Andrzej Wilkowski  
(Wrocław)

ELLiptic CURves AND THEIr USES  
IN INTERNET SECURITY

Abstract. In this paper we present how to use elliptic curves in safe data circulation on the Internet. We also talk about adding points of elliptic curves. Groups of these points are the basis to construct cryptosystems which have shorter keys than present ones. Finally, we discuss the ways to protect the Internet user from the criminal activities known as phishing.

Key words: antiphishing filter, ElGamal system for elliptic curve, elliptic curves, phishing.

1. Elliptic curves are widely and extensively documented in professional literature. As put by Serge Lang, in preface to his book: “it is possible to write endlessly on elliptic curves” (S. Lang (1978)). Theory of elliptic curves, a jewel of 19th century mathematics, was developed by Abel, Gauss, Jacobi and Legendre. The curves are a natural companion of numerous subfields of mathematics, and their name is derived from the closely related theory of elliptic functions. One of the most recent applications of elliptic curves is the proof to Fermat’s Last Theorem (A. Wiles (1995)) (still considered a valid approach to proving the famous conjecture, although one that Fermat himself would probably disapprove of, as lengthy and too elaborate). Contrary to common perception, elliptical curves are not ellipsoids, but non-singular third-degree curves. What is, then, the origin of the elliptic label? As we know, arc length of an ellipse can be naturally expressed as an integral of a fourth-degree polynomial square root function. Such an integral may, obviously, be expressed as an integral of a third-degree polynomial square root function, commonly referred to as elliptic integral. Hence, the adjective elliptic refers to the theory of elliptic integrals, i.e. integrals that are expressed as

\[ \int R(x, y)dx, \]

where \( R(x, y) \) is a rational function of two variables, and \( y^2 \) is a polynomial of degree 3 (or 4) of \( x \) variable, with no repeated roots. Such integrals, on complex plane, are multiple valued functions, and are defined only as modulo lattice of periods. Thus, it may be assumed that values of elliptic integrals fall on the surface of a torus. The reverse of elliptic integral is a 2-period function referred to as
elliptic function (any 2-period meromorphic function on complex plane may be derived in this way). Elliptic curves are parameterized by the above functions.

From ca. 1985 onwards, the theory of elliptic curves over finite fields has been utilized in solving many cryptographic problems, such as prime factorization of natural numbers, primality tests, and construction of asymmetric cryptosystems.

Groups of elliptic curve points over finite fields are similar to multiplicative groups on finite fields. However, they are superior in two ways: firstly, they are greater in number, and secondly, they seem to provide similar level of security at a limited length of keys. For example, the conventional RSA cryptosystems, commonly used on the Internet, typically use keys of 1024 or 4096 byte length. This level of security can be obtained using keys of ca. 173 and 313 bytes, respectively, if the cryptosystem is based on elliptic curves (I. Blake, G. Seroussi, N. Smart (2004)), which may be a significant difference in more demanding applications (the RSA algorithm is relatively slow).

2. First, let us define the basic concept of this paper.

**Definition.** Elliptic curve $E$, over body $K$, is a set of

$$E(K) = \{(x, y) \in K^2 : y^2 = x^3 + ax + b; \quad a, b \in K\} \cup \{O_E\},$$

where $O_E$ is referred to as point in infinity, right-hand polynomial is with no repeated roots, and the $K$ body characteristic is different from 2 and from 3.

It is worth noting here that for infinite fields $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ the characteristic equals 0, while for finite fields $F_q$, of $q = p^j$ elements, or for fields $\mathbb{Z}/p\mathbb{Z}$, where $p$ is a prime number, the characteristic equals $p$. When the characteristic of field $K$ equals 2 or 3, the equation presented above varies slightly. For this reason, the respective cases will be disregarded for the purpose of this paper.

**Example 1 (S.Y. Yan (2006)).** Let $E$ be an elliptic curve $y^2 = x^3 + 3x$ over field $F_5$. In this case, the curve $E$ consists of 10 points:

$$E(F_5) = \{O_E, (0, 0), (1, 2), (1, 3), (2, 2), (2, 3), (3, 1), (3, 4), (4, 1), (4, 4)\}.$$

The table below shows the number of points for various elliptic curves over a field of $F_5$.

<table>
<thead>
<tr>
<th>Elliptic curve</th>
<th>Number of points</th>
</tr>
</thead>
<tbody>
<tr>
<td>$y^2 = x^3 + 2x$</td>
<td>2</td>
</tr>
<tr>
<td>$y^2 = x^3 + 4x + 2$</td>
<td>3</td>
</tr>
<tr>
<td>$y^2 = x^3 + x$</td>
<td>4</td>
</tr>
<tr>
<td>$y^2 = x^3 + 3x + 2$</td>
<td>5</td>
</tr>
<tr>
<td>$y^2 = x^3 + 1$</td>
<td>6</td>
</tr>
<tr>
<td>$y^2 = x^3 + 2x + 1$</td>
<td>7</td>
</tr>
<tr>
<td>$y^2 = x^3 + 4x$</td>
<td>8</td>
</tr>
<tr>
<td>$y^2 = x^3 + x + 1$</td>
<td>9</td>
</tr>
<tr>
<td>$y^2 = x^3 + 3x$</td>
<td>10</td>
</tr>
</tbody>
</table>

Source: own study based on (S.Y. Yan (2006)).
As seen in the Table 1, the number of points varies in the range of 2 to 10. For standard cases, the following estimation is valid.

