

ELLIPTIC CURVES AND CRYPTOGRAPHY

WALTER MOREIRA

INTRODUCTION

The main goal of this work is to present the basic definitions and properties of the elliptic curves, in order to understand why they are a relevant object in the field of public key cryptography.

The amount of work on this field is immense, a Google search of the string

"elliptic curves" cryptography

results in more than 30.000 hits, ranging from articles in mathematics and computer science to books and implementations and applications of the algorithms. There are a lot of different algorithms and also different ways to attack them, many of them involving a huge amount of theory. Moreover, the fact that still there is not an accepted standard on elliptic curve cryptography benefits the existence of a myriad of different implementations.

Thus, the final destination of this work is to just introduce the necessary terms and definitions to understand a very concrete example of public key cryptography using elliptic curves. We also want to sketch some of the possible vulnerabilities and ways to break the system, which is in fact what motivates most of the research on the field. Naturally, there are a lot of important points that are not touched here, and the general tone of the exposition is somewhat colloquial. However, the information is enough to actually understand and execute successfully a real-world cryptographic example.

The outline of the exposition is as follows. The first and second sections are the general introduction to elliptic curves over arbitrary fields. Most of the proofs are not done, but some of them are sketched. We are primarily concerned with introducing the group structure on the points of an elliptic curve. Section 3 presents some results in the particular case of a finite field, which is the case used in the practice. Section 4 shows some generalities about public key cryptography and then the particular case about using elliptic curves. Finally, the last session is an example of signature using a free and open source implementation.

1. ELLIPTIC CURVES

There are several ways to introduce the elliptic curves and to define the abelian group structure on its points. We saw in class an analytic way, by using elliptic functions and

Date: December 6, 2004.

inducing the group structure from quotients of the complex numbers by lattices. Just to see a different approach, we follow here a more algebraic line. Since our main goal is the use of the elliptic curves in cryptography, it is important to note that we need to be able to define these objects and its group structure over arbitrary fields, in particular over finite fields, including characteristic 2 and 3. This is a natural condition, arising from the fact that it is necessary to use an “exact” and efficient arithmetic on the algorithms. For example, even when it is possible to compute over \mathbb{Q} , it is difficult to have control over the size of the numerator and denominator, yielding practical problems. There are other reasons involving security arguments which makes \mathbb{Q} not suitable for the cryptographic algorithms.

We will follow a road similar to the one in [Eng99]. Let \mathbb{k} be an algebraically closed field and let's consider the projective plane $\mathbb{P}_2(\mathbb{k})$. We will remove the algebraic closure condition later. A (*projective*) *elliptic curve* E is a non-singular curve on $\mathbb{P}_2(\mathbb{k})$ given by the cubic equation

$$Y^2Z + a_1XYZ + a_3YZ^2 = X^3 + a_2X^2Z + a_4XZ^2 + a_6Z^3.$$

The mnemotechnic rule for the subindices in the coefficients is that x has weight 2, y has weight 3, z has weight 0, and a_i has weight i , with total weight 6 in each monomial. This polynomial is called *Weierstrass equation*.

There are several invariants associated to an elliptic curve E . The invariants we are going to use are Δ , called *discriminant*; and j , properly called j -invariant; defined by

$$\begin{aligned} b_2 &= a_1^2 + 4a_2 \\ b_4 &= 2a_4 + a_1a_3 \\ b_6 &= a_3^2 + 4a_6 \\ b_8 &= a_1^2a_6 + 4a_2a_6 - a_1a_3a_4 + a_2a_3^2 - a_4^2 \\ c_4 &= b_2^2 - 24b_4 \\ \Delta &= -b_2^2b_8 - 8b_4^3 - 27b_6^2 + 9b_2b_4b_6 \\ j &= c_4^3/\Delta, \text{ for } \Delta \neq 0. \end{aligned}$$

As usual in mathematics, we are not only interested in the objects themselves, but also in the equivalences and transformations between them. In this case we consider equivalent two elliptic curves E and E' related by the change of coordinates

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \mapsto \begin{pmatrix} u^2X + rZ \\ u^3Y + u^2sX + tZ \\ Z \end{pmatrix} \quad (1)$$

where u, r, s and t are in \mathbb{k} and $u \neq 0$. We will see in section 2 the reason for considering this change of variable as an equivalence between curves. We also say that E and E' are isomorphic.

Under this equivalence, the discriminant of E changes by the multiplication of u^{-12} and j remains invariant. Moreover, it can be proved that two elliptic curves with the same j -invariant are isomorphic.

Using this notion of isomorphism we can find simplified forms for the Weierstrass equation of an elliptic curve E . We list the different cases according to the characteristic here:

$$\begin{aligned}
 \text{char}(\mathbb{k}) \neq 2 \text{ or } 3 : & \quad Y^2Z = X^3 + a_4XZ^2 + a_6Z^3, \\
 \text{char}(\mathbb{k}) = 3, j \neq 0 : & \quad Y^2Z = X^3 + a_2X^2Z + a_6Z^3, \\
 \text{char}(\mathbb{k}) = 3, j = 0 : & \quad Y^2Z = X^3 + a_4XZ^2 + a_6Z^3, \\
 \text{char}(\mathbb{k}) = 2, j \neq 0 : & \quad Y^2Z + XYZ = X^3 + a_2X^2Z + a_6Z^3, \\
 \text{char}(\mathbb{k}) = 2, j = 0 : & \quad Y^2Z + a_3YZ^2 = X^3 + a_4XZ^2 + a_6Z^3.
 \end{aligned} \tag{2}$$

The procedure to get these canonical forms is straightforward. Note that the coefficients in these forms are different from the original ones, but can easily be computed from the change of variable.

