play an important role later and should not be forgotten, but will be put aside for a bit now.

The second case occurs if c is not equal to zero. Then one can divide the equation by c^3 and get the solution $\left(\frac{a}{c}, \frac{b}{c}\right)$.

This shows that if (a, b, c) is a solution to the extended equation, also (ta, tb, tc) is a solution for it when $t \neq 0$. For the case that $c \neq 0$ the results are similar. Let $tc \neq 0$, then $\left(\frac{a}{c}, \frac{b}{c}\right) = \left(\frac{ta}{tc}, \frac{tb}{tc}\right)$.

Definition 4.5. The points (a, b, c) and (a', b', c') are called equivalent in $\mathbf{F} \times \mathbf{F} \times \mathbf{F}$ if there exist $t \in \mathbf{F}\backslash\{0\}$ such that:

$$a = ta', \qquad b = tb', \qquad c = tc'.$$

The equivalence is denoted by the symbol $\sim$, written $(a, b, c) \sim (a', b', c')$.

Definition 4.6. The two dimensional projective space $\mathbf{P}^2(\mathbf{F})$ is defined as the quotient of $\mathbf{F} \times \mathbf{F} \times \mathbf{F} \backslash\{(0, 0, 0)\}$ with the equivalence relation $\sim$:

$$\mathbf{P}^2(\mathbf{F}) = \mathbf{F} \times \mathbf{F} \times \mathbf{F} \backslash\{(0, 0, 0)\}/\sim$$

So the projective space $\mathbf{P}^2(\mathbf{F})$ is a set of equivalence classes of $\sim$. Which means that every $(a, b, c) \neq (0, 0, 0)$ is a point in $\mathbf{P}^2(\mathbf{F})$, denoted $[a : b : c]$. Two points $[a : b : c]$ and $[a' : b' : c']$ are called equivalent if and only if it holds that $a = ta', b = tb'$ and $c = tc'$ for some $t \neq 0$.

It is possible to formulate a mapping between the affine space and the projective space.

$$\begin{aligned} i : \mathbf{A}^2(\mathbf{F}) &\to \mathbf{P}^2(\mathbf{F}) \\ i(a, b) &= [a : b : 1] \end{aligned}$$

Using this definition, one can see that the mapping is injective:

$$\begin{aligned} &i(a, b) = i(a', b') \\ \Rightarrow &[a : b : 1] = [a' : b' : 1]. \end{aligned}$$

From this it can be deduced that $t = 1$, because the equations $a = ta', b = tb'$ and $1 = 1t$ have to hold. This shows that $(a, b) = (a', b')$. When using the mapping i, $\mathbf{A}^2(\mathbf{F})$ can be viewed as a subset of $\mathbf{P}^2(\mathbf{F})$.

All points in the form $[a : b : c]$ with $c \neq 0$ can be written as points in $\mathbf{A}^2(\mathbf{F})$. But $\mathbf{P}^2(\mathbf{F})$ is bigger than that. For example the point $[a : b : 0]$ is an element in $\mathbf{P}^2(\mathbf{F})$ but not in $\mathbf{A}^2(\mathbf{F})$. If it was, it would induce that $t0 = 1$ which is clearly a contradiction. This observation makes it necessary to define a second mapping. This time from $\mathbf{F}$ to $\mathbf{P}^2(\mathbf{F})$:

$$j : \mathbf{F} \to \mathbf{P}^2(\mathbf{F})$$
$$j(a) = [a : 1 : 0]$$

Like before this mapping is injective and the image of j contains all points $[a : b : 0]$ in $\mathbf{P}^2(\mathbf{F})$ when $b \neq 0$. But still one point is missing from the set, and that is $[1 : 0 : 0]$. So $[a : 0 : 0] = [1 : 0 : 0]$ for all $a \neq 0$.

Combining those thoughts on i, j and the point $[1 : 0 : 0]$, it can be concluded that $\mathbf{P}^2(\mathbf{F})$ can be written as a union of $i(\mathbf{A}^2(\mathbf{F}))$, $j(\mathbf{F})$ and $[1 : 0 : 0]$:

$$\mathbf{P}^2(\mathbf{F}) = i(\mathbf{A}^2(\mathbf{F})) \cup j(\mathbf{F}) \cup \{[1 : 0 : 0]\}$$

Definition 4.7. Let g be a polynomial in X, Y and Z over $\mathbf{F}$. g is called homogeneous of degree d if:

$$g(X, Y, Z) = \sum_{\mu_1,\mu_2,\mu_3 \geqslant 0} \alpha_{\mu_1,\mu_2,\mu_3} X^{\mu_1} Y^{\mu_2} Z^{\mu_3}$$

with $\alpha_{\mu_1,\mu_2,\mu_3}$ not all zero, and $\mu_1 + \mu_2 + \mu_3 = d$ for the case that $\alpha_{\mu_1,\mu_2,\mu_3} \neq 0$

An example of a homogeneous polynomial of degree 3 would be $g(X, Y, Z) = Y^2 Z - X^3 - Y Z^2 - Z^3$.

Lemma 4.8. *Let $g \in \mathbf{F}[X, Y, Z]$ be a homogeneous polynomial of degree d and $a, b, c \in \mathbf{F}$ and $t \in \mathbf{F}\backslash\{0\}$ Then it holds that:*

$$g(a, b, c) = 0 \Leftrightarrow g(ta, tb, tc) = 0.$$

Proof. Let

$$g = \sum_{\mu_1,\mu_2,\mu_3 \geqslant 0} \alpha_{\mu_1,\mu_2,\mu_3} X^{\mu_1} Y^{\mu_2} Z^{\mu_3}.$$

From this follows then that

$$g(ta, tb, tc) = \sum_{\mu_1,\mu_2,\mu_3 \geqslant 0} \alpha_{\mu_1,\mu_2,\mu_3} (ta)^{\mu_1} (tb)^{\mu_2} (tc)^{\mu_3}.$$

Using the exponent rules gives that

$$= \sum_{\mu_1,\mu_2,\mu_3 \geqslant 0} \alpha_{\mu_1,\mu_2,\mu_3} t^d (a)^{\mu_1} (b)^{\mu_2} (c)^{\mu_3} = t^d g(a, b, c).$$

The sum of the exponents equals d in every nonzero term and the statement is proved because $g(a, b, c) = 0$ by assumption and therefore $t^d 0 = 0$. □

This tells that when (a, b, c) is zero for the polynomial g also the multiplies of this zero (ta, tb, tc) have to be a zeros in $\mathbf{F}$.

This leads to the next definition of *projective plane curves.*

Definition 4.9 (Projective Plane Curve)**.** Set $g \in \mathbf{F}[X, Y, Z]$ to be a homogeneous polynomial. The set of roots of g in $\mathbf{P}^2(F)$ shall be denoted $C_g(\mathbf{F})$

$$C_g(\mathbf{F}) = \{[a : b : c] \in \mathbf{P}^2(\mathbf{F}) \mid g(a, b, c) = 0\}.$$

The set $C_g(\mathbf{F})$ of roots is then called *projective plane curve.*

Revisiting the example from before, there are the two polynomials $f(x, y) = y^2 - x^3 - x - 1 \in \mathbf{A}^2(\mathbf{F})$ and now also $g(X, Y, Z) = Y^2Z - X^3 - XZ^2 - Z^3 \in \mathbf{P}^2(\mathbf{F})$ and their sets of zeros $C_f(\mathbf{F})$ and $C_g(\mathbf{F})$. From the previous pages it can be concluded that every solution in $C_f(\mathbf{F})$ is also contained in $C_g(\mathbf{F})$. This is true because the mapping $i : (a, b) \to [a : b : 1]$ is an injective mapping from $\mathbf{A}^2(\mathbf{F})$ to $\mathbf{P}^2(\mathbf{F})$.

At the beginning of this section, the equation of an affine curve got extended to introduce the idea of the projective space. The theory developed until now is based on the case $c \neq 0$. For $c = 0$ the curve has an additional zero namely $[0 : 1 : 0]$, which is not in $C_f(\mathbf{F})$ but in $C_g(\mathbf{F})$. So:

$$C_g(\mathbf{F}) = i(C_f(\mathbf{F})) \cup \{[0 : 1 : 0]\}.$$

Until now it is shown that the affine curve $C_f(\mathbf{F})$ can be embedded into the projective curve $C_g(\mathbf{F})$. But this projective curve $C_g(\mathbf{F})$ contains an additional point which is not part of $C_f(\mathbf{F})$. This point is, in elliptic curve cryptography, usually refereed to as *point at infinity* and will be denoted $\mathcal{O}$.

Proposition 4.10. *Let f be any nonzero polynomial in $\mathbf{F}[x, y]$, additionally $f(x, y) = \sum_{\mu_1,\mu_2 \geqslant 0, \mu_1+\mu_2 \leqslant d} \alpha_{\mu_1,\mu_2} x^{\mu_1} y^{\mu_2}$ where the coefficients α_{μ_1,μ_2} are in $\mathbf{F}$, a polynomial of degree d. d is the maximum of all $\mu_1 + \mu_2$ for any nonzero α_{μ_1,μ_2}. The following polynomial of degree d:*

$$g(X, Y, Z) = \sum_{\mu_1,\mu_2 \geqslant 0, \mu_1+\mu_2 \leqslant d} \alpha_{\mu_1,\mu_2} X^{\mu_1} Y^{\mu_2} Z^{d-\mu_1-\mu_2}.$$

is then homogeneous and satisfies the condition $g(a, b, 1) = f(a, b)$ for every pair $(a, b) \in \mathbf{A}^2(\mathbf{F})$. The mapping $i : \mathbf{A}^2(\mathbf{F}) \to \mathbf{P}^2(\mathbf{F})$ maps $C_f(\mathbf{F})$ to $C_g(\mathbf{F})$. If a point $[a : b : c] \in \mathbf{P}^2(F)$ can be written as $i(x)$ for some $x \in \mathbf{A}^2(\mathbf{F})$, it follows that x is in $C_f(\mathbf{F})$.

Proof. The polynomial g is by assumption homogeneous and of degree d. Also the following equality holds:

$$g(a, b, 1) = f(a, b)$$

and it follows that $i(a,b) = [a:b:1] \in C_g(\mathbf{F})$ for all points $(a,b) \in C_f(\mathbf{F})$. If for choosing an arbitrary $(a,b) \in \mathbf{A}^2(\mathbf{F})$ the point $i(a,b) = [a:b:1]$ is in $C_g(\mathbf{F})$, so is $g(a,b,1) = 0$ as well as $f(a,b)$. From this follows that $(a,b) \in C_g(\mathbf{F})$. □

Other mappings between $\mathbf{A}^2(\mathbf{F})$ and $\mathbf{P}^2(\mathbf{F})$ are also possible. For example:

$$i_1(a,b) = [1:a:b] \text{ or } i_2(a,b) = [a:1:b].$$

All three copies of $i(\mathbf{A}^2(\mathbf{F})), i_1(\mathbf{A}^2(\mathbf{F})), i_2(\mathbf{A}^2(\mathbf{F}))$ overlap. So $i(a,b)$ can be rewritten in terms of $i_1(b/a, 1/a)$ or $i_2(a/b, 1/b)$ for all $a, b \neq 0$. Those three sets combined contain all points $[a:b:c]$ in $\mathbf{P}^2(\mathbf{F})$, because for each one of those sets, either a, b or c is not equal to zero.