**Theorem (Hasse 1933).**

\[
|E(F_p)| \leq 1 + p + 2\sqrt{p}.
\]

**Example 2.** Let us analyze an elliptic curve \(E\) over the field of real numbers \(\mathbb{R}\):

\[
E(\mathbb{R}) = \{(x, y) \in \mathbb{R}^2 : y^2 = x^3 - 4x + 2\} \cup \{O_E\}.
\]

Such curve, obviously, contains an infinite number of points on the plane. Figure 1 below shows its graphical representation.

![Graph of the elliptic curve](image)

**Fig. 1.** Elliptic curve \(y^2 = x^3 - 4x + 2\).

Source: own study.
As Figure 1 shows, any straight line non-parallel to y axis may cross the curve in three points (tangency is counted twice). For this curve, a point in infinity $O_E$ should be represented as an infinitely distant point on y axis, towards the limiting direction of increasingly steep tangentials, such point is the \textit{third point of crossing} for any vertical straight line crossing or tangent with the curve $E$.

The most basic operation on an elliptic curve is addition of points to the curve. To perform such an operation, it is convenient to utilize geometric intuition, based on Fig. 1. In such a case, the rule of adding points to the elliptic curve can be summarized as follows:

\textbf{The sum of three points, in which the straight line crosses the curve, equals $O_E$.}

The geometric rule of point addition allows to add two points on the elliptic curve in such a way that the third point can be established. To perform this operation numerically, one needs an algebraic formula. Basic algebraic formulas are shown, valid for any fields of characteristic different from 2 and 3.

Let

$$P = (x_1, y_1), \quad Q = (x_2, y_2) \in E(K) = \{ (x, y) \in K^2 : y^2 = x^3 + ax + b ; \quad a, b \in K \} \cup \{ O_E \},$$

then

$$P + Q = \begin{cases} O_E, & \text{for } x_1 = x_2 \text{ and } y_1 = -y_2, \\ (x_3, y_3), & \text{for other cases,} \end{cases}$$

where

$$(x_3, y_3) = (d^2 - x_1 - x_2, d(x_1 - x_3) - y_1) \in E(K),$$

and

$$d = \begin{cases} 3x_1^2 + a, & \text{for } P = Q, \\ \frac{y_2 - y_1}{x_2 - x_1}, & \text{for other cases.} \end{cases}$$

\textbf{Comment.} Adding points to elliptic curve $E$ leads to creation of Abelian group structure with neutral element $O_E$. In 1922, Mordell presented a proof that an Abelian group of points in any elliptic curve over a field of rational numbers $\mathbb{Q}$ is a direct sum of a finite subgroup of finite order (torsion subgroup) and a subgroup generated by finite number of points of infinite order.

This property allows us to utilize elliptic curves in cryptography. The most important works in this field are (V. Miller (1986) and N. Koblitz (1987)). From
then on, elliptic curves have been the subject of extensive studies in cryptography. Many secure encryption methods and digital signature tools were developed based on the curves, adopted for the Internet and widely used. At present elliptic curve-based cryptography is a standard application (the front runner in the field being the Canadian IT company Certicom, owner of over 130 patented solutions). Basic building blocks of a cryptosystem based on an elliptic curve $E$, over a finite field $F_q$, are summations in the form of

$$P + P + \ldots + P = kP,$$

where $P$ represents a point on curve $E$, while $k$ is an integer. As it turned out, the calculation can be performed using a recurrent doubling method, with the use of

$$O(\log_2 k(\log_2 q)^3)$$

byte operations (this is a very fast algorithm that can be used widely in computing). Security of such a cryptosystem results from the fact that with a given curve $E$, point $P$ on the curve and point $kP$ on the same curve, it is difficult to establish the value of $k$ integer. This difficulty is referred to as the elliptic curve discrete logarithm problem. At present, it is widely assumed that for a properly selected curve $E$ and field $F_q$, the solution to discrete logarithm problem of $E(F_q)$ requires computing complexity of order exponentially proportionate to the size of the field (hence, their respective algorithms have no practical use).

**Example 3.** Almost every cryptosystem used on the Internet and based on public keys has its elliptic curve counterpart. Let us examine the elliptic curve equivalent of the ElGamala cryptosystem (S.Y. Yan (2006)).

- Alice and Bob publicly disclose their choice of elliptic curve $E$ over field $F_q$, where $q = p^r$ and $p$ is a large prime number, together with a randomly selected point $P \in E$.
- Alice randomly selects an integral number $r_A$ (Alice’s private key) and computes the point $r_A P$ (Alice’s public key); Bob, similarly, randomly selects an integral number $r_B$ (Bob’s private key) and computes the point $r_B P$ – Bob’s public key (the numbers $r_A$ and $r_B$ are secret, while points $r_A P$ and $r_B P$ are publicly disclosed).
- To send Bob a message – point $M$, Alice randomly selects an integral number $k$ (secret) and sends a pair of points $(kP, M + k(r_B P))$.
- To decipher the message $M$, Bob computes $M + k(r_B P) - r_B (kP) = M$.