The condition on non-singularity can be formulated in terms of the discriminant of the Weierstrass equation as shown in the following theorem.

Theorem 1.1. *The Weierstrass equation is irreducible and is singular if and only if its discriminant is 0.*

Sketch of the proof. There are two cases to consider. When $\text{char}(\mathbb{k}) \neq 2$, then the discriminant Δ is proportional to the discriminant of the polynomial $X^3 + a_2X^2 + a_4X + a_6$, and the derivative of this polynomial vanishes precisely when there are multiple roots, which is the condition tested by its discriminant. For the case $\text{char}(\mathbb{k}) = 2$ the invariant Δ reduces to a_6 or a_3^4 , according to whether $j \neq 0$ or $j = 0$, and one can test that the derivatives do not vanish directly.

The proof of the irreducibility consists in consider the affine case first, as a polynomial in $\mathbb{k}(X)[Y]$. Under the assumption of reducibility we can factor the Weierstrass equation in two linear factors and easily get a contradiction by counting the degrees. These two arguments remain true under homogeneization and this proves the theorem. \square

This theorem is what justifies the name ‘‘curve’’. There are other equivalent definitions in terms of geometry. For example, an elliptic curve can be defined as a complete non-singular curve of genus 1, and from there the Weierstrass equation can be deduced. In [Mil96] several of these definitions are introduced and proved equivalents.

2. GROUP STRUCTURE ON THE ELLIPTIC CURVES

In the same way as with the definition of an elliptic curve, there are many ways to introduce the abelian group structure on its points. The geometric way to define the addition of points is very suggestive, but it poses the problem of proving that the operation is associative. We will show however the geometric picture and we will show

an sketch of the formal path to prove that it is an associative commutative product. Finally we will prove that they give rise to the same operation.

Observe that the point at infinity $\mathcal{O} = (0 : 1 : 0) \in \mathbb{P}_2(\mathbb{k})$ always belong to any elliptic curve E . We choose \mathcal{O} as a distinguished point, which will be the 0 of the abelian group of points. Consider the elliptic curve E given by

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

We will compute a concrete encryption example using this curve in section 5, over the field $\mathbb{F}_{2^{163}}$. It is in normal form according to the table (2), and its invariants are $\Delta = 1$ and $j = 1$. In the next figure we show its graph and two points P and Q .

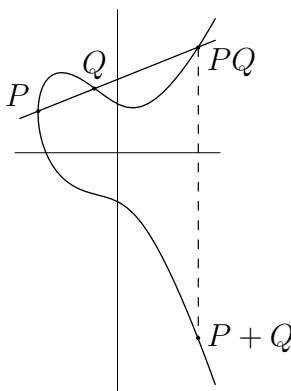


FIGURE 1

By the Bezout's theorem in the projective plane, two curves C and C' with no common irreducible components must intersect in $(\deg C)(\deg C')$ points, counting multiplicities. In the case of a line through the points P and Q this gives three points of intersection with the cubic; we call PQ to the third point. Then we consider the line through PQ and \mathcal{O} , which is just a vertical line in the affine picture passing by PQ , and its third intersection is the point $P + Q$.

The intersections must be interpreted properly when the multiplicities are greater than 1. If $P = Q$, we define PQ as the intersection of the tangent line at P with the cubic. If the line through P and Q is tangent to the cubic with define PQ as the point of tangency. And if $P = Q$ is the inflection point, then $PQ = P$. All this cases make the point \mathcal{O} the identity of the operation $+$.

There is nothing special about the point \mathcal{O} . We could choose another point on the curve and perform the previous construction, and we would get an isomorphic group structure. The choice of \mathcal{O} as zero just makes the construction of $P + Q$ from PQ very easy, since it is a reflection over the curve with respect to the x -axis.

Now we concentrate in the formal construction of the group structure. We first define the *field of rational functions* on an elliptic curve E as the subfield of the field

of fractions of the coordinate ring

$$\mathbb{k}[E] = \frac{\mathbb{k}[X, Y, Z]}{(E)},$$

consisting of the quotients of homogeneous elements with the same degree. We denote it by $\mathbb{k}(E)$. The *local ring of E at a point P* is the subring of $\mathbb{k}(E)$ formed by quotients f/g where $g(P) \neq 0$, and it is denoted by $\mathcal{O}_P(E)$. This is the beginning of an scheme-theoretic treatment of the algebraic curves, as done in [EH00]. We will not follow that road here, though.