Instead of writing the mapping in terms of i, i_1 or i_2 one can also denote it as:

$$C_g(\mathbf{F}) \cap \mathbf{A}^2(\mathbf{F}) = C_f(\mathbf{F})$$

or to point out the mapping i_1 and i_2

$$C_g(\mathbf{F}) \cap i_1(\mathbf{A}^2(\mathbf{F})) = C_f(\mathbf{F})$$
$$C_g(\mathbf{F}) \cap i_2(\mathbf{A}^2(\mathbf{F})) = C_f(\mathbf{F})$$

Those two mappings i_1 and i_2 have a similar result as i, namely:

Proposition 4.11. *Let the homogeneous polynomial of degree d be defined by $g = \sum_{\mu_1,\mu_2,\mu_3 \geqslant 0} \alpha_{\mu_1,\mu_2,\mu_3} X^{\mu_1} Y^{\mu_2} Z^{\mu_3}$ with $\mu_1 + \mu_2 + \mu_3 = d$ for all nonzero coefficients. Then*

$$C_g(\mathbf{F}) \cap i_1(\mathbf{A}^2(\mathbf{F})) = i_1(C_{f_1}(\mathbf{F}))$$

for $f_1(x,y) = \sum_{\mu_2,\mu_3 \geqslant 0, \mu_2+\mu_3 \leqslant d} \alpha_{d-\mu_2-\mu_3,\mu_2,\mu_3} x^{\mu_2} y^{\mu_3}$ and

$$C_g(\mathbf{F}) \cap i_2(\mathbf{A}^2(\mathbf{F})) = i_2(C_{f_2}(\mathbf{F}))$$

for $f_1(x,y) = \sum_{\mu_1,\mu_3 \geqslant 0, \mu_1+\mu_3 \leqslant d} \alpha_{d-\mu_1-\mu_3,\mu_1,\mu_3} x^{\mu_1} y^{\mu_3}$.

Proof. This proof is very similar to the proof of proposition 4.10. □

Definition 4.12. Let g denote a homogeneous polynomial in $\mathbf{F}[X,Y,Z]$ of degree d.

1. A projecive plane curve $C_g(\mathbf{F})$ is called singular in a point $[a:b:c] \in C_g(\mathbf{F})$ if all derivatives of g in [a : b : c] are zero.

$$\frac{\partial g}{\partial X}(a,b,c) = \frac{\partial G}{\partial Y}(a,b,c) = \frac{\partial g}{\partial Z}(a,b,c) = 0$$

2. The curve $C_g(\mathbf{F})$ shall be called non-singular if $C_g(\bar{\mathbf{F}})$ does not contain any singular points.

Lemma 4.13. *Let $g(X, Y, Z) = \sum_{\mu_1,\mu_2,\mu_3 \geqslant 0} \alpha_{\mu_1,\mu_2,\mu_3} X^{\mu_1} Y^{\mu_2} Z^{\mu_3}$ be a homogeneous polynomial of degree d and $f(x, y) = \sum_{\mu_1,\mu_2 \geqslant 0, \mu_1+\mu_2 \leqslant d} \alpha_{\mu_1,\mu_2,d-\mu_1-\mu_2} x^{\mu_1} y^{\mu_2}$. For all points $P \in C_g(\mathbf{F})$ it then holds that: If $P = i(Q)$ is in $i(\mathbf{A}^2(\mathbf{F}))$, then $C_g(\mathbf{F})$ is singular in P if and only if the affine curve $C_f(\mathbf{F})$ is singular in Q.*

Proof. Let $Q = (a, b)$, then proposition 4.10 ensures that Q is on the affine curve $C_f(\mathbf{F})$. It follows then that $P = i(Q) = [a : b : 1]$.

$$\frac{\partial g}{\partial X}(X, Y, Z) = \sum_{\mu_1 > 0, \mu_2, \mu_3 \geqslant 0} \mu_1 \alpha_{\mu_1,\mu_2,\mu_3} X^{\mu_1 - 1} Y^{\mu_2} Z^{\mu_3}$$

such that $\frac{\partial g}{\partial X}(a, b, 1) = \frac{\partial f}{\partial x}(a, b)$ as well as $\frac{\partial g}{\partial Y}(a, b, 1) = \frac{\partial f}{\partial y}(a, b)$. In addition:

$$\frac{\partial g}{\partial Z}(X, Y, Z) = \sum_{\mu_1, \mu_2 \geqslant 0, \mu_3 > 0} \mu_3 \alpha_{\mu_1,\mu_2,\mu_3} X^{\mu_1} Y^{\mu_2} Z^{\mu_3 - 1}$$

From which it follows that:

$$\frac{\partial g}{\partial Z}(a, b, 1) = \sum_{\mu_1, \mu_2 \leqslant 0, \mu_3 > 0} \mu_3 \alpha_{\mu_1,\mu_2,\mu_3} a^{\mu_1} b^{\mu_2}.$$

If $\mu_3 = 0$ the respective addend disappears. As $\mu_1 + \mu_2 + \mu_3 = d$ it follows:

$$\begin{aligned} \frac{\partial g}{\partial Z}(a, b, 1) &= \sum_{\mu_1, \mu_2 \leqslant 0, \mu_1 + \mu_2 \leqslant 0} \alpha_{\mu_1,\mu_2,d-\mu_1\mu_2} (d - \mu_1 - \mu_2) a^{\mu_1} b^{\mu_2} \\ &= df(a, b) - a\frac{\partial f}{\partial x}(a, b) - b\frac{\partial f}{\partial y}(a, b). \end{aligned}$$

□

4.3 Elliptic Curves

Now that affine and projective curves have been introduced, it is time to focus on a special kind of projective curves, the so called *elliptic curves*. What makes them interesting is the fact that a group law can be defined such that the points on the curve form an abelian group.

Definition 4.14. A non-singular projective plane curve $C_g(\mathbf{F})$ with g being a homogeneous polynomial of degree three of the form:

$$g(X, Y, Z) = Y^2 Z + a_1 XYZ + a_3 YZ^2 - X^3 - a_2 X^2 Z - a_4 XZ^2 - a_6 Z^3$$

with $a_1, a_2, a_3, a_4, a_6 \in \mathbf{F}$ is called an *elliptic curve*.

An elliptic curve written in this form is called a Weierstraß equation.

Proposition 4.15. *Let $C_g(\mathbf{F})$ be an elliptic curve where g of the form:*

$$g(X,Y,Z) = Y^2Z + a_1XYZ + a_3YZ^2 - X^3 - a_2X^2Z - a_4XZ^2 - a_6Z^3.$$

Case 1: *char(*$\mathbf{F}$*)* $\neq 2$
The mapping

$$\begin{aligned} \Phi : \mathbf{P}^2(\mathbf{F}) &\to \mathbf{P}^2(\mathbf{F}) \\ [r:s:t] &\to [r : s + \frac{a_1}{2}r + \frac{a_3}{2}t : t] \end{aligned}$$

is bijective and it holds that:

$$\Phi(C_g(\mathbf{F})) = C_{h_1}(\mathbf{F})$$

where $h_1(X,Y,Z) = Y^2Z - X^3 - \frac{1}{4}b_2X^2Z - \frac{1}{2}b_4XZ^2 - \frac{1}{4}b_6Z^3$ *with* $b_2 = a_1^2 + 4a_2, b_4 = 2a_4 + a_1a_3$ *and* $b_6 = a_3^2 + 4a_6$*.* $C_{h_1}(\mathbf{F})$ *is then also an elliptic curve.*

Case 2: *char(*$\mathbf{F}$*)* $\neq 2, 3$
The mapping

$$\begin{aligned} \Psi : \mathbf{P}^2(\mathbf{F}) &\to \mathbf{P}^2(\mathbf{F}) \\ [r:s:t] &\to [36r + 3b_2t : 216s : t] \end{aligned}$$

is bijective and it holds that:

$$\Psi(C_g(\mathbf{F})) = C_{h_2}(\mathbf{F})$$

where $h_2(X,Y,Z) = Y^2Z - X^3 + 27c_4XZ^2 + 54c_6Z^3$ *with* $c_4 = b_2^2 - 24b_4$ *and* $c_6 = -b_2^3 + 36b_2b_4 - 216b_6$*.* $C_{h_2}(\mathbf{F})$ *is then also an elliptic curve.*

Case 3: *char(*$\mathbf{F}$*)* $= 2$ *and* $a_1 \neq 0$
The mapping

$$\begin{aligned} \Theta : \mathbf{P}^2(\mathbf{F}) &\to \mathbf{P}^2(\mathbf{F}) \\ [r:s:t] &\to [\frac{1}{a_1^2}r + \frac{a_3}{a_1}t : a_1^3s + \frac{a_1^2a_4 + a_3^2}{a_1^3}t : t] \end{aligned}$$

is bijective and it holds that:

$$\Theta(C_g(\mathbf{F})) = C_{h_2}(\mathbf{F})$$

where $h_3(X,Y,Z) = Y^2Z + XYZ - X^3 - a_2'X^2Z - a_6Z^3$ *with* $a_2' = \frac{a_3+a_1a_2}{a_1^3}$ *and* $a_6' = \frac{a_1^6a_6+a_1^5a_3a_4+a_1^4a_2a_3^2+a_1^4a_4^2+a_1^3a_3^3+a_3^4}{a_1^{12}}$*.* $C_{h_2}(\mathbf{F})$ *is then also an elliptic curve.*

This proposition shows that in the case char($\mathbf{F}$)$\neq 2$ the Weierstraß equation can be turned into the simplified Weierstraß equation of the form:

$$Y^2Z = X^3 + a_2X^2Z + a_4X^2 + a_6Z^3$$

with new coefficients a_i. In the case that char($\mathbf{F}$)$\neq 2, 3$ the equation gets even simpler

$$Y^2Z = X^3 + a_4XZ^2 + a_6Z^3$$

Proof. **Case 1:** char($\mathbf{F}$) $\neq 2$
From the definition of the mapping it is clear that char($\mathbf{F}$) has to be unequal to two, otherwise the mapping would make no sense. It is easy to find the inverse mapping

$$\Phi^{-1}([r : s : t] = [r : s - \frac{a_1}{2}r - \frac{a_3}{2}t : t])$$

which means that the mapping is bijective. Φ and Φ^{-1} will also be used for the mapping from $\mathbf{F}^3$ to $\mathbf{F}^3$ with $\Phi(s, r, t) = (r, s + \frac{a_1}{2}r + \frac{a_3}{2}t, t)$ and $\Phi^{-1}(s, r, t) = (r, s - \frac{a_1}{2}r - \frac{a_3}{2}t, t)$. Now $h_1(X, Y, Z) = g(X, Y - \frac{a_1}{2}X - \frac{a_3}{2}Z, Z)$:

$$\begin{aligned}
g(X, Y - &\frac{a_1}{2}X - \frac{a_3}{2}Z, Z)\\
&= (Y - \frac{a_1}{2}X - \frac{a_3}{2}Z)^2Z + a_1X(Y - \frac{a_1}{2}X - \frac{a_3}{2}Z)Z\\
&+ a_3(Y - \frac{a_1}{2}X - \frac{a_3}{2}Z)Z^2 - X^3 - a_2X^2Z - a_4XZ^2 - a_6Z^3\\
&= \left[Y^2 - 2Y(\frac{a_1}{2}X + \frac{a_3}{2}Z) + (\frac{a_1^2}{4}X^2 + 2\frac{a_1a_3}{4}XZ + \frac{a_3^2}{4}Z^2)\right]Z\\
&+ a_1XYZ - \frac{a_1^2}{2}X^2Z - \frac{a_1a_3}{2}XZ^2 + a_3YZ^2 - \frac{a_1a_3}{XZ^2}\\
&- \frac{a_3^2}{2}Z^3 - X^3 - a_2X^2Z - a_4XZ^2 - a_6Z^3\\
&= Y^2Z - X^3 + (-\frac{a_1^2}{4} - a_2)X^2Z + (-\frac{a_1a_3}{2} - a_4)XZ^2\\
&+ (-\frac{a_3^2}{4} - a_6)Z^3\\
&= Y^2Z - X^3 - \frac{1}{4}b_2X^2Z - \frac{1}{2}b_4XZ^2 - \frac{1}{4}b_6Z^3\\
&= h_1(X, Y, Z)
\end{aligned}$$

This shows $h_1(r, s, t) = g(\Phi^{-1}(r, s, t))$ and therefore that $g(r, s, t) = 0$ if and only if $h_1(\Phi(r, s, t)) = 0$. From this it follows:

$$\Phi(C_g(\mathbf{F})) = C_{h_1}(\mathbf{F})$$

The polynomial h_1 is then an elliptic curve if it can be shown that $C_{h_1}(\mathbf{F})$ is non-singular. Using the chain rule gives:

$$\begin{aligned}
\frac{\partial h_1}{\partial X}(r,s,t) &= \frac{\partial g}{\partial X}(\Phi^{-1}(r,s,t)) - \frac{a_1}{2}\frac{\partial g}{\partial Y}(\Phi^{-1}(r,s,t)) \\
\frac{\partial h_1}{\partial Y}(r,s,t) &= \frac{\partial g}{\partial Y}(\Phi^{-1}(r,s,t)) \\
\frac{\partial h_1}{\partial Z}(r,s,t) &= -\frac{a_3}{2}\frac{\partial g}{\partial Y}(\Phi^{-1}(r,s,t)) + \frac{\partial g}{\partial Z}(\Phi^{-1}(r,s,t))
\end{aligned}$$

This shows that for every point $P = [r:s:t]$ in $C_{h_1}(\bar{\mathbf{F}})$ there exists $\Phi^{-1}[r:s:t]$ a point in $C_g(\bar{\mathbf{F}})$. The derivatives of g don't vanish all simultaneously in this point and therefore also not all derivatives of h_1 in (r,s,t). This proves that C_{h_1} is non-singular over the closure $\bar{\mathbf{F}}$ and therefore an elliptic curve.