Bob uses a similar procedure to message Alice. Elliptic curve $E$, point $P$ and public keys of Alice and Bob are available to any Internet user. As a result, any person can send encrypted messages to both Alice and Bob. However, if the transmission is tapped, the eventual eavesdropper has to compute the discrete logarithm over curve $E$. As there is no reliable and effective method for computing those, the system described above can be regarded secure.

To conclude this part of the paper, it is noteworthy to mention the subject of quantum cryptography, a field that combines the disciplines of cryptography and
quantum mechanics. The basic tool of that field is the hypothetical quantum computer (Ch. Monroe, D. Wineland (2008)), a physical system designed in such a way that the outcome of its evolution, in line with quantum mechanics rules, represents the solution for a given computing problem. Such a system would offer “truly random” generation of random numbers (A. Mitra (2009)) or prime factorization of a natural number $N$ in the timeframe of $O(\log^3 N)$ and memory requirement of $O(\log N)$. Postulates for respective algorithms are currently available (P.W. Shor (1996)). Such a system would, in theory, pose a threat to the security of the widely used RSA cryptosystem (with its public key being the product of two prime numbers and its security based on the computing problem of prime factorization of the product). The most advanced quantum computation so far is the factorization of $15 = 3 \times 5$, accomplished in 2001 by a joint team of IBM and Stanford University. It is hard to predict the efficiency of quantum computers in respect to the elliptic curve discrete logarithm problem, so the actual threat to security of such systems cannot be studied as yet.

3. This section discusses the ways to protect the Internet user from the criminal activities known as phishing. After Wikipedia, “in the field of computer security, phishing is the criminally fraudulent process of attempting to acquire sensitive information such as usernames, passwords and credit card details by masquerading as a trustworthy entity in an electronic communication. Phishing is an example of social engineering technique. The term was coined in mid 90s by crackers exploiting AOL (a large US Internet provider). A phisher might pose as an AOL staff member and send an instant message to a potential victim, asking him to reveal his password. In order to lure the victim into giving up sensitive information the message might include imperatives like ‘verify your account’ or ‘confirm billing information’. Once the victim had revealed the password, the attacker could access and use the victim’s account for fraudulent purposes or spamming”.

In the US alone, losses incurred in 2007 by phishing are estimated at 3.2 billion USD (L.F. Cranor (2009)). Below are some common-sense advices to protect every Internet user from falling prey to phishing attempts.

- If you receive an e-mailed request to visit and log on to your service provider, verify the authenticity of the message with the administrator of the service.
- Do not open any hypertext links directly from the e-mail body text.
- If the webpage address looks suspicious, verify its authenticity using a standard search engine (fraudulent addresses will not rank high on results list).
- Make sure your system and software is updated on regular schedule.
- Do not send any sensitive information openly via e-mail (this applies to such information as passwords, credit card numbers, and so on).
- Banks and other financial institutions use HTTPS protocol on every page that requires system log-on. If the login page in question does not include HTTPS in its address, immediately contact system administration authorities and refrain from entering any data on such a page.
Consider the use of OpenDNS (a free server system and communications protocol to translate domain names meaningful to humans into the numerical (binary) identifiers associated with networking equipment).

Table 2. Basic principles of anti-phishing filter

<table>
<thead>
<tr>
<th>Heuristics</th>
<th>Suspicious parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>Domain age</td>
<td>12 months or less</td>
</tr>
<tr>
<td>Known logos</td>
<td>Page contains known logos, but the domain itself does not belong to their owners</td>
</tr>
<tr>
<td>Suspicious URL</td>
<td>URL contains such characters as @, hyphen, IP address or more than 5 dots</td>
</tr>
<tr>
<td>Suspicious link</td>
<td>Link on page contains @ or hyphen</td>
</tr>
<tr>
<td>Form</td>
<td>Page contains a separate field for text entry</td>
</tr>
<tr>
<td>Lexical analysis</td>
<td>URL does not correspond with Google-suggested address</td>
</tr>
</tbody>
</table>

Source: own research based on (L.F. Cranor (2009)).

For more on the subject, see (L.F. Cranor (2009)) and webpage: http://apwg.org/advice.

Practical skills in recognizing suspicious URLs (webpage addresses) can be verified in an online game: ANTI-PHISHING PHIL (http://cups.cs.cmu.edu/anti-phishing_phil).

The user plays a small fish that needs to decide whether a worm associated with a given address is edible or if it should be avoided. After each round, user skill is ranked and suggestions are given for improving anti-phishing skills. Apart from common-sense methods, special filtering tools are available. An efficient filter of this kind should be flexible, to allow for update of the ever-changing phishing tactics. In (L.F. Cranor (2009)), one of the most recent filtering tools is discussed, with a laboratory-test efficiency of 95% of fake webpage address recognition. Basic principles underlying this tool are presented below.

**Literature**