Now we can give a reason for the equivalence of curves defined in section 1. It is easy to see that the change of variables (1) in $\mathbb{k}[X, Y, Z]$ descends to the quotient by the ideals (E) and (E') , yielding an isomorphism φ between $\mathbb{k}[E]$ and $\mathbb{k}[E']$. But moreover, since it sends a function which does not vanish in P to a function which does not vanish in P' , the image of P by the change of variables, it also induces an isomorphism

$$\varphi : \mathcal{O}_P(E) \rightarrow \mathcal{O}_{P'}(E').$$

This suggests that the change of variable is the “correct” notion of isomorphism, since it preserves the local rings.

It can also be proved that the local ring $\mathcal{O}_P(E)$ is a discrete valuation ring. This gives us an order function which we call ord , namely, for $f \in \mathbb{k}[E]$ with f not null, $\text{ord}_P(f)$ is the integer d such that $f = s^d u$, with u an unit and s belonging to the maximal ideal of $\mathcal{O}_P(E)$. The order function is naturally extended to the field of fractions $\mathbb{k}(E)$. We say that a function has a *zero at P* when the order at P is positive, and we say that it has *pole at P* when the order is negative. The absolute value of the order is the *multiplicity* of the zero or pole. The main property of this concepts is the following theorem:

Theorem 2.1. *A function $f \in \mathbb{k}(E)$ has as many zeros as poles, counting multiplicities, and there is a finite number of them.*

We define the *group of divisors of E* as the free abelian group generated by the points of E :

$$\text{Div}(E) = \langle \langle P \rangle \mid P \in E \rangle_{\mathbf{AbGrp}}.$$

The *degree* of a divisor $D \in \text{Div}(E)$ is the image of D under the evaluation map $\langle P \rangle \mapsto 1$, that is, just the sum of the coefficients of the formal sum. We denote by $\text{Div}^0(E)$ the subgroup of elements of $\text{Div}(E)$ of degree 0. The isomorphism given by the change of variable between two equivalent elliptic curves E and E' can be extended naturally to an isomorphism of groups

$$\varphi : \text{Div}(E) \rightarrow \text{Div}(E').$$

There is an important subgroup of $\text{Div}(E)$ which consists in the elements of the form

$$\text{div } r = \sum_{P \in E} \text{ord}_P(r) \langle P \rangle,$$

where $r \in \mathbb{k}(E)$. Theorem 2.1 guarantees that the sum is finite, hence $\operatorname{div} r$ is in $\operatorname{Div}(E)$, and we denote this subgroup as $\operatorname{Prin}(E)$, the group of *principal divisors*. The subgroup of elements of $\operatorname{Prin}(E)$ with degree zero is denoted as $\operatorname{Prin}^0(E)$. Actually, Theorem 2.1 also says that $\operatorname{Prin}^0(E) \subset \operatorname{Prin}(E) \subset \operatorname{Div}^0(E)$.

Finally, we define the *Picard group* and the *zero part of the Picard group* as the quotients

$$\operatorname{Pic}(E) = \frac{\operatorname{Div}(E)}{\operatorname{Prin}(E)}, \quad \operatorname{Pic}^0(E) = \frac{\operatorname{Div}^0(E)}{\operatorname{Prin}(E)},$$

respectively. Our goal is to establish a bijection between $\operatorname{Pic}^0(E)$ and the points of E .

Now, given a divisor $D \in \operatorname{Div}(E)$, we define the set

$$L(D) = \{f \in \mathbb{k}(E) \mid \operatorname{div} f + D \geq 0\} \cup \{0\},$$

where the inequality means that all the coefficients of the divisor are greater or equal than 0. For example, if $D = \langle P \rangle + 2\langle Q \rangle$, the set $L(D)$ consists of the functions which have at most one pole at P , at most a double pole at Q , and no other poles. The set $L(D)$ is a finite dimensional vector space over \mathbb{k} . We define

$$\ell(D) = \dim_{\mathbb{k}}(L(D)).$$

The following theorem is the key point to define the desired bijection. Remember that we are in the hypothesis of \mathbb{k} algebraically closed.

Theorem 2.2 (Riemann–Roch). *If g is the genus of a curve C , then we have*

$$\ell(D) \geq \deg D + 1 - g$$

for all divisors $D \in \operatorname{Div}(C)$ and the equality holds when $\deg D > 2g - 2$.

The theorem is stated for any plane curve, clearly the notion of divisors does not depend on the non-singularity of the curve. However, in the case of an elliptic curve E , we have that the genus is 1 (in fact it is one of the possible definitions, see [Mil96]), and the Riemann–Roch Theorem states that $\ell(D) = \deg D$ for any divisor with degree greater than zero.

Theorem 2.3. *Let E be an elliptic curve. The map*

$$\begin{aligned} E &\longrightarrow \operatorname{Pic}^0(E) \\ P &\longmapsto [P] - [\mathcal{O}] \end{aligned}$$

is a bijection. Here $[P]$ denote the class of $\langle P \rangle$ in the Picard group.