Case 2: char($\mathbf{F}$) $\neq 2,3$
This mapping is also bijective with

$$\Psi^{-1}([r:s:t]) = [\frac{1}{36}r - \frac{b_2}{12}t : \frac{1}{216}s : t]$$

being the mappings inverse. All the denominators contain powers of 2 and 3, therefore char($\mathbf{F}$) gives no problem. Similar to the previous case, one can show that:

$$h_2(X,Y,Z) = 2^6 3^6 h_1(\frac{1}{36}X - \frac{b_2}{12}Z, \frac{1}{216}Y, Z).$$

This shows that $h_1(r,s,t) = 0$ if and only if $h_2(\Psi(r,s,t)) = 0$ indicating that:

$$\Psi(C_{h_1}(\mathbf{F})) = C_{h_2}(\mathbf{F}).$$

The polynomial h_2 has therefore the desired form of an elliptic curve. Repeating the step of calculating the derivatives using the chain rule reveals that the curve $C_{h_2}(\mathbf{F})$ is like $C_{h_1}(\mathbf{F})$ non-singular and therefore and elliptic curve.

Case 3: char($\mathbf{F}$) $= 2$ and $a_1 \neq 0$
The mapping Θ has an inverse given by:

$$[r:s:t] \rightarrow [a_1^2 r + \frac{a_3}{a_1}t : a_1^3 s + \frac{a_1^2 a_4 + a_3^2}{a_1^3}t : t].$$

Repeating the same procedure as before shows that:

$$a_1^6 h_3(X,Y,Z) = g(a_1^2 X + \frac{a_3}{a_1}Z, a_1^3 Y + \frac{a_1^2 a_4 + a_3^2}{a_1^3}Z, Z).$$

Like in the two previous cases, using the chain rule to calculate the derivatives of h_3 reveals that $C_{h_3}(\mathbf{F})$ is also an elliptic curve.

□

Definition 4.16. Let $g(X,Y,Z) = Y^2Z + a_1XYZ + a_3YZ^2 - X^3 - a_2X^2Z - a_4XZ^2 - a_6Z^3$ be a Weierstraß polynomial. Then the discriminant of the curve $C_g(\mathbf{F})$ is defined as:

$$\Delta = -b_2^2b_8 - 8b_4^3 - 27b_6^2 + 9b_2b_4b_6$$

where:

$$\begin{aligned} b_2 &= a_1^2 + 4a_2 \\ b_4 &= 2a_4 + a_1a_3 \\ b_6 &= a_3^2 + 4a_6 \\ b_8 &= a_1^2a_6 + 4a_2a_6 - a_1a_3a_4 + a_2a_3^2 - a_4^2 \end{aligned}$$

The discriminant is used to examine if a curve in Weierstraß form is non-singular or not.

Proposition 4.17. *Let $g(X,Y,Z) = Y^2Z + a_1XYZ + a_3YZ^2 - X^3 - a_2X^2Z - a_4XZ^2 - a_6Z^3$ be an Weierstraß polynomial. The curve $C_g(\mathbf{F})$ is non-singular if and only if the discriminant $\Delta = -b_2^2b_8 - 8b_4^3 - 27b_6^2 + 9b_62b_4b_6$ is not equal to zero.*

Proof. From the definition of singularity, one recalls that the elliptic curve $C_g(\mathbf{F})$ is singular exactly when the affine curve $C_f(\mathbf{F})$ is non-singular.

$$f(x,y) = y^2 + a_1xy + a_3y - x^3 - a_2x^2 - a_4x - a_6$$

This is only the case if $C_f(\bar{\mathbf{F}})$ is non-singular. $C_f(\mathbf{F})$ is singular if there are elements $r, s \in \bar{\mathbf{F}}$ such that.

$$f(r,s) = \frac{\partial f}{\partial x}(r,s) = \frac{\partial f}{\partial y}(r,s) = 0 \tag{1}$$

with:

$$\frac{\partial f}{\partial x}(r,s) = a_1s - 3r^2 - 2a_2r - a_4 \tag{2}$$
$$\frac{\partial f}{\partial y}(r,s) = 2s + a_1r + a_3 \tag{3}$$

Case 1: char($\mathbf{F}$)$= 2$ and $a_1 = 0$

Using the fact that multiplies of two equal zero in $\mathbf{F}_2$ reduces the discriminant from $\Delta = -b_2^2b_8 - 8b_4^3 - 27b_6^2 + 9b_62b_4b_6$ to $\Delta = -27a_3^4 = a_3^4$ as $b_2 = b_4 = 0$ and $b_6 = a_3^2$. Additionally the partial derivative $\frac{\partial f}{\partial y}$ equals a_3. This shows that the curve $C_f(\bar{\mathbf{F}})$ contains a singular point for $a_3 = 0$, which also forces $\Delta = 0$.

If one assumes that the converse, $\Delta = 0$ is true, also $\frac{\partial f}{\partial y} = 0$. As $\bar{\mathbf{F}}$ is the algebraic closure there exists an $r \in \bar{\mathbf{F}}$ such that the following equation, coming from 2, has a solution.

$$r^2 + a_4 = 0$$

Then there exists an $s \in \bar{\mathbf{F}}$ such that:

$$s^2 + a_3 s = r^3 + a_2 r^2 + a_4 r + a_6.$$

This shows that (r, s) is singular point on the curve $C_f(\bar{\mathbf{F}})$.

Case 2: $\text{char}(\mathbf{F}) = 2$ and $a_1 \neq 0$
The calculations rules implied by the fields characteristic, again reduce the discriminant. This time to:

$$\Delta = a_1^6 a_6 + a_1^5 a_3 a_4 + a_1^4 a_2 a_3^2 + a_1^4 a_4^2 + a_1^3 a_3^3 + a_3^4.$$

If $C_f(\bar{\mathbf{F}})$ contains a singular point, $r, s \in \bar{\mathbf{F}}$ can be obtained from the equations:

$$\begin{aligned} f(r,s) &= 0 \\ a1_s + r^2 + a_4 &= 0 \\ a_1 r + a_3 &= 0 \end{aligned}$$

In this case $a_3 \neq 0$, so it follows that

$$r = \frac{a_3}{a_1} \text{ and } s = \frac{a_3^2 + a_1^2 a_4}{a_1^3}.$$

Inserting this into $f(r, s)$ gives:

$$f(r,s) = \frac{\Delta}{a_1^6}$$

From which it follows that $\Delta = 0$. Again, assuming that $\Delta = 0$ gives:

$$r = \frac{a_3}{a_1} \text{ and } s = \frac{a_3^2 + a_1^2 a_4}{a_1^3},$$

and as before, $f(r, s) = \frac{\Delta}{a_1^6}$ which leads to the conclusion that also $f(r, s) = 0$ in this case. This leads to the conclusion that $C_f(\mathbf{F})$ contains a singular point.

Case 3: $\text{char}(\mathbf{F}) = 3$
In this case, the discriminant simplifies to

$$\Delta = -b_2^2 b_8 - 8b_4^3$$

Let the mapping:

$$\Phi : C_g(\mathbf{F}) \to C_{h_1}(\mathbf{F})$$
$$[r : s : t] \to [r : s + \frac{a_1}{2}r + \frac{a_3}{2}t : t]$$

be the same mapping as in the proof of proposition 4.15, but $h_1(X, Y, Z) = Y^2Z - X^3 - \frac{1}{4}b_2X^2Z - \frac{1}{2}b_2XZ^2 - \frac{1}{4}b_6Z^3$. Calculating the derivatives of h_1 shows, like in the proof of 4.15, that the curve $C_g(\mathbf{F})$ is non-singular if and only if $C_{h_1}(\mathbf{F})$ is non-singular. When calculating the discriminant, following definition 4.16 one has to set $a'_1 = a'_3 = 0, a'_2 = \frac{1}{4}b_2, a'_4 = \frac{1}{2}b_4$ and $a'_6 = \frac{1}{4}b_6$. This shows that for $i = 2, 4, 6, 8$ it holds that $b'_i = b_i$, which means that $C_g(\mathbf{F})$ and $C_{h_1}(\mathbf{F})$ have the same discriminant. Therefore it is sufficient to prove the statement for one of the curves.

The curve $C_{h_1}(\bar{\mathbf{F}})$ contains a singular point if and only if there exist elements $r, s \in \bar{\mathbf{F}}$ such that:

$$\begin{aligned} s^2 - r^3 - \frac{1}{4}b_2r^2 - \frac{1}{2}b_4r - \frac{1}{4}b_6 &= 0 \\ 3r^2 + \frac{1}{2}b_2r + \frac{1}{2}b_4 &= 0 \\ 2s &= 0 \end{aligned}$$

If there exists such r in $\bar{\mathbf{F}}$, then the polynomial $\sigma(x) = x^3 + \frac{1}{4}b_2x^2 + \frac{1}{2}b_4x + \frac{1}{4}b_6$ and its derivative vanish at the same point. Over the algebraic closure $\bar{\mathbf{F}}$ the function σ factors to:

$$\sigma(x) = (x - \alpha_1)(x - \alpha_2)(x - \alpha_3)$$

for some $\alpha_1, \alpha_2, \alpha_3 \in \bar{\mathbf{F}}$. Differentiating the equation reveals that function and derivative can only be equal to zero at the same point if and only if the polynomial has a double root. To examine if the polynomial has a double root, one can use the discriminant of the polynomial. For $\sigma(x) = (x - \alpha_1)(x - \alpha_2)(x - \alpha_3)$ it is defined as:

$$D\sigma = (\alpha_1 - \alpha_2)^2(\alpha_1 - \alpha_3)^2(\alpha_2 - \alpha_3)^2$$

This leaves the task to show that the discriminant Δ equals 0 exactly when $D\sigma = 0$. In the general case the discriminant of a cubic is defined as:

$$D(ax^3 + bx^2 + cx + d) = b^2c^2 - 4ac^3 - 4b^3d - 27a^2d^2 + 18abcd$$

Because $\mathbf{F}$ has a characteristic of three, the equation simplifies to:

$$D(\sigma) = \frac{1}{64}b_2^2b_4^2 - \frac{1}{64}b_2^3b_6 - \frac{1}{2}b_4^3$$

Using the relation $4b_8 = b_2b_6 - b_4^2$ leads to

$$D(\sigma) = \frac{1}{16}(-b_2^2b_8 - 8b_4^3)\frac{1}{16}\Delta$$

which proves the proposition for a field of characteristic three.