Proof. Clearly the image of the map belongs to $\operatorname{Pic}^0(E)$, since it has degree 0. Given an element $[D]$ of $\operatorname{Pic}^0(E)$, we have that D has degree 0 and therefore, $D + \langle \mathcal{O} \rangle$ has degree 1. According to the remark following the Riemann–Roch theorem, $\ell(D + \langle \mathcal{O} \rangle) = 1$ and thus there is only one function f , up to multiplication by a scalar, such that $\operatorname{div} f + D + \langle \mathcal{O} \rangle \geq 0$. Since the degree of a principal divisor is zero, by theorem 2.1, we conclude that $\operatorname{div} f + D + \langle \mathcal{O} \rangle$ has degree 1. But a divisor of degree 1 greater or equal

to 0 has to be of the form $\langle P \rangle$ for some point $P \in E$. And moreover, $[P] = [D] - [\mathcal{O}]$, because they differ by a principal divisor. Thus, the map of the theorem has an inverse. \square

Using Theorem 2.3 we can finally reach our goal and define the addition of two points in E as the structure which makes the bijection of the theorem a group isomorphism. A nice argument is presented in [Mil96] to relate this algebraic construction to the geometric picture of the group law. Suppose that $P + Q = S$ according to figure 1. Denote by L_1 and L_2 the lines through P and Q and through S and \mathcal{O} , respectively. From the Bezout's theorem we know that $f = L_1/L_2$ has its zeros at P , Q , and PQ , and its poles at \mathcal{O} , S , and PQ . Thus

$$\operatorname{div} f = \langle P \rangle + \langle Q \rangle + \langle PQ \rangle - \langle \mathcal{O} \rangle - \langle S \rangle - \langle PQ \rangle = \langle P \rangle + \langle Q \rangle - \langle S \rangle - \langle \mathcal{O} \rangle.$$

This implies that $[P] + [Q] - 2[\mathcal{O}] = [S] - [\mathcal{O}]$, and this is just the definition of the addition $P + Q = S$ according to theorem 2.3.

To get rid of the hypothesis of \mathbb{k} algebraically closed, we have to consider the definitions of functions and divisors over an algebraic closure \mathbb{k}^{al} of \mathbb{k} , which we will assume to be perfect. Then, we define a natural action of the Galois group $\Gamma = \operatorname{Gal}(\mathbb{k}^{\text{al}}/\mathbb{k})$ on the points of E and we still have a bijection between the points and the fixed part of $\operatorname{Pic}^0(E)$ by Γ , where E is now defined over \mathbb{k}^{al} . Hence, in what follows we will assume that we have defined the group structure on elliptic curves over arbitrary perfect fields.

There is still one point to be considered in this section of generalities. Now that the addition of points has been shown to be same to the geometrical way, we can easily compute the coordinates of points, which is important to implement the addition on the computer. The procedure is just to compute the intersections as in the figure 1. Since the only point of the line at infinity of $\mathbb{P}_2(\mathbb{k})$ which belongs to an elliptic curve is \mathcal{O} , and it is the zero of the addition, we can consider only the affine space $(x : y : 1)$ inside $\mathbb{P}_2(\mathbb{k})$. As an example, we show the coordinates of the addition for the normal form of an elliptic curve in the case $\operatorname{char} \mathbb{k} = 2$ and $j \neq 0$ (see table (2)). For $P = (x_1 : y_1 : 1)$ and $Q = (x_2 : y_2 : 1)$, we have $P + Q = (x_3 : y_3 : 1)$ with

$$x_3 = \begin{cases} \left(\frac{y_1 + y_2}{x_1 + x_2} \right)^2 + \left(\frac{y_1 + y_2}{x_1 + x_2} \right) + a_2 + x_1 + x_2, & \text{if } P \neq Q \\ x_1^2 + \frac{a_6}{x_1^2}, & \text{if } P = Q \end{cases}$$

$$y_3 = \begin{cases} \frac{y_1 + y_2}{x_1 + x_2}(x_1 + x_3) + y_1 + x_3, & \text{if } P \neq Q \\ \frac{x_1^2 + y_1}{x_1}x_3 + x_1^2 + x_3, & \text{if } P = Q \end{cases}$$

where the curve has the form $Y^2Z + XYZ = X^3 + a_2X^2Z + a_6Z^3$. Observe that after we know that these formulæ define an associative operation on the points, we could

use it to define the addition for curves on any subfield of an algebraically closed field, yielding another argument to consider arbitrary fields.

3. ELLIPTIC CURVES OVER FINITE FIELDS

In this section we work with finite fields $\mathbb{k} = \mathbb{F}_{p^k}$ of characteristic p . The main goal of this section is to present some results about the cardinality of the group of rational points of an elliptic curve. We will see in section 4.2 that the number of elements has a great impact on the security of the system.

Let E an elliptic curve with coefficients in $\mathbb{k} = \mathbb{F}_{p^k}$. We denote also by E the points of the curve in an algebraic closure of \mathbb{k} , and we denote by $E_{\mathbb{k}}$ the k -rational points of E . Clearly $E_{\mathbb{k}}$ is a finite group. The main theorems we present are the Hasse's Theorem and the structure of the torsion subgroup of E .

We call α a *rational map* on the curve E if α is a pair of functions in $\mathbb{k}(E) \times \mathbb{k}(E)$ such that $\alpha(P)$ belongs to E for all P in E . So, a rational map can be viewed as a point in $E_{\mathbb{k}(E)}$, in affine notation.