Case 4: $\text{char}(\mathbf{F}) > 3$
Reusing the bijection from 4.15:

$$\Psi \circ \Phi : C_g(\mathbf{F}) \to C_{h_2}(\mathbf{F})$$

with $h_2(X, Y, Z) = Y^2Z - X^3 + 27c_4XZ^2 + 54c_6Z^3$ reveals also for this case, after calculating the derivatives, that $C_g(\mathbf{F})$ is non-singular if and only if $C_{h_2}(\mathbf{F})$ is non-singular as well. Calculaiting the discriminant of $C_{h_2}(\mathbf{F})$ gives:

$$2^6 3^9(c_4^3 - c_6^2) = 2^{12}3^{12}\Delta$$

It is therefore again sufficient to show the proposition for one of the curves. Like in the previous case $C_{h_2}(\bar{\mathbf{F}})$ has a singular point if and only if the polynomial $x^3 - 27c_4x - 54c_6$ has a double root. This is the case if the discriminant vanishes. This is the case when $4 \cdot 27^3c_4^3 - 27 \cdot 54^2c_6^2 = 0$, which only happens for $c_4^3 - c_6^2 = 0$. From this, the proposition follows.

□

Definition 4.18. Let a homogeneous polynomial of degree 1 be called $g \in \mathbf{F}[X, Y, Z]$.

$$g(X, Y, Z) = \alpha X + \beta Y + \gamma Z \qquad \alpha, \beta, \gamma \in \mathbf{F}$$

When the coefficients α, β, γ are not all equal to zero, this curve is called a projective line. It can be written as $L(\alpha, \beta, \gamma)$ instead of $C_g(\mathbf{F})$.

Recalling the definition of singularity, it is clear that a line is a non-singular curve. All the derivatives are constants for any point P on the line and therefore never equal to zero at the same time.

Recalling proposition 4.10, the intersection of the line $C_g(\mathbf{F})$ with $i(\mathbf{A}^2(\mathbf{F}))$ gives an affine curve in $\mathbf{A}^2(\mathbf{F}) = \mathbf{F} \times \mathbf{F}$. The line in the affine space be given by the equation:

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

In the case that $\alpha = \beta = 0$ the value of γ has to be unequal to zero. This means that $C_g(\mathbf{F})$ is the empty set. If α and β do not vanish simultaneously, the sets follow the rules of ordinary lines in a plane.

$$C_f(\mathbf{F}) = \left\{(x, y) \in \mathbf{F} \times \mathbf{F} \mid y = -\frac{\alpha}{\beta}x - \frac{\gamma}{\beta}\right\}$$

and

$$C_f(\mathbf{F}) = \left\{(x, y) \in \mathbf{F} \times \mathbf{F} \mid x = -\frac{\gamma}{\alpha}\right\} \quad \text{for } \beta = 0 \text{ and } \alpha \neq 0$$

What happens with parallel lines in $\mathbf{A}^2(\mathbf{F}) = \mathbf{F} \times \mathbf{F}$ when they get mapped to $\mathbf{P}^2(\mathbf{F})$? Let f and f_c be two lines in $\mathbf{A}^2(\mathbf{F})$

$$f(x, y) = y - ax \qquad f_c(x, y) = y - ax - c \quad \text{with } a \in \mathbf{F}, c \neq 0$$

Let the sets of their points be denoted by C_f and C_{f_c} with:

$$\begin{aligned} C_f &= \{(x, y) \in \mathbf{F} \times \mathbf{F} \mid y = ax\} \\ C_{f_c} &= \{(x, y) \in \mathbf{F} \times \mathbf{F} \mid y = ax + c\} \end{aligned}$$

Mapping C_f and C_{f_c} into the projective space gives C_f and C_{g_c} with the corresponding projective lines:

$$\begin{aligned} g(X, Y, Z) &= Y - aX \\ g_c(X, Y, Z) &= Y - aX - cZ \end{aligned}$$

Recalling that the projective space contains the affine space, one can write $C_g(\mathbf{F}) \cap \mathbf{A}^2(\mathbf{F}) = C_f(\mathbf{F})$ and $C_{g_c}(\mathbf{F}) \cap \mathbf{A}^2(\mathbf{F}) = C_{f_c}(\mathbf{F})$

Calculating their intersection using linear algebra gives a point that does not lie in the affine plane $\mathbf{A}^2(\mathbf{F})$ but in the projective plane.

$$\left[\begin{vmatrix} 1 & 0 \\ 1 & -c \end{vmatrix} : \begin{vmatrix} 0 & -a \\ -c & -a \end{vmatrix} : \begin{vmatrix} -a & 1 \\ -a & 1 \end{vmatrix}\right] = [-c : -ac : 0] = [1 : a : 0]$$

This intersection at *infinity* leads to the next lemma.

Lemma 4.19. *For projective lines the two following statements hold:*

1. *The line going through two points in the projective plane* $\mathbf{P}^2(\mathbf{F})$ *is unique.*

2. *Two dissimilar projective lines intersect in exactly one point in* $\mathbf{P}^2(\mathbf{F})$

Proof. 1. Let there be two different points $P_1 = [a_1 : b_1 : c_1]$ and $P_2 = [a_2 : b_2 : c_3]$ in $\mathbf{P}^2(\mathbf{F})$. To find a solution (α, β, γ) such both points are on one line, one has to solve a system of linear equations:

$$a_1\alpha + b_1\beta + c_1\gamma = 0$$
$$a_2\alpha + b_2\beta + c_2\gamma = 0$$

This can be rewritten to a matrix:

$$\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{pmatrix}$$

As the points P_1 and P_2 are different, the lines of the matrix are linear independent, revealing a rank of 2. Therefore the rank-nullity theorem gives that the nullity is one because rank and nullity have to sum up to the dimension of the matrix which is three. This implies the existence of a solution $(\alpha, \beta, \gamma) \neq 0$ such that $P_1 \in L(\alpha, \beta, \gamma)$ and $P_2 \in L(\alpha, \beta, \gamma)$ are on the same line. Every other solution $(\alpha', \beta', \gamma')$ is a multiple of (α, β, γ).

2. Let $L_1 = L(\alpha_1, \beta_1, \gamma_1)$ and $L_1 = L(\alpha_2, \beta_2, \gamma_2)$ be two different projective lines. Then the matrix $\begin{pmatrix} \alpha_1 & \beta_1 & \gamma_1 \\ \alpha_2 & \beta_2 & \gamma_2 \end{pmatrix}$ has rank two and its kernel is one dimensional. Therefore there exists a solution not equal to zero. Let that solution be called $P = [a : b : c] \in \mathbf{P}^2(\mathbf{F})$. P then lies on both projective lines $L_1 = L(\alpha_1, \beta_1, \gamma_1)$ and $L_1 = L(\alpha_2, \beta_2, \gamma_2)$. Every other point $P' = [a' : b' : c'] \neq [0 : 0 : 0]$ that is in the kernel has to be a multiple of P.

□

This result shows that there are no parallel lines in the projective space.

Definition 4.20. Let $P = [a : b : c]$ be a non-singular point on the projective plane curve $C_g(\mathbf{F})$. The projective line:

$$L(\frac{\partial g}{\partial X}(a, b, c), \frac{\partial g}{\partial Y}(a, b, c), \frac{\partial g}{\partial Z}(a, b, c))$$

is then called the tangent line at P.

After having defined lines and tangent lines in the projective space, it is of interest how often lines and elliptic curves intersect. A result from algebraic geometry comes in handy. Bézout's theorem states that two plane projective curves g, f over a field $\mathbf{F}$, which do not have a common component, have a total number of intersections (including their multiplicities) over the closure $\bar{\mathbf{F}}$ which is at most $\deg(g) \cdot \deg(f)$.

So the next step will be to define what is meant by the multiplicity of the intersection of a projective line and curve.

Definition 4.21. Let $L(\alpha, \beta, \gamma)$ denote a projective line and $C_g(\mathbf{F})$ a projective curve. The point $P = [a : b : c] \in L(\alpha, \beta, \gamma)$ shall be fixed and a point $P' = [a' : b' : c']$ be chosen arbitrarily from $L(\alpha, \beta, \gamma)$. It follows that the multiplicity of the intersections of $L(\alpha, \beta, \gamma)$ and $C_g(\mathbf{F})$ in P is defined as the order of vanishing at t=0:

$$\psi(t) = g(a + ta', b + tb', c + tc')$$

This will from now on be denoted by $m(P, L(\alpha, \beta, \gamma), C_g(\mathbf{F})$.

Remark. For any point on $P \notin L(\alpha, \beta, \gamma)$ the multiplicity $m(P, L(\alpha, \beta, \gamma), C_g(\mathbf{F})$ equals 0.

But how many times do a projective line and an elliptic curve intersect? Bézout's theorem tells that it happens at most three times. But does it happen never, once, twice or three times?

Proposition 4.22. *The sum of all multiplicities of and projective curve L and an elliptic curve $E(\mathbf{F})$ is denoted:*

$$\sum_{P \in \mathbf{P}^2(\mathbf{F})} m(P, L, E(\mathbf{F}))$$

and equals either 0,1 or 3.

Proof. Let $L(\alpha, \beta, \gamma)$ be a line and $g(X, Y, Z) = Y^2Z + a_1XYZ + a_3YZ^2 - X^3 - a_2X^2Z - a_4XZ^2 - a_6Z^3$ an elliptic curve in the projective plane. In case that g and L do not intersect, $m(P, L, E(\mathbf{F})) = 0$.

Case 1: $\alpha = \beta = 0$
With this conditions the point of intersection is $\mathcal{O} = [0 : 1 : 0]$. When using definition 4.21 to calculate the multiplicity a second point is needed. Choosing that point to be $[1 : 0 : 0]$ gives $g(0 + t, 1 + 0t, 0 + 0t) = \psi(t) = -t^3$. From this follows that the order of vanishing at zero equals to 3. This means that the sum of multiplicities is three.

Case 2: $\alpha \neq 0, \beta = 0$
Let L contain a point $P = [x : y : z]$. From this it follows that $\alpha x = -\gamma y$ and leaves two possibilities, either $z = 0$ or $z \neq 0$. For $z = 0$, $P = \mathcal{O} = [0 : 1 : 0]$ and for $z \neq 0$ $P = [-\frac{\gamma}{\alpha} : y_0 : 1]$ for some $y_0 \in \mathbf{F}$. Let the arbitrary point on L be equal to $[-\gamma : 0 : \alpha]$ so that $\psi(t) = g(-\gamma t, 1, \alpha t) = t(c_2t^2 + c_1t + c_0)$ for some constants $c_2, c_1, c_0 \in F$ of which $c_0 \neq 0$. It follows that $m(\mathcal{O}, L, E(\mathbf{F})) = 1$.

In case $z \neq 0$ the point P is on $E(\mathbf{F})$ if and only if y_0 is a zero of the polynomial $h(y) = g(-\frac{\gamma}{\alpha}, y, 1)$. Using again $\mathcal{O}$ as a second point to calculate the multiplicity gives $\psi(t) = h(y_0 + t)$. Setting $t = 0$ results in $h(y) = (y - y_0)^k h^*(y)$

with k being the order of the zero at y_0 of h and h^* a polynomial not equal to zero in y_0. Rewriting this to examine it better leads to:

$$\psi(t) = h(y_0 + t) = t^k h^*(y_0 + t)$$

from which it can be seen that k is also the order of vanishing in zero of ψ. So the $\sum\limits_{P \in \mathbf{P}^2(\mathbf{F})} m(P, L, E(\mathbf{F}))$ equals one plus the sum of the orders of the zeros of h in $\mathbf{F}$. Calculating:

$$\begin{aligned} h(y) &= g(-\frac{\gamma}{\alpha}, y, 1) \\ &= y^2 + a_1(-\frac{\gamma}{\alpha})y + a_3 y - (-\frac{\gamma}{\alpha})^3 - a_2(-\frac{\gamma}{\alpha})^2 - a_4(-\frac{\gamma}{\alpha}) - a_6 \end{aligned}$$

one sees that h is of degree 2. This means that h has either no zero in $\mathbf{F}$, one zero of order two or two zeros of order one in $\mathbf{F}$.

Case 3: $\beta \neq 0$
In this case, the intersection $L \cap E(\mathbf{F})$ is contained in the affine space $\mathbf{A}^2(\mathbf{F})$, because $\mathcal{O}$ can not be a point on the line L. Let $P = [x_0, y_0, 1]$ denote a point that is in $L \cap E(\mathbf{F})$, this holds if and if $y_0 = -\frac{\gamma}{\alpha} - \frac{\alpha}{\beta}x_0$ and x_0 a root of $h(x) = g(x, \frac{\gamma}{\alpha} - \frac{\alpha}{\beta}x, 1)$ is. To calculate the multiplicity for such an P one takes again an arbitrary point on the line L, for example $[-\beta, \alpha, 0]$, and evaluates the function ψ to get:

$$\begin{aligned} \psi(t) &= g(x_0 - t\beta, y_0 + t\alpha, 1) \\ &= g(x_o - t\beta, -\frac{\gamma}{\beta} - \frac{\alpha}{\beta}(x_0 - t\beta), 1) = h(x_0 - t\beta) \end{aligned}$$

Like before $m(P, L, E(\mathbf{F}))$ equals the order of the zeros x_0 in h. $h(x) = g(x, -\frac{\gamma}{\beta} - \frac{\alpha}{\beta}x, 1)$ itself is a polynomial of degree three with the highest coefficient being -1. By the definition of what the algebraic closure $\bar{\mathbf{F}}$ is, the polynomial can be split up:

$$h(x) = -(x - x_1)(x - x_2)(x - x_3)$$

all x_1, x_2 and x_3 are in $\bar{\mathbf{F}}$, not necessarily different. The sum of orders of the zeros of h in $\mathbf{F}$ therefore equals the number of x_i in $\mathbf{F}$. This number is in any case smaller or equal to three. Multiplying this out:

$$h(x) = -x^3 - (x_1 + x_2 + x_3)x^2 + (x_2 x_3 - x_1 x_2 - x_1 x_2)x + x_1 x_2 x_3$$

one sees that the coefficient of x^2 in h is $x_1 + x_2 + x_3$ which is an element in $\mathbf{F}$. Therefore the number of x_i which lie in $\mathbf{F}$ can't be equal to two. If that is the case for two x_i, then it also has to hold that the third x_i is an element of $\mathbf{F}$. This concludes the assumption.

□

Corollary 4.23. *For an elliptic curve $E(\mathbf{F})$ it holds that:*

1. *Let L be the line going through the points Q and P which lay on $E(\mathbf{F})$, then L has (counting in multiplicities) three intersections with the elliptic curve.*

2. *Let L denote the tangent line on the curve $E(\mathbf{F})$ going through the point P, then L has (counting in multiplicities) three intersections with the elliptic curve when counting P twice.*

The terminology *counting in multiplicities* means in this case that every point Q gets counted $m(Q, L, C_g(\mathbf{F}))$ times.

Proof. 1. The result from proposition 4.22 gives that:

$$\sum_{P \in \mathbf{P}^2(\mathbf{F})} m(P, L, E(\mathbf{F})) = 3$$

This leads to two options. Option one is that there is a point $R \in L \cap E(\mathbf{F})$ which is neither Q nor P, then all three points P, Q, R have multiplicity one and R is the additional intersection point. Option two if one of the points P, Q has multiplicity two and therefore the other point has to have multiplicity one.

2. It is clear that P must have a multiplicity of at least two. From proposition 4.22 it can be concluded that there is either a second point $Q \in L \cap E(\mathbf{F}) \neq P$ which has multiplicity one or that P has multiplicity three. If P has multiplicity three, the third intersection is again the point P, if not then Q is the third intersection of the line and the elliptic curve.

□

Now that curves and lines got formally introduced and their interaction examined, a group law can be defined under which the points on an elliptic curve form an abelian group.

Definition 4.24 (Addition law for elliptic curves)**.** As before, let $E(\mathbf{F})$ be an elliptic curve and P and Q two points on the curve. Then the addition of those two points will be denoted by $P \oplus Q$ and carried out in the following way:

Draw a line L_1 through P and Q. The intersection of this line with $E(\mathbf{F})$ is guaranteed by corollary 4.23 and then denoted $P * Q$. Now a second line L_2 is drawn intersecting $P*Q$ and $\mathcal{O} = [0 : 1 : 0]$ and then also $E(\mathbf{F})$. The intersection with $E(\mathbf{F})$ is then $P \oplus Q$. In the case that $P * Q = \mathcal{O}$ the tangent on $E(\mathbf{F})$ in $\mathcal{O}$ is chosen to be L_2

In a similar manner the doubling of a point $P \oplus P$ on $E(\mathbf{F})$ is performed. The line L_1 is the tangent line in P on $E(\mathbf{F})$ and the third intersection of L_1

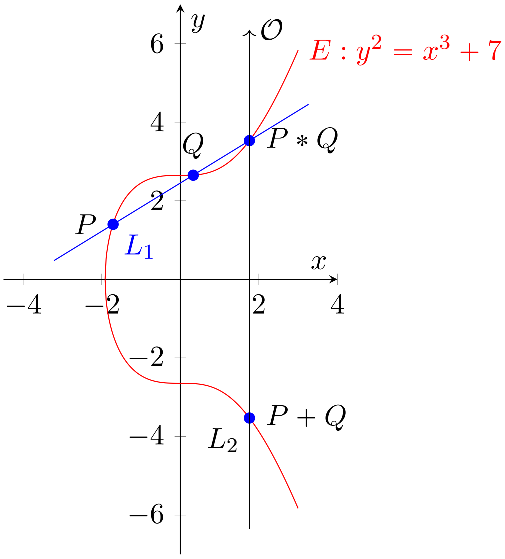

Figure 1: Addition two points P and Q over $\mathbf{R}$

with $E(\mathbf{F})$ is denoted $P * P$. Drawing again the line L_2 going through $P * P$ and $\mathcal{O}$ the third intersection of L_2 with $E(\mathbf{F})$ is then the point $P \oplus P$.

This procedure can be visualized very well when $\mathbf{F} = \mathbf{R}$ as seen in figure 1.

Lemma 4.25. *Let P, Q, R be three points on an elliptic curve $E(\mathbf{F})$. If all lay on the same projective line L, then it holds that*

$$(P \oplus Q) \oplus R = \mathcal{O}$$

This holds also if some or all of P, Q, R coincide but only if they occur as often as it matches their multiplicity.

Proof. When calculating the addition of P and Q one gets the third intersection of the line L_1 with the elliptic curve $E(\mathbf{F})$ at the point R. From this, it follows that the intersection of the elliptic curve with a line L_2 going through R and $\mathcal{O}$ equals $Q \oplus P$. To add R to $Q \oplus P$ one draws a line L_1' through $Q \oplus P$ and R. This line L_1' is the same as L_2, therefore the third intersection of the line with the curve must be $\mathcal{O}$. Inspecting the tangent line L_2' at $\mathcal{O}$ on $E(\mathbf{F})$ one sees that the third point of intersection is again $\mathcal{O}$. This proofs the lemma. □

As mentioned before, a group can be formed from the points on an elliptic curve. In the next proposition, this is formulated and proven.

Proposition 4.26. *Let $E(\mathbf{F})$ be an elliptic curve. Under the previously defined addition law (Def 4.24), the points on the curve and the neutral element $\mathcal{O}$ form an abelian group.*

1. *For every $P, Q \in E(\mathbf{F})$ also $P \oplus Q$ is an element of $E(\mathbf{F})$. (Closure)*

2. *$P \oplus \mathcal{O} = P$ for all $P \in E(\mathbf{F})$. (Identity element)*

3. *There exists a point $-P$ for all $P \in E(\mathbf{F})$ such that $P \oplus (-P) = \mathcal{O}$. (Inverse element)*

4. *$P \oplus Q = Q + P$ for all $P, Q \in E(\mathbf{F})$. (Commutativity)*

5. *$(P \oplus Q) + R = P \oplus (Q \oplus R)$ for all $P, Q, R \in E(\mathbf{F})$. (Associativity)*

Proof. Let P, Q, R be points on the elliptic curve $E(\mathbf{F})$ given by a Weierstraß equation. Let furthermore L denote a projective line.

Closure

The closure is given by the definition of the group law and lemma 4.25.

Identity Element

To show that $\mathcal{O}$ is the identity element, two cases have to be examined, first $P = \mathcal{O}$ and then for P an arbitrary point other than $\mathcal{O}$.

Case 1: $P = \mathcal{O}$
Let $P = \mathcal{O}$, then one needs to calculate the tangent in the point and gets:

$$\frac{\partial g}{\partial X}(0,1,0) = 0 \quad \frac{\partial g}{\partial Y}(0,1,0) = 0 \quad \frac{\partial g}{\partial Z}(0,1,0) = 1$$

The result is the line $L(0,0,1)$ given by the equation $Z = 0$. But this line is not in the affine space $\mathbf{A}^2(\mathbf{F})$ and therefore can only intersect the elliptic curve in $\mathcal{O}$. The third intersection (counting in multiplicities) denoted $\mathcal{O} * \mathcal{O}$, is again $\mathcal{O}$. Following the procedure of the group law, putting another tangent through $\mathcal{O}$ gives that the third intersection with $E(\mathbf{F})$ is $\mathcal{O}$. So it follows that $\mathcal{O} \oplus \mathcal{O} = \mathcal{O}$.

Case 2: $P \neq \mathcal{O}$
Let $P \neq \mathcal{O}$ be an arbitrary point on the elliptic curve $E(\mathbf{F})$. Adding the point $\mathcal{O}$ translates to drawing a projective line L_1 through P and $\mathcal{O}$. The resulting third intersection with $E(\mathbf{F})$ is then denoted by $P * \mathcal{O}$. To construct then $P \oplus \mathcal{O}$ one has to draw a line L_2 through $\mathcal{O}$ and $P * \mathcal{O}$. This line L_2 is the same line as L_1. Therefore the third intersection of L_2 and $E(\mathbf{F})$ has to be P.

From this, it follows that $\mathcal{O}$ is the neutral element of the group.

Inverse Element

Let $\ominus P$ denote the third intersection of the line L going through P and $\mathcal{O}$. By the definition, the point $\ominus P$ is an element of $E(\mathbf{F})$ which is on the same line as P and $\mathcal{O}$. By lemma 4.25 it then holds that:

$$\mathcal{O} = (P \oplus \mathcal{O}) \oplus (\ominus P) = P \oplus (\ominus P)$$

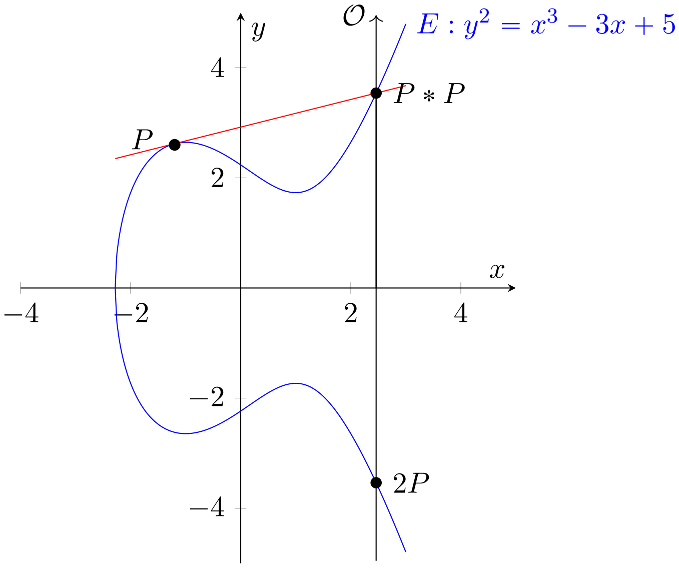

Figure 2: Point duplication over $\mathbf{R}$

Commutativity

When drawing the line through P and Q to obtain the point $P * Q$, one also obtains the point $Q * P$ because the line through P and Q is unique. Therefore $P \oplus Q = Q \oplus P$

Associativity

This part of the theorem is the most tedious and would extend this work with several pages just dedicated to this proof. Therefore the reader is advised to consult [Kna92] Chapter 3 if interested. □

Remark. From this point on the symbol $\oplus$ will be replaced by an ordinary $+$ when the addition law is used on elliptic curves.