The rational maps form a group, since they are points in an elliptic curve. It can be proved that the addition in the curve corresponds just to pointwise addition as functions. We need to consider particular rational maps, those who are also group homomorphisms. Let's denote by $\text{End}(E)$ the set of these maps. This set has more structure than the abelian group structure mentioned before, it is in fact a ring with multiplication the composition of functions.

In the ring $\text{End}(E)$ there is an important rational map: the multiplication by an integer. Let $m \in \mathbb{Z}$ and write

$$[m] : E \rightarrow E$$

the map which sends P to mP , the action of \mathbb{Z} in the abelian group of points. This map induces an action of \mathbb{Z} in the ring $\text{End}(E)$ and, thus, we have the following theorem:

Theorem 3.1. *For an elliptic curve E , the set of rational maps $\text{End}(E)$ is a ring and a \mathbb{Z} -module, under the multiplication with the map $[m]$.*

The kernel of the multiplication map $[m]$ is called the *m -torsion points* of E , and it is denoted by $E[m]$. Even when E may be infinite, the subgroup $E[m]$ is finite for any integer $m \neq 0$. This generates the natural question of studying the structure of $E[m]$ as a finitely generated abelian group. The main theorem is:

Theorem 3.2. *For a positive integer $m = p^{\nu} m'$, with p the characteristic of \mathbb{F}_q and $p \nmid m'$, we have $E[m] \cong \mathbb{Z}_{m'} \times \mathbb{Z}_{m'}$ if $E[p] = 0$, and $E[m] \cong \mathbb{Z}_{m'} \times \mathbb{Z}_m$ otherwise.*

We will not prove the Hasse's theorem, but the previous statement is an important ingredient for its proof. Let's formulate it, though. There is a particular rational map which is called the *Frobenius endomorphism* $\varphi = (X^q, Y^q)$, in a field of q elements. The verification that it is an rational map is trivial, using that $u \mapsto u^p$ is an homomorphism in \mathbb{F}_q .

Theorem 3.3 (Hasse). *Let $t = q + 1 - |E_{\mathbb{F}_q}|$. Then:*

- (1) *The Frobenius endomorphism verifies $\varphi^2 - [t]\varphi + [q] = 0$, in $\text{End}(E)$.*
- (2) *$|t| \leq 2\sqrt{q}$.*

The main importance of this theorem in cryptography is that it gives a bound for the number of elements of the group $E_{\mathbb{F}_q}$. We will see later why this is important to the security of the cryptosystems. Observe that part (2) is an inequality, the actual cardinality in general is difficult to compute. There are some special cases in which this order can be computed, for example, in the case of a finite extension of \mathbb{k} and if $|E_{\mathbb{k}}|$ is known. Other cases are the *supersingular curves*, although fixed and special families of curves are usually discouraged in cryptography.

Finally, let's mention a theorem about the structure of the group $E_{\mathbb{k}}$:

Theorem 3.4 (Rück). *Let E an elliptic curve over the finite field \mathbb{F}_q . Then*

$$E_{\mathbb{F}_q} \cong \mathbb{Z}_{n_1} \times \mathbb{Z}_{n_2}$$

where $n_1|n_2$ and $n_1|q - 1$.

4. PUBLIC KEY CRYPTOGRAPHY

4.1. Generalities. In this subsection we treat the general problem of public key cryptography, and in the next subsection we specialize some of the algorithms using elliptic curves.

The general problem of cryptography is the transmission of information via insecure channels. We want to be able to transmit a message which can only be understood by the selected receiver, even when an intruder could intercept the message. The traditional way to do this is by encoding the information with a certain procedure and by interchanging the inverse procedure between the persons who wishes to communicate. The drawback of this system is widely known in these days, namely, the two persons must have a way to *securely* interchange its inverse functions, which in some scenarios may be not be possible, or at least very expensive.

The solution proposed by Diffie and Hellman [DH76] consists in the notion of *trap-door one-way functions*. The basic idea is as follows. Consider the two usual characters of the crypto-world: Alice and Bob. Imagine that Bob has, for every key k , a function f_k from which is very difficult to compute the inverse f_k^{-1} . Then, Bob could send the function f_k to Alice over an insecure channel (he can even publish the function f_k , hence the name *public key*). Alice could apply the function f_k to the message she wish to send to Bob, and then send the result to Bob also via an insecure or public channel. Finally, using the key k Bob can compute an inverse to f_k (hence the name *trap-door*) and decrypt the message from Alice. Since all these transactions can be done over public channels, the cost of the communications is extremely low.

The key point in the described situation is that the function f_k should be “difficult” to invert. The word “difficult” is formalized using complexity theory. An algorithm which can be computed in polynomial time is *easy*, while exponential time algorithms

are *hard*. Of course, there are other possibilities in-between, and actually that is what makes the elliptic curve methods currently better than the integer factorization and discrete logarithm over finite fields.

It should be noted, although it is widely known, that all these methods are *believed* to be difficult. It has not been proved that there do not exist polynomial time algorithms for these problems. In fact, it is related to one of the greatest unsolved problems: $P \neq NP$.