Remark. When a point is added several times to itself the following notation will be used:

$$\underbrace{P + P + \cdots + P}_{\text{n times}} = nP$$

In books, it is sometimes referred to as *scalar multiplication.*

One can also derive formulas for direct calculations. Point duplication as seen in figure 2 is of special interest for public-key cryptographic.

Proposition 4.27. *Let* $P = [x_1 : y_1 : 1]$ *be a point on an elliptic curve* $g(X, Y, Z) = Y^2Z + a_1XYZ + a_3YZ^2 - X^3 - a_2X^2Z - a_4XZ^2 - a_6Z^3$*. Then* *the coordinates of the point* $2P$ *are given by* $x_3 = \lambda^2 + a_1\lambda - a_2 - 2x_1$ *and* $y_3 = -(\lambda + a_1)x_3 - \nu - a_3$. *With:*

$$\lambda = \frac{3x_1^2 + 2a_2x_1 + a_4 - a_1y_1}{2y_1 + a_1x_1 + a_3}$$
$$\nu = \frac{-x_1^3 + a_4x_1 + 2a_6 - a_3y_1}{2y_1 + a_1x_1 + a_3}$$

Proof. Let tangent line in P be $L = L(\lambda', \mu', \nu')$ with λ', μ' and ν' in P are given by the derivatives in the point:

$$\begin{aligned}
\lambda' &= \frac{\partial g}{\partial X}(x_1, y_1, 1) = a_1y_1 - 3x_1^2 - 2a_2x_1 - a_4 \\
\mu' &= \frac{\partial g}{\partial Y}(x_1, y_1, 1) = 2y_1 + a_1x_1 + a_3 \\
\nu' &= \frac{\partial g}{\partial Z}(x_1, y_1, 1) = y_1^2 + a_1x_1y_1 + 2a_3y_1 - a_2x_1^2 - 2a_4x_1 - 3a_6
\end{aligned}$$

Case 1: $\mu' = 0$
In case that $\mu' = 0$ the point $\mathcal{O}$ is on the line. This implies that $P + P = \mathcal{O}$ and therefore $P = -P$.

Case 2: $\mu' \neq 0$

$$\begin{aligned}
\lambda = -\frac{\lambda'}{\mu'} &= \frac{3x_1^2 + 2a_2x_1 + a_4 - a_1y_1}{2y_1 + a_1x_1 + a_3} \\
\nu = -\frac{\nu'}{\mu'} &= \frac{-y_1^2 - a_1x_1y_1 - 2a_3y_1 + a_2x_1^2 + 2a_4x_1 + 3a_6}{2y_1 + a_1x_1 + a_3} \\
&= \frac{-x_1^3 + a_4x_1 + 2a_6 - a_3y_1}{2y_1 + a_1x_1 + a_3}
\end{aligned}$$

Using now the affine form of the Weierstraß equation $f(x, y) = 0$ and substituting $y = \lambda x + \nu$ into it, gives the following equation:

$$(\lambda x + \nu)^2 + a_1x(\lambda x + \nu) + a_3(\lambda x + \nu) - x^3 - a_2x^2 - a_4x - a_6 = 0$$

which every point that lies on both E and L has to satisfy. One point satisfying this equation, namely x_1, is already known. This makes it possible to rewrite the equation to $c(x - x_1)(x - x_2')(x - x_3')$ with $x_2', x_3' \in \bar{\mathbf{F}}$ and $c \in \mathbf{F}$. Comparing the coefficients of both equations gives that $c = -1$ and $\lambda^2 + a_1\lambda - a_2 = x_1 + x_2' + x_3$ From the fact that L is a tangent in P it is clear that x_1 is a double root, so either $x_2' = x_1$ or $x_3' = x_1$. Rearranging the equation gives that $x_3' = \lambda^2 + a_1\lambda - a_2 - 2x_1$. This means that x_3' is also a solution im $\mathbf{F}$. The third intersection of L and $E(\mathbf{F})$ is then $P' = (x_3', y_3')$ with $y_3' = \lambda x_3' + \nu$. If $P_3' = P_1$, the polynomial $-(x - x_1)(x - x_2')(x - x_3')$ can be rewritten to $-(x - x_1)^3$ which means that the intersection has multiplicity three at P. For the case $P_1 \neq 2P'$ it follows that

$-(P+P) = P'$. To get the coordinate y_3 one sets $x = x_3$ into the Weierstraß equation and receives a quadratic equation in y in the form $y^2 + cy + d = 0$. $c = a_1x_3 + a_3$ and $d = -x_3^3 - ax_2x_3^2 - a_4x_3 - a_6$ lie in $\mathbf{F}$. Therefore this equation has two solution in the closure $\bar{\mathbf{F}}$. One of the solutions, namely y_3' is already known. So the equation factors to:

$$y^2 + cy + d = (y - y_3)(y - y_3')$$

where y_3 lies in $\bar{\mathbf{F}}$. Multiplying out both sides and comparing the coefficients reveals:

$$y_3 = -y_3' - c = -y_3' - a_1x_3 - a_3$$

The result is then $x_3 = \lambda^2 + a_1\lambda - a_2 - 2x_1$ and $y_3 = -(\lambda + a_1)x_3 - \nu - a_3$ where λ and ν are defined as before.

□

The next proposition applies to field with a characteristic bigger than three. This condition makes it possible to transform the Weierstraß equation into the simplified Weierstraß equation, which makes the derivation of the formulas shorter. For fields of characteristic two and three, this process can be adapted and applied to the normal Weierstraß equation.

Proposition 4.28. *Let $P = [x_1, : y_1 : 1]$ be a point on an elliptic curve $g(X, Y, Z) = Y^2Z - X^3 - a_4XZ^2 - a_6Z^3$*

1. *For a $P = (x_1, y_1) \in C_f(\mathbf{F})$ the point $-P$ equals $(x_1, -y_1)$*
2. *For $P_1 = (x_1, y_1), P_2 = (x_2, y_2)$ both in $C_f(\mathbf{F})$ and $x_1 \neq x_2$ their sum $P_1 + P_2 = P_3 = (x_3, y_3)$ with:*

$$x_3 = \lambda^2 - x_1 - x_2$$
$$y_3 = \lambda(x_1 - x_3) - y_1$$

where λ is given by:

$$\lambda = \frac{y_2 - y_1}{x_2 - x_1}$$

Proof. 1. The point $(-P)$ is on the line going through P and $\mathcal{O}$. This line is vertical through P, so P and $-P$ have to have the same x-coordinate x_1. Examining the equation $y^2 = x_1^3 - a_4x_1 - a_6$ shows that there are only two possible values, y_1 and $-y_1$. P already has the coordinates (x_1, y_1), so $-P$ has to be $(x_1, -y_1)$.

2. Let L be the line connecting P_1 and P_2. Then the line parameters λ', ν' and μ' are unknown at first. The points on the line have to fulfil the line equation:

$$\lambda' x + \mu' y + \nu' = 0$$

In case that $x_1 = x_2$ the value μ' equals zero and the line connecting both points also intersects $\mathcal{O}$ which indicates that $P_2 = -P_1$. With $\lambda = \frac{\lambda'}{\mu'}$ and $\nu = \frac{\nu'}{\mu'}$ the line equation can be rewritten to:

$$y = \lambda x + \nu$$

So the y-coordinates of P_1 and P_2 are $y_1 = \lambda x_1 + \nu$ and $y_2 = \lambda x_2 + \nu$ respectively. Combining the formulas gives:

$$\begin{aligned} \lambda(x_2 - x_1) &= y_2 - y_1 \\ \lambda &= \frac{y_2 - y_1}{x_2 - x_1} \end{aligned}$$

Further more this leads to a formula to calculate ν:

$$\begin{aligned} \nu = y_1 - \lambda x_1 &= y_1 - \frac{y_2 - y_1}{x_2 - x_1} x_1 \\ &= \frac{y_1(x_2 - x_1) - x_1(y_2 - y_1)}{x_2 - x_1} \\ &= \frac{y_1 x_2 - y_2 x_1}{x_2 - x_1} \end{aligned}$$

Like in the proof for the duplication formula, $y = \lambda x + \nu$ is now put into the Weierstraß equation to obtain the point of intersection.

$$-x^3 + \lambda^2 x^2 + (2\lambda\nu - a_4)x + \nu^2 - a_6 = 0$$

Rewriting the polynomial of degree three, of which two solutions x_1, x_2 are already known, gives the following over the algebraic closure $\bar{\mathbf{F}}$:

$$c(x - x_1)(x - x_2)(x - x') = 0 \qquad c \in \mathbf{F}, x' \in \bar{\mathbf{F}}$$

Comparing both forms of the polynomial gives to $c = -1$ and $\lambda^2 = x_1 + x_2 + x'$ and leads to $x' = \lambda^2 - x_1 - x_2$ which is an element if $\mathbf{F}$. The set of intersections $\mathbf{A}^2(\mathbf{F}) \cap L \cap E(\mathbf{F})$ then contains P_1, P_2 and $P' = (x', \lambda x' + \nu)$. If $P' \neq P_1 \wedge P_2$ then $P' = -(P_1 + P_2)$, otherwise if $P' = P_1 \vee P_2$ the multiplicity of the point has to be examined. Let $P' = P_1$, then the multiplicity of P_1 in $L \cap E(\mathbf{F}$ equals the order of vanishing of x_1 like in the proof of proposition 4.22. The order of vanishing is then two which implies that $x_1 = x'$ and therefore $P' = -(P_1 + P_2)$ holds. The same holds for the case $P' = P_2$. Inverting the y coordinates of P' and replacing ν with $y_1 - \lambda x_1$ gives then P_3 with:

$$x_3 = \lambda^2 - x_1 - x_2$$
$$y_3 = \lambda(x_1 - x_3) - y_1$$

with λ as defined previously in the proof.

□

A theorem not necessary for elliptic curve cryptography, but still remarkable is Mordell's theorem for elliptic curves over the rational numbers:

Theorem 4.29 (Mordell)**.** *Let E denote a non-singular cubic curve with rational coefficients and a rational point, then the group of rational points is finitely generated.*

Proof. A proof can be found in [Sil09] □

4.4 Elliptic Curves Over Finite Fields

After the theory for elliptic curves over arbitrary fields was introduced before, this part will be a brief introduction to elliptic curves over finite fields. First the procedure of point addition on curves over finite fields will be presented, followed by a short discussion of determining the group order.

Until now no calculation example for the formulas in proposition 4.28 was given. In converse to most standard literature about elliptic curves and elliptic curve cryptography, a visualization of point addition over a finite field is shown.

Example 4. Consider the function from example 2 and its set of solutions $C_f(\mathbf{F}_5) = \{(0,1), (0,4), (2,1), (2,4), (3,1), (3,4), (4,2), (4,3)\}$. The task is to add the point $Q = (2,1)$ and $P = (4,2)$ by drawing lines in the plane and then checking the result with the formulas.