It is important to note, also, that cryptography is concerned with another problem, together with the encryption/decryption: the *signature problem*. It consists in the possibility to certify that a certain person is the author of a given message. Most of the algorithms used nowadays have an involution property which allows Bob to use the key k to sign a message, and Alice can use the public function f_k to verify that the signature comes from Bob. For signing a message it is enough to compute a hash function of it and then encrypt the hash with the private key. Later, with the public key any person can decrypt the signature and compare it with his own computation of the hash.

We saw in class the RSA algorithm, which relies in the difficulty of factorizing integers. We will describe here the discrete logarithm problem and other algorithms for encryption and signature which can be used later with elliptic curves.

Let G be a finite abelian group, written additively. Given two elements α and β in G , the *discrete logarithm problem* consists in finding an integer k (when exists) such that $k\alpha = \beta$. It is clearly necessary to be able to efficiently add elements of the group G and also to test for equality. Given these conditions, the difficulty of the discrete logarithm depends on the order and structure of the group.

The most usual algorithm based on the discrete logarithm is the ElGamal cryptosystem [ElG85]. It is usually used with the discrete logarithm over the multiplicative group of a finite field \mathbb{F}_q^* , but we formulate it for an arbitrary abelian group G :

- Let G be the plaintext space, that is, the message is encoded as a sequence of elements of G .
- Let $G \times G$ be the cryptospace, that is, the encrypted message is a sequence of pairs of elements of G .
- Choose an element $\alpha \in G$ and an integer a such that $1 \leq a \leq \text{ord}_G \alpha$. Let $\beta = a\alpha$.
- The values α and β constitute the public key. The value a is the secret key (say Alice's secret key). Of course, the group G must also be agreed between Alice and Bob, but it can be done in a public way.
- Bob chooses a random integer k and sends to Alice the pairs $(k\alpha, x + k\beta)$, where $x \in G$ is the encoded message.
- Alice applies the function $(c_1, c_2) \mapsto c_2 - ac_1$ and retrieves x .

Given that the computation of the discrete logarithm is believed to be hard, an intruder cannot deduce x in polynomial time from the pair $(k\alpha, x + k\beta)$. In the case of $G = \mathbb{F}_p$, the first step of the algorithm is

- Pick a large prime p , so that the discrete logarithm problem on \mathbb{F}_p^* is difficult.

and then the prime p is part of the public key. In the general case, the selection of the group G is critical for the security of the system. For example, the discrete logarithm is easy over $\mathbb{Z}/n\mathbb{Z}$ and some attacks are successful over other groups.

For completeness we show also the signature scheme of ElGamal, since we are going to show a practical example of signature. In this case Alice wants to send to Bob a signed message which can be verified to effectively belong to her. Assume that $g : G \rightarrow \{1, \dots, |G| - 1\}$ is an efficiently computable bijection.

- Alice choose an integer k , coprime with the order of the group G , and compute $\gamma = k\alpha$.
- She solves the congruence in s :

$$g(m) \equiv ag(\gamma) + ks \pmod{|G|}, \tag{3}$$

where m is the message (remember that a is Alice's secret key). Observe that it has a unique solution since k and $|G|$ are coprime.

- The signature pair is (r, s) .
- Bob computes $g(m)\alpha$ and $g(\gamma)\beta + s\gamma$, where $\beta = a\alpha$ is Alice's public key. They should be equal since

$$g(\gamma)\beta + s\gamma = (ag(\gamma))\alpha - s(k\alpha) + s\gamma = g(m)\alpha.$$

If they coincide, Bob can be fairly sure that the message was signed by Alice.

One possible method to try to solve the discrete logarithm problem is by testing all the values $2\alpha, 3\alpha, \dots$ until getting β . This method clearly has the complexity proportional to the order of the cyclic group generated by α , hence exponential complexity on the number of bits of the key. However, there are other possible attacks to the discrete logarithm. We mention some of them.

The first improvement to the brute force testing of all the values $i\alpha$ is the *baby-step giant-step algorithm* proposed by Shanks [Sha71]. The basic idea is to compute a list of values $i\alpha$ for i less than some integer b and to order them according to some linear order. These are the baby-steps. Then, start computing $ja, 2ja, 3ja, \dots$ (the giant-steps) for some suitable integer j and test whether they lie in the list of baby-steps. Since that list is ordered, one can perform a binary search and get a better algorithm than the naïve one. However, this attack still have exponential complexity, and the main reason is that it is too generic, it works for any group.

Another method which achieves a sub-exponential complexity in the case $G = \mathbb{F}_q^*$ is the *index calculus method*. We write now the group G multiplicatively. Select a number of elements in G , which we call a *factor base*:

$$\Gamma = \{\gamma_1, \gamma_2, \dots, \gamma_t\} \subset G.$$

Choose random integers s between 1 and the order of the group and compute α^s . Then, try to factorize the result in Γ :

$$\alpha^s = \prod_{i=1}^t \gamma_i^{s_i}.$$

When the factorization is possible and enough integers s have been found with that property, $\log_\alpha \gamma_i$ can be computed by solving the linear system in \mathbb{Z} :

$$s = \sum_{i=1}^t s_i \log_\alpha \gamma_i.$$

After we have computed the logarithms of Γ , select again random integers s and try to factorize $\beta\alpha^{-s}$ in Γ . If the factorization is possible for some s we have

$$\log_\alpha \beta - s = \sum_{i=1}^t s_i \log_\alpha \gamma_i,$$

and we can use the previously computed logarithms. This method clearly depends on the existence of “good” factor bases, which is the case of the group \mathbb{F}_q^* . The main conclusion is that the discrete logarithm for the multiplicative group of finite fields is not “too” hard.