Drawing a line in the plane through Q and P is an easy task to do, but the line ends in at $x = 4$. This seems like a contradiction to the before developed theory as there has to be a third intersection with the elliptic curve. Recalling that this is done over a finite field, the procedure has to be adapted to the properties of a finite field. After the line ends at $x = 4$ the next integer x-coordinate one would expect is $x = 5$ which is congruent to $0 \pmod 5$. So the line continues at $x = 0$, and has the equation $y = 3x \pmod 5$. Evaluating the line equation at an arbitrary point in $\mathbf{F}_5$, for example $x = 1$ gives the y value of three which is on the line. So drawing a line through $(1,3)$ with slope three gives the second line. If, for example in a bigger field, there was still no intersection, one has to continue drawing lines that are parallel until an intersection is found.

The last step is drawing a line through $P * Q$ and $\mathcal{O}$ which gives then the intersection with the curve at the point $Q + P = (3,1)$.

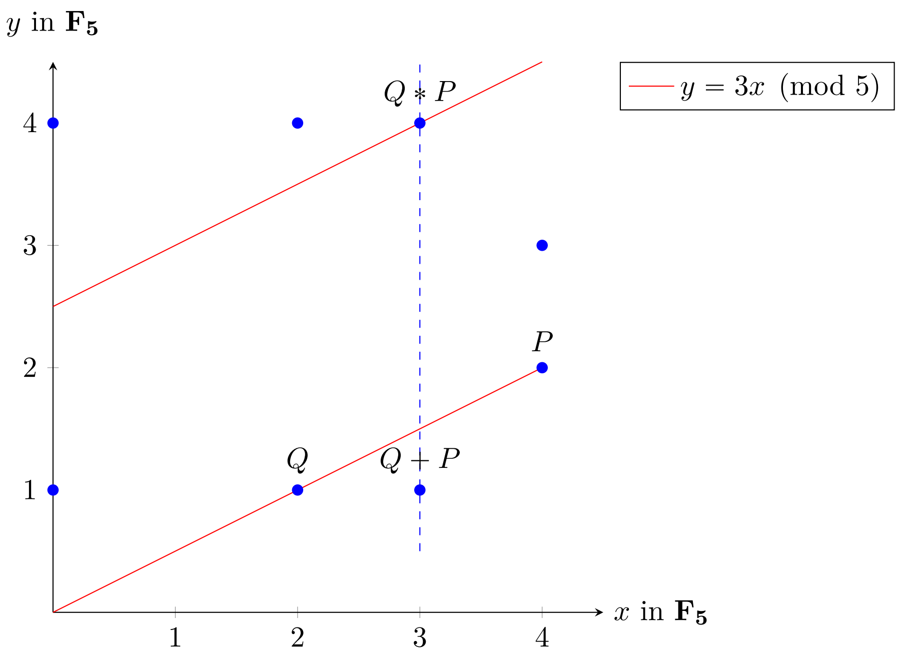

Figure 3: $y^2 = x^3 + x + 1$ over $\mathbf{Z}_5$

Let $Q + P = R$, then the formulas give:

$$\begin{aligned}
\lambda &= \frac{y_q - y_p}{x_q - x_p} = \frac{1-2}{2-4} = \frac{-1}{-2} = \frac{4}{3} = 4 \cdot 3^{-1} = 8 = 3 \pmod 5 \\
x_r &= \lambda^2 - x_p - x_q = 9 - 4 - 2 = 3 \pmod 5 \\
y_r &= \lambda(x_p - x_r) - y_p = (3(4-3) - 2) = 1 \pmod 5
\end{aligned}$$

So both times one gets at $Q + P = (3, 1)$.

For applying elliptic curves on cryptography problems it becomes important to have an idea of how many points there are on a curve. For curves over the real numbers this can be easily answered with infinity, but how many points are there over a finite field? Those points are elements of the group and therefore determine the order of the group. A first approximation on how many points there are on an elliptic curve over a finite field is given by the following theorem, which was conjectured by Emil Artin and proven by Helmut Hasse in the 1930s:

Theorem 4.30 (Hasse's theorem on elliptic curves)**.** *Let E be a non-singular elliptic curve defined over the finite field $\mathbf{F}_q$. Then the number of points on E which are contained in $\mathbf{F}_q$ equals $p + 1 - \epsilon$, where ϵ denotes and* error term *with the property $|\epsilon| \leqslant 2\sqrt{q}$.*

$$-2\sqrt{q} \leqslant \#E(\mathbf{F}_q) - q - 1 \leqslant 2\sqrt{q}$$

A proof of this theorem would require too much theoretical background. The interested reader is advised to consult [Sil09].

But the bound is not accurate enough to use it for cryptography. In section 3.1 the issue with subgroups of small order has been mentioned. This is not only a problem with finite fields, but also for elliptic curves over finite fields. One can try out all combinations of example 2 or 4 with pen and paper without investing huge amounts of time. Blake and Smart [Bla+99] have a chapter on determining the order of a group from which the following idea of an *naive approach* is taken from.

Theorem 4.31. *Let $f = y^2 - x^3 + ax + b$ be an elliptic curve over a finite field with characteristic p. The sum of all rational points on the curve is given by:*

$$\#E(\mathbf{F}_p) = p + 1 + \sum_{x=0}^{p-1} \left(\frac{x^3 + ax + b}{p} \right)$$

Proof. Assuming that for every element x in $\mathbf{F}_p$ plus $\mathcal{O}$ there is a pair (x, y) which is a solution to the equation of the elliptic curve, one needs to start with $p+1$ points. The equation of the elliptic curve is quadratic in y which means that for every solution (x, y), there is a second point $(-x, y)$ which is also a solution. Having all possible points already counted once, one only needs to add a point if $x^3 + ax + b$ is a quadratic residue and subtract one point if $x^3 + ax + b$ is a quadratic nonresidue. □

This is of course not a very practical approach to the problem of finding the group order. There are other methods like *Schoof's Algorithm* and the *Schoof-Elkies-Atkin Algorithm* which are out of the scope of this work but necessary to mention.

5 Elliptic curve Diffie-Hellman Key Exchange

Definition 5.1. Let E be an elliptic curve over the field $\mathbf{F}_p$, with $P, Q \in E(\mathbf{F}_p)$ under the group law defined before. Then n is called a solution to the *Elliptic Curve Discrete Logarithm Problem* if:

$$\begin{aligned} nP &= Q \\ n &= \log_P(Q) \end{aligned}$$

The Diffie-Hellman problem on elliptic curves is similar but needs more public parameters. While for classical Diffie-Hellman key exchange over finite fields, Alice and Bob only had to agree on a field and a primitive root, now they also have to agree on the parameters of a curve.

Definition 5.2. After agreeing on a field $\mathbf{F}$, an elliptic curve $E(\mathbf{F})$ and a primitive root P, they calculate their Q_A and Q_B respectively.

$$\begin{aligned} Q_A &= n_A P \\ Q_B &= n_B P \end{aligned}$$

As before, both participants again exchange those values publicly over the insecure channel and repeat the previous procedure:

$$\begin{aligned} Q_{AB} &= n_A Q_B = n_A n_B P \\ Q_{AB} &= n_B Q_A = n_A n_B P \end{aligned}$$

The following example is taken from [HPS14]:

Example 5. Let $p = 3851$ and the curve $E : Y^2 = X^3 + 324X + 1287$ with the generator $P = (920, 303) \in E(\mathbf{F}_{3851})$. Let Alice choose her $n_A = 1194$ and Bob his to be $n_b = 1759$. Each of them calculates their value Q_A and Q_B and sends it to the other person:

$$\begin{aligned} n_A P = Q_A &= 1194P = (2067, 2178) \\ n_B P = Q_B &= 1759P = (3684, 3125) \end{aligned}$$

If both use their secret values to compute the exchanged secret both arrive at the following:

$$\begin{aligned} n_A Q_B &= 1194(3684, 3125) = (3347, 1242) \\ n_B Q_A &= 1759(2067, 2178) = (3347, 1242) \end{aligned}$$

5.1 An Algorithm To Break The ECDLP

Notes On Index Calculus For Elliptic Curves

Index calculus and elliptic curve cryptography is a difficult topic. In [SS98; Mil86b] Miller and Silverman present arguments why index calculus should not work for the ECDLP. Silverman even presented in [Sil00] an alternative algorithm to the index calculus approach for elliptic curves But Claus Diem presents an index calculus algorithm over finite extension fields [Die11], as well as Gary McGuire and Daniela Mueller in [MM17].

What all of those approaches have in common, is their exponential time complexity, which makes them uninteresting for practical attacks on the ECDLP.

The Pohlig-Hellman algorithm

With the *Pohlig-Hellman* algorithm, it is possible to reduce the complexity of the DLP in a cyclic group of composite order to the DLP in several cyclic groups of prime power order [HPS14].

In order to solve the set of smaller problems, one needs another method to solve the DLP in those cyclic groups and later combines the partial solutions with the Chinese Remainder theorem. If the group order is prime, the Pohlig-Hellman algorithm gives no advantage over other methods.

Theorem 5.3 (The Pohlig-Hellman algorithm)**.** *Let G denote an arbitrary group which contains an element g of order N. Let N factor into prime powers such that:*

$$N = q_1^{e_1} q_2^{e_2} \dots q_t^{e_t}$$

Then the following procedure solves the discrete logarithm problem in G.

1. *Factor N into its prime power factorization*

2. *Compute all values $g_i = g^{N/q_i^{e_i}}$ and $h_i = h^{N/q_i^{e_i}}$ for $1 \leqslant i \leqslant t$ and solve the discrete logarithm problem*

$$g_i^y = h_i$$

where y_i denotes the solution of each individual problem.

3. *With the Chinese remainder theorem the system of congruences can be solved:*

$$x \equiv y_1 \pmod{q_1^{e_1}},\ x \equiv y_2 \pmod{q_2^{e_2}},\ \dots,\ x \equiv y_t \pmod{q_t^{e_1 t}}$$

giving the final solution.

Proof. Let x denote the solution to the system of congruences. Then for each i the solution can be written as

$$x = y_i + q_i^{e_i} z_i \quad \text{for some } z_i$$

This then leads to:

$$\begin{aligned}
(g^x)^{\frac{N}{q_i^{e_i}}} &= (g^{y_i + q_i^{e_i} z_i})^{\frac{N}{q_i^{e_i}}} \\
&= (g^{\frac{N}{q_i^{e_i}}})^{y_i} \cdot g^{N z_i} \\
&= (g^{\frac{N}{q_i^{e_i}}})^{y_i} \\
&= g_i^{y_i} = h_i = h^{\frac{N}{q_i^{e_i}}}
\end{aligned}$$

This can be rewritten to a discrete logarithm with basis g:

$$\frac{N}{q_i^{e_i}} \cdot x \equiv \frac{N}{q_i^{e_i}} \cdot \log_g(h) \pmod{N}$$

The discrete logarithm to the basis g is defined $\bmod N$ because g^N is the identity element of the group. The next step is to observe that $\frac{N}{q_i^{e_i}}$ for $i = 1, \dots, t$ are all coprime. When using the extended Euclidean Algorithm repeatedly, it is possible to find c_i such that:

$$\sum_{i=1}^{t} \frac{N}{q_i^{e_i}} \cdot c_i = 1$$

This makes it then possible to add the congruences multiplies by their c_i respectively and sum them up:

$$\sum_{i=1}^{t} \frac{N}{q_i^{e_i}} \cdot c_i \cdot x \equiv \sum_{i=1}^{t} \frac{N}{q_i^{e_i}} \cdot c_i \cdot \log_g(h) \pmod{N}$$

Which then collapses to:

$$x \equiv log_g(h) \pmod{N}$$

and thereby completes the proof. □

The following example shows the Pohlig-Hellman algorithm applied to the DLP over finite fields. It's taken from [HPS14].