4.2. Elliptic curves in cryptography. We are now in position to comment on the use of elliptic curves in cryptography. The basic idea is to use the ElGamal cryptosystem over the group of rational points of an elliptic curve over a finite field, that is, select $G = E_{\mathbb{F}_q}$ for a suitable elliptic curve E .

The main advantage of the use of the group $E_{\mathbb{F}_q}$ is that no sub-exponential algorithm for computing the discrete logarithm is known until now. There are some arguments which suggest that the index calculus method is not possible over this group. They are roughly based in the fact that the factorizations mentioned before would be done by lifting the points of $E_{\mathbb{F}_p}$ to $E_{\mathbb{Z}_{(p)}}$, where $\mathbb{Z}_{(p)}$ is the set of rationals a/b with $p \nmid b$. But then the efficiency of the method would depend on the size of a and b , and it can easily be shown that the coordinates of nP grows quadratically with n . This suggests that the index method cannot run in sub-exponential time over the group of a generic elliptic curve.

Since the difficulty of the problem over the group $E_{\mathbb{F}_q}$ is greater than over \mathbb{F}_q^* , we can use smaller keys to achieve the same degree of security. For example, the usual size of key on the ElGamal algorithm which is used in the PGP or GPG implementations is at least of 1024 bits. With respect to elliptic curve cryptography, the last broken key has 109 bits, using parallel computations over thousands of machines. Keys of lengths 163 and greater are not expected to be broken in a near future, hence we can use keys of an order of magnitude smaller.

However, it is necessary to have some caution when choosing the elliptic curves, since there are special cases in which some of the algorithms succeed in producing sub-exponentials running times. For example, when the order of the group is not prime, Pohlig and Hellman [DH76] proposed a method to reduce the discrete logarithm to special Sylow subgroups of G . This is what justify the importance of knowing the order of the group $E_{\mathbb{F}_q}$, but as we pointed after Theorem 3.3, it is difficult to compute in the generic case.

Another case in which the attacks are successful is the supersingular case. An elliptic curve is *supersingular* if the ring $\text{End}(E)$ is non-commutative. In this situation, Menezes, Okamoto and Vanstone [MOV93] proposed an algorithm based on the Weil pairing. The Weil pairing is a particular bilinear form

$$e_n : E[n] \times E[n] \rightarrow \overline{\mathbb{F}_q}^*$$

with the property that if $E[n] \subset \mathbb{F}_{q^k}$, then $e_n(P_1, P_2) \in \mathbb{F}_{q^k}$ for all $P_1, P_2 \in E[n]$ (here $\overline{\mathbb{F}_q}$ is the algebraic closure of \mathbb{F}_q). The following theorem is the main tool:

Theorem 4.1. *For a basis $\{P, Q\}$ of $E[m]$ (remember Theorem 3.4) as \mathbb{Z}_m -module, $e_m(P, Q)$ is a primitive m -th root of unity.*

Let P a point of order n in $E_{\mathbb{F}_q}$ and $R = \ell P$. With this theorem and notations, the MOV algorithm is:

- Choose the minimum k such that $E[n] \subset E_{\mathbb{F}_{q^k}}$.
- Compute a point Q such that $e_n(P, Q)$ is a n -th root of unity.
- Let $\beta = e_n(R, Q)$.
- Then, $\ell = \log_\alpha \beta$ in \mathbb{F}_{q^k} .

Thus, we have reduced the discrete logarithm over the group of an elliptic curve to the problem over a finite field, in which there are sub-exponential algorithms. Observe that, by linearity,

$$\beta = e_n(\ell P, Q) = e_n(P, Q)^\ell = \alpha^\ell,$$

which proves the last step of the algorithm. For most types of curves, finding the integer k of the algorithm is not possible, but for supersingular curves k is small and can be found in probabilistic polynomial time.

As an example, in the case of characteristic 2 or 3, a curve E is supersingular if and only if the j -invariant is 0. This shows that the curve of figure 1 is not supersingular, since $j = 1$.

5. PRACTICAL EXAMPLE

In this section we present an example of use of a free and open source implementation of the algorithms using elliptic curves. The software can be found in

<http://users.pandora.be/stes/ecc.html>

and is due to David Stes. The application is neatly written in Objective-C and the algorithms are very readable, thus it was possible to verify that the program is using exactly the same ElGamal algorithm we presented here.