Example 6 (Finite fields)**.** Let the characteristic of the field be $p = 11251$ and take the primitive root $g = 23$. Then goal is to find the exponent x for which $g \equiv h = 9689 \pmod{11251}$. The order of the group is $N = p - 1 = 11250$. 11250 factors into $2 \cdot 3^2 \cdot 5^4$. Those give $N/2 = 5625, N/3^2 = 1250$ and $N/5^4 = 18$ as exponents

$$\begin{aligned} 23^{5625x} &\equiv 9689^{5625} \pmod{11251} \\ 11250^x &\equiv 11250 \pmod{11251} \\ x &= 1 \pmod 2 \end{aligned}$$

$$\begin{aligned} 23^{1250x} &\equiv 9689^{1250} \pmod{11251} \\ 5029^x &\equiv 10724 \pmod{11251} \\ x &= 4 \pmod{3^2} \end{aligned}$$

$$\begin{aligned} 23^{18x} &\equiv 9689^{18} \pmod{11251} \\ 5448^x &\equiv 6909 \pmod{11251} \\ x &= 511 \pmod{5^4} \end{aligned}$$

Solving the DLP over $\pmod 2$ and $\pmod{3^2}$ is easy. Finding the solution $\pmod{5^4}$ requires more work, but significantly less than finding a solution $\pmod{11251}$. There exist algorithms to solve this faster than with an exhaustive search, for example in [HPS14].

The next step is to solve the system of congruences using the Chinese Remainder Theorem:

$$x \equiv 1 \pmod 2, \quad x \equiv 4 \pmod{3^2}, \quad x \equiv 511 \pmod{5^4}$$

This gives $x = 4261$ as the smallest solution. Checking the answer shows that $23^{4261} \equiv 9689 \pmod{11251}$.

The ECDLP can be solved with the same procedure. The following example for elliptic curves is taken from [Bla+99].

Example 7 (Elliptic curves)**.** Let $P = (1, 237)$ and $Q = (190, 271)$ be points on the elliptic curve $E : Y^2 = X^3 + 71X + 602$ defined over the finite field $\mathbf{F_{1009}}$. $E(\mathbf{F_{1009}})$ has the group order 1060, which factors to $2^2 \cdot 5 \cdot 53$. The solution to the problem $Q = mP$ can then be reduced to calculation m mod $2^2, 5$ and 53. The point P has order 530, therefore it is sufficient to calculate the solution modulo 2 instead of modulo 4. Starting with the points in the subgroup of order 2:

$$\begin{aligned} P_2 &= 265P = (50, 0) \\ Q_2 &= 256Q = (50, 0) \end{aligned}$$

from this one sees that $Q_2 = (m \pmod 2)P$ which leads to $m \equiv 1 \pmod 2$. Now the points have to be multiplied by $530/5 = 106$, which gives:

$$\begin{aligned} P_5 &= 106P = (639, 160) \\ Q_5 &= 106Q = (639, 849) \end{aligned}$$

Q_5 and P_5 have the same x-coordinate and therefore it holds that $P_5 = -Q_5$. Which shows that $m \equiv 4 \pmod 5$. The last congruence equation to be solved for is $\pmod{53}$. Here the points have to be multiplied by ten as $530/53 = 10$.

$$\begin{aligned} P_{53} &= 10P = (32, 737) \\ Q_{53} &= 10Q = (592, 97) \end{aligned}$$

By, for example exhaustive search, one gets that $Q_{53} = -5P_{53}$ and therefore $m \equiv -5 \equiv 48 \pmod{53}$. This three congruences will then be combined to the following system:

$$\begin{aligned} x &\equiv 1 \pmod 2 \\ x &\equiv 4 \pmod 5 \\ x &\equiv 48 \pmod{53} \end{aligned}$$

Using the Chinese Remainder Theorem to find a solution gives that $x = 419$. And in fact $Q = 419P$.

Theorem 5.4. *The discrete logarithm problem can be solved in*

$$\mathcal{O}\left(\sum_{i=1}^{r} S_{q_i^{e_i}} + r\log^2(q_r)\right) \text{ steps} \qquad \text{with } q_r^{e_r} \text{ the biggest divisor of } N$$

when the Pohlig-Hellman algorithm is used.

Remark. $S_{q_i^{e_i}}$ is a place holder for the number of steps an algorithm would take to solve the discrete logarithm $\pmod{q_i^{e_i}}$. In general on cane use the *Baby-Step-Giant-Step* algorithm which has a time complexity of $\mathcal{O}(\sqrt{q_r^{e_r}})$.

Proof. The first step of the algorithm is to factorize the number N to its prime power factorization $q_1^{e_1}q_2^{e_2}\dots q_r^{e_r}$ where $q_1^{e_1} < p_2^{e_2} < \dots < q_r^{e_r}$. In this proof the factorization seen as given as it is out of scope of this work to talk about factoring algorithms. But it can be bounded by approximately $e^{(1+o(1))\sqrt{\ln N \ln\ln N}} = L_N[1/2, 1]$ when using the Quadratic Sieve.

Computing the different values of $g_i = g^{N/q_i^{e_i}}$ and $h_i = h^{N/q_i e_i}$ has the cost of $\sum_{i=1}^{r} \mathcal{O}(\log(N/q_i^{e_i})\log^2(N))$ bit operations where the calculation of $g_1 = g^{N/q_1^{e_1}}$ dominates this step, hence the bound $\mathcal{O}(\log(N/q_1^{e_1})\log^2(N))$.

It is clear that each DLP takes $S_{q_i^{e_i}}$ steps to be solved, which sums to $\sum_{i=1}^{r} S_{q_i^{e_i}}$.

When all DLPs are solved, it remains to find the solution by using the Chinese Remainder Theorem. Using a classical schoolbook approach for this means that one calculates:

$$x = \sum_{i=1}^{r} a_i N_i N_i^{-1} \pmod{q_i^{e_i}}$$

This is two multiplications ($\mathcal{O}(\log^2(q_i^{e_i}))$ and one inversion ($\mathcal{O}(\log^2(q_i^{e_i}))$ for each partial solution. Finally adding all of those numbers up has to be done r times. Leading to $r \cdot \mathcal{O}(\log^2(q_r^{e_r}))$, which can be bound by $\mathcal{O}(r \log^2(q_r^{e_r}))$. This gives a total expected running time of $\mathcal{O}\left(\sum_{i=1}^{r} S_{q_i^{e_i}} + r \log^2(q_r^{e_r})\right)$. □

This result shows that point counting on elliptic curves is important for cryptography. If $\#E(\mathbf{F}_q)$ is not a prime, or even worse B-smooth for small B, the DLP is fast to compute.

Bibliography

[Ben+08] Jens Bender et al. “Sicherheitsmechanismen für kontaktlose Chips im deutschen elektronischen Personalausweis”. In: *Datenschutz Und Datensicherheit - Dud* 32 (Mar. 2008), pp. 173–177. DOI: `10.1007/s11623-008-0026-7`.

[Bla+99] I. Blake et al. *Elliptic Curves in Cryptography.* Lecture note series. Cambridge University Press, 1999. ISBN: 9780521653749.

[Bur11] D.M. Burton. *Elementary Number Theory.* Mcgraw-Hill, 2011. ISBN: 9780077418120.

[Die11] Claus Diem. “On the discrete logarithm problem in elliptic curves”. In: *Compositio Mathematica* 147.1 (2011), pp. 75–104. DOI: `10.1112/S0010437X10005075`.

[Eng12] A. Enge. *Elliptic Curves and Their Applications to Cryptography: An Introduction.* Springer US, 2012. ISBN: 9781461552079.

[FSK10] Niels Ferguson, Bruce Schneier, and Tadayoshi Kohno. *Cryptography Engineering: Design Principles and Practical Applications.* Wiley Publishing, 2010. ISBN: 0470474246.

[Gal12] Steven D. Galbraith. *Mathematics of Public Key Cryptography.* 1st. USA: Cambridge University Press, 2012. ISBN: 1107013925.

[Gra08] Andrew Granville. “Smooth numbers: Computational number theory and beyond”. In: *Math. Sci. Res. Inst. Publ.* 44 (Jan. 2008).

[HPS14] J. Hoffstein, J. Pipher, and J.H. Silverman. *An Introduction to Mathematical Cryptography.* Undergraduate Texts in Mathematics. Springer New York, 2014. ISBN: 9781493917112.

[Kna92] A.W. Knapp. *Elliptic Curves.* Mathematical Notes - Princeton University Press. Princeton University Press, 1992. ISBN: 9780691085593.

[Kob08] Neal Koblitz. “Cryptography”. In: *Random Curves: Journeys of a Mathematician.* Berlin, Heidelberg: Springer Berlin Heidelberg, 2008, pp. 297–329. ISBN: 978-3-540-74078-0. DOI: `10.1007/978-3-540-74078-0_14`. URL: `https://doi.org/10.1007/978-3-540-74078-0_14`.

[Kob87] Neal Koblitz. “Elliptic Curve Cryptosystems”. In: *Mathematics of Computation* 48.177 (1987), pp. 203–209. ISSN: 00255718, 10886842. URL: `http://www.jstor.org/stable/2007884`.

[Lan78] S. Lang. *Elliptic Curves: Diophantine Analysis.* Grundlehren der mathematischen Wissenschaften. Springer-Verlag, 1978. ISBN: 9783540084891.

[Len86] H.W. Lenstra. “Elliptic Curves and Number-Theoretic Algorithms”. In: *Report 86-19, Mathematisch Instituut Amsterdam (1986)* (Jan. 1986).

[Mil86a] Victor S. Miller. “Use of Elliptic Curves in Cryptography”. In: *Advances in Cryptology — CRYPTO ’85 Proceedings*. Ed. by Hugh C. Williams. Berlin, Heidelberg: Springer Berlin Heidelberg, 1986, pp. 417–426. ISBN: 978-3-540-39799-1.

[Mil86b] Victor S. Miller. “Use of Elliptic Curves in Cryptography”. In: *Advances in Cryptology — CRYPTO ’85 Proceedings*. Ed. by Hugh C. Williams. Berlin, Heidelberg: Springer Berlin Heidelberg, 1986.

[MM17] Gary McGuire and Daniela Mueller. *A New Index Calculus Algorithm for the Elliptic Curve Discrete Logarithm Problem and Summation Polynomial Evaluation*. Cryptology ePrint Archive, Report 2017/1262. `https://eprint.iacr.org/2017/1262`. 2017.

[Ngu11] Kim Nguyen. “Index Calculus Method”. In: *Encyclopedia of Cryptography and Security*. Ed. by Henk C. A. van Tilborg and Sushil Jajodia. Boston, MA: Springer US, 2011, pp. 597–600. ISBN: 978-1-4419-5906-5.

[Pom08] Carl Pomerance. “Smooth numbers and the quadratic sieve”. In: *Algorithmic Number Theory: Lattices, Number Fields, Curves and Cryptography*. Vol. 44. MSRI Book Series. Cambridge University Press, 2008, pp. 69–81.

[Sil00] Joseph H. Silverman. “The Xedni Calculus and the Elliptic Curve Discrete Logarithm Problem”. In: *Des. Codes Cryptography* 20.1 (2000). ISSN: 0925-1022. DOI: `10.1023/A:1008319518035`.

[Sil09] Joseph Silverman. *The Arithmetic of Elliptic Curves*. Vol. 106. Jan. 2009. DOI: `10.1007/978-0-387-09494-6`.

[SS98] Joseph H. Silverman and Joe Suzuki. “Elliptic Curve Discrete Logarithms and the Index Calculus”. In: *Advances in Cryptology — ASIACRYPT’98*. Ed. by Kazuo Ohta and Dingyi Pei. Berlin, Heidelberg: Springer Berlin Heidelberg, 1998, pp. 110–125. ISBN: 978-3-540-49649-6.

[ST15] Joseph H. Silverman and John T. Tate. *Rational Points on Elliptic Curves*. 2nd. Springer Publishing Company, Incorporated, 2015. ISBN: 331918587X.

[Sut19] Andrew Sutherland. *Lecture notes in 18.783 – Elliptic Curves*. 2019.

[Wer13] A. Werner. *Elliptische Kurven in der Kryptographie*. Springer-Lehrbuch. Springer Berlin Heidelberg, 2013. ISBN: 9783642563515.