First of all, we have to generate the public and private pair of keys, by selecting an elliptic curve and two points on it. The application has several different curves over different fields. We select the curve of figure 1, which we know is not supersingular, as we say in the previous section. Some information is displayed by the application:

```

1  walterm@hobbes ecc-0.2.2 $ ec-keygen -i sect163k1
2  Nickname : Sect163k1
3  Description : SECG curve over finite field with 2^163 elements
4  Minimal polynomial : w^163 + w^7 + w^6 + w^3 + 1 == 0
5  Curve : Y^2 Z + X Y Z == X^3 + X^2 Z + Z^3
6  Characteristic prime field : 2
7  Discriminant : 1
8  j-Invariant : 1
9  Order of basepoint : 5846006549323611672814741753598448348329118574063
10 Basepoint : (w^161 + w^159 + w^158 + w^157 + w^156 + w^155 + ...
11 Multiplying basepoint by order ...
12 Resulting point is : (0,1,0) (verified on curve)
13 Equal to point at infinity : YES
14 Order of basepoint is probably prime (Rabin Miller test).
```

The curve with nickname Sect163k1 is the curve displayed in line 5 over the field $\mathbb{F}_{2^{163}}$, whose minimal polynomial is shown in line 4. The base point α of the ElGamal algorithm is shown in line 10 and it is truncated since it takes several lines to print it. The last lines are a verification that $(\text{ord}_{E_G} \alpha)\alpha = \mathcal{O}$, for $G = E_{\mathbb{F}_{2^{163}}}$. Note that it also verifies that the order of the point is pseudo-prime, to avoid some of the attacks mentioned before.

Then we generate the keys over this curve:

```
walterm@hobbes ecc-0.2.2 $ ec-keygen -e sect163k1 -o key
```

Two files `key.pub` and `key.prv` are created. Now we generate a file with with name `msg` and the content `Hello, world`, and we use the ElGamal algorithm with the private key to sign the file:

```

1  walterm@hobbes ecc-0.2.2 $ ec-sign -v -G -c sha1 -k key.prv
2                                     -f msg -s msg.sign
3  7b4758d4baa20873585b9597c7cb9ace2d690ab8          sha1 digest
4  Message: 703796955989529448569322802441099278483500632760
5  Private key: 1281854678590587452193278074179895231173252807099
6  Random number, relative prime to order of basepoint:
7     1166401717724322665623331734394570904071554327546
8  Bezout inverse: -1385274205420100050396760535772411461052840496545
9  Random point on curve: (w^161 + w^160 + w^158 + w^155 + w^154 + ...
```

```

10 ElGamal signature: 2570653213999036177680740713879861118148222510293
11 ec-sign: writing sect163k1 signature to 'msg.sign'

```

Here we are signing the message using the SHA1 algorithm (that we did not mention) to generate a hash number. In line 5 it is displayed the private key, that is, the integer a from ElGamal algorithm. The Bezout inverse in line 8 is the solution of the congruence (3). The random point in line 9 is the point β . These two objects are saved in a file `msg.sign` which can be used to test the correctness of the signature, using the public key:

```

1 walterm@hobbes ecc-0.2.2 $ ec-verify -k key.pub -f msg -s msg.sign
2 walterm@hobbes ecc-0.2.2 $ echo $?
3 0

```

The number 0 in line 3 indicates that the verification was successful.

REFERENCES

- [ASM98] K. Araki, T. Satoh, and S. Miura. Overview of elliptic curve cryptography. In *Public Key Cryptography*, volume 1431 of *LNCSS*, pages 29–48. Springer-Verlag, Berlin, 1998.
- [DH76] Whitfield Diffie and Martin E. Hellman. New directions in cryptography. *IEEE Trans. Information Theory*, IT-22(6):644–654, 1976.
- [EH00] David Eisenbud and Joe Harris. *The geometry of schemes*, volume 197 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 2000.
- [ElG85] Taher ElGamal. A public key cryptosystem and a signature scheme based on discrete logarithms. In *Advances in cryptology (Santa Barbara, Calif., 1984)*, volume 196 of *Lecture Notes in Comput. Sci.*, pages 10–18. Springer, Berlin, 1985.
- [Eng99] Andreas Enge. *Elliptic curves and their applications to cryptography*. Kluwer Academic Publishers, Netherlands, 1999. An introduction.
- [Hus87] Dale Husemoller. *Elliptic curves*, volume 111 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1987. With an appendix by Ruth Lawrence.
- [Kob84] Neal Koblitz. *Introduction to elliptic curves and modular forms*, volume 97 of *Graduate Texts in Mathematics*. Springer-Verlag, New York, 1984.
- [Mil96] J. S. Milne. Elliptic curves. Notes in <http://www.jmilne.org/math/CourseNotes/math679.html>, August 1996.
- [MM97] Henry McKean and Victor Moll. *Elliptic curves*. Cambridge University Press, Cambridge, 1997. Function theory, geometry, arithmetic.
- [MOV93] Alfred J. Menezes, Tatsuaki Okamoto, and Scott A. Vanstone. Reducing elliptic curve logarithms to logarithms in a finite field. *IEEE Trans. Inform. Theory*, 39(5):1639–1646, 1993.
- [Sha71] Daniel Shanks. Class number, a theory of factorization, and genera. In *1969 Number Theory Institute (Proc. Sympos. Pure Math., Vol. XX, State Univ. New York, Stony Brook, N.Y., 1969)*, pages 415–440. Amer. Math. Soc., Providence, R.I., 1971